What Lies Beneath: An Examination of the Underlying Components of Deception Detection

by

Elizabeth Elliott

A thesis submitted to the
School of Graduate and Postdoctoral Studies in partial
fulfillment of the requirements for the degree of

Doctor of Philosophy in Forensic Psychology

Faculty of Social Science and Humanities

University of Ontario Institute of Technology (Ontario Tech University)

Oshawa, Ontario, Canada

December 2022

Thesis Examination Information

Submitted by: **Elizabeth Elliott**

Doctor of Philosophy in **Forensic Psychology**

Thesis title: What Lies Beneath: An Examination of the Underlying Components of Deception Detection

An oral defense of this thesis took place on December 12th, 2022 in front of the following examining committee:

Examining Committee:

Chair of Examining Committee	Dr. Leigh Harkins
Research Supervisor	Dr. Amy-May Leach
Examining Committee Member	Dr. Joseph Eastwood
Examining Committee Member	Dr. Karla Emeno
University Examiner	Dr. Lindsay Malloy
External Examiner	Dr. Marguerite Ternes

The above committee determined that the thesis is acceptable in form and content and that a satisfactory knowledge of the field covered by the thesis was demonstrated by the candidate during an oral examination. A signed copy of the Certificate of Approval is available from the School of Graduate and Postdoctoral Studies.

Abstract

Little is known about the underlying components that are responsible for the processes involved with decision-making in deception detection. Throughout this dissertation, I explored previous (mis)conceptions about decision-making in deception and its detection. In Study 1, I used self-report and objective measures (i.e., coding of witness statements) to examine a long-held dichotomy in deception detection research – that those labelled "truth-tellers" are exclusively honest and those labelled "lie-tellers" fabricate their entire accounts. My findings revealed that both groups incorporated truthful and deceptive elements into their accounts, in fact. Although it has been well-established that truth-tellers provide more information than lie-tellers (DePaulo et al., 2003; Vrij, 2008), I found that truth-tellers provided more *accurate* information than lie-tellers, whereas the groups did not differ in utterances of inaccurate details. Rather, lie-tellers omitted significantly more details than truth-tellers, particularly during the free recall phase of the interview. In Study 2, I examined whether deception detection is static, as has been implied by the field's focus on post hoc decision-making (Shanks, 2017). Using a novel, dynamic approach to measuring deception detection, I found that decisions and biases changed over time. Observers in the control condition – who made decisions after viewing an interview, as is typical in the literature – were able to discern between lie-tellers and truth-tellers and they exhibited a truth-bias, replicating previous research. However, observers who rendered their decisions continuously held no biases and were insensitive to veracity. Lastly, in Study 3, I examined the effect of varying the focal element of the deception detection task (i.e., person, event, or detail) because the impact of question phrasing on veracity decisions was unexamined in previous research. There were no differences between question

phrasing conditions, which implies that phrasing is not the source of variability in discrimination or accuracy results throughout the literature. Overall, this dissertation serves to validate and challenge long-standing notions within deception detection research.

Keywords: Deception Detection; Decision-Making; Details; Interviewing; Phrasing

Author's Declaration

I hereby declare that this thesis consists of original work of which I have authored. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I authorize the University of Ontario Institute of Technology (Ontario Tech University) to lend this thesis to other institutions or individuals for the purpose of scholarly research. I further authorize University of Ontario Institute of Technology (Ontario Tech University) to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research. I understand that my thesis will be made electronically available to the public.

The research work in this thesis that was performed in compliance with the regulations of Research Ethics Board under **REB Certificate numbers: 14219, 14220, and 15794**.

_____ ELIZABETH ELLIOTT

Statement of Contributions

I hereby certify that I am the sole author of this thesis and that Chapters 2 and 3 of this thesis have been submitted for publication in *Applied Cognitive Psychology* and *Journal of Experimental Psychology: General*, respectively. I have used standard referencing practices to acknowledge ideas, research techniques, or other materials that belong to others. Furthermore, I hereby certify that I am the sole source of the creative works and/or inventive knowledge described in this thesis.

Acknowledgements

My life has completely changed over the course of the last seven years. I came into the PhD program bright eyed and confident but as any academic could probably attest, the process has completely humbled me. Throughout my PhD, I have travelled to three continents for conferences and research, was placed on bedrest after going into pre-term labor following stressful comprehensive exams, had a child and went on leave, survived a global pandemic, took another leave to try my hand at industry work, tried to publish manuscripts, and finally, moved to another country to pursue a postdoctoral position in academia. No other job in the world would grant me so much freedom and simultaneously put me through so much stress. Several important people helped me along the way, and I want to thank them here.

First of all, I want to acknowledge the support, guidance, knowledge, mentorship, and training that my supervisor, Dr. Amy-May Leach, has given me. Over the years, I have written and expressed many heartfelt 'thank yous' to Amy via emails, texts, and in person. I want her to know that I am infinitely grateful for all her time, patience, composure, and truthfully, effort. I am so lucky to have had a supervisor and mentor who truly cared about my growth. She challenged me on everything from study feasibility and methodology to manuscript writing and the presentation of results, and in turn, she has made me a better researcher and communicator. People outside of academia are seldom aware that supervisors are not paid for their mentorship, but rather devote their time and energy selflessly. And Amy, she went above and beyond. Simply put: this work would not have been possible without you, Amy. Thank you.

I am thankful for my fellow graduate students, especially Lyndsay Woolridge, Christina Connors, and Lindsay Groat, who have cheered me on and provided an ear when I need to vent or work through a problem.

Next, I want to sincerely thank my committee members for their guidance throughout the years, their time, and feedback on the final version of the dissertation.

I am also very grateful for the financial support that the Social Sciences and Humanities Research Council granted me to conduct this work.

Thank you to my family, my parents, Nataliya and Genadi, and my brother, Edward. Thank you, Edward, for paving the way for me to be curious, always encouraging me without judgment, and showing me that changing my mind is not only okay but is often beneficial. My parents have made countless sacrifices, not the least of which were two immigrations, to provide a life for my brother and I that they could only dream about achieving. They came from traditional upbringings and restrictive regimes where the prospect of studying psychology for a living was derisory. Nonetheless, they have always encouraged me to explore my options, supported me, and taken an interest in my work. Thank you, mama and papa. Слава Україні!

Lastly, and most importantly, I want to state my gratitude for my husband and best friend, Spencer, who has stood by my side and supported my career aspirations from the moment we met. He has never complained and has done nothing but encourage me to follow my dreams. I know that no matter what is next to come, he will be by my side providing much needed support (and humor). I am forever in awe of you, Spencer.

Table of Contents

List of Tables

Chapter 3

Chapter 4

List of Figures

Chapter 2

Chapter 3

List of Abbreviations and Symbols

ALIED	Adaptive Lie Detector
ANOVA	Analysis of Variance
BF_{01}	Bayes Factors, "01" subscript implies support for null hypothesis
c	A measure of response bias
CI	Confidence Interval
CIA	Central Intelligence Agency
d	A standardized measure of effect size
d′	A measure of discrimination
DOI	Digital Object Identifier
e.g.	exempli gratia, meaning "for example"
η_p^2	eta-squared, inferential statistic that measures the strength of a relationship
F	Inferential statistic that represents the ratio of explained variance to unexplained variance
FBI	Federal Bureau of Investigations
i.e	id est, meaning "that is"
ICC	Interclass Correlation
M	Mean, average
min	Minute(s)
mph	Miles per hour
N	Overall number of participants
n	Number of participants in a group
p	p-value, a measure of statistical significance
PEACE	Preparation and Planning; Engage and Explain; Account, Clarification, Challenge; Closure; Evaluation
RM	Reality Monitoring
s	Second(s)
SD	Standard Deviation
SDT	Signal Detection Theory

SUE	Strategic Use of Evidence
t	Inferential statistic used to determine the differences between two group means
ϕ	Standard normal probability density function
V	Pillai's Trace, used with uneven cell sizes
χ^2	Chi-squared, inferential statistical used to assess goodness of fit
Z_{Hit}	Standardized measure of Hits (i.e., instances when observer correctly identifies a lie as a 'lie')
$Z_{False\ Alarm}$	Standardized measure of False Alarms (i.e., instances when observer incorrectly identifies a truth as a 'lie')

Chapter 1: General Introduction

What if everything we "know" about deception detection is wrong? Deception detection research relies on the assumption that lie-tellers and truth-tellers are distinct from one another (e.g., Nahari et al., 2019). Researchers meticulously analyze and enhance nonverbal and verbal behaviors in search of differences between the groups to aid in the detection of lies, for example (for a review see Vrij et al., 2010). Although there are countless meta-analytic investigations of deception detection, the veracity question posed to participants varies in the literature (e.g., Bond & DePaulo, 2006; Hartwig & Bond, 2011; Palena et al., 2021). Furthermore, researchers almost exclusively assess deception detection post hoc. Despite some advances, observers continue to perform relatively poorly on deception detection tasks (Bond & DePaulo, 2006; Vrij et al., 2017). Why? Over a decade ago, in their seminal meta-analysis, DePaulo et al. (2003) stated that deception detection studies created a clear dichotomy between lies and truths that was not present outside of the laboratory.

1.1 What is Deception?

Researchers have yet to identify the exact composition of a lie. That is, there is no specific number of cues, behaviors, or verbal patterns that may immediately be recognized as deceit. In fact, we cannot, with certainty, determine whether a specific statement is a *lie* or whether a person is *lying* (Nahari et al., 2019). Instead, researchers have relied on assumptions about deception to study the phenomenon (Leins et al., 2013).

The first assumption is that lying is more cognitively demanding than telling the truth (Vrij, 2008; Vrij et al., 2008). Lie-tellers, unlike truth-tellers, must work to earn and maintain their credibility (Zuckerman et al., 1981). Formulating a lie is more taxing than simply recalling a memory, for example. A truthful memory is activated for both lie- and

truth-tellers, but only the former need to actively suppress it to deceive convincingly (Vrij, Mann, et al., 2010; Walczyk et al., 2013). In a recent review, Sporer (2016) suggested re-focusing research away from cognitive load manipulations towards the differences in cognitive processes.

Secondly, it is a long-standing assumption that lying and truth-telling involve different processes. Unlike truth-tellers, deceivers must carefully manage their behavior to avoid detection (DePaulo, 1992). Not only are lie-tellers supposed to be aware of their own bodies, but they are also expected to monitor the reactions of others (Zuckerman et al., 1981). Because of lie-tellers' potential to leak nonverbal cues to deceit, they can exhibit different nonverbal behaviors than truth-tellers (Ekman & Friesen, 1974). More recently, however, researchers have argued that lie-tellers and truth-tellers engage in similar forms of impression management (e.g., appearing calm and relaxed), but differ in terms of information management (Hartwig et al., 2010). Lie-tellers report preparing and strategizing more than truth-tellers (Hartwig et al., 2007). Specifically, lie-tellers have reported using strategies, such as avoiding lying, denying guilt, not providing incriminating details, and appearing pleasant, whereas truth-tellers generally rely on being perceived as honest by just telling the truth.

Thus, in recent years, deception detection researchers have turned their attention to examinations of verbal indicators of deceit (Nahari et al., 2019). The premise that has received the most empirical support is that lie-tellers tend to provide less information than truth-tellers (e.g., Amado et al., 2015, 2016; Masip et al., 2005; Vrij, 2008). Lie-tellers are perceived as less cooperative and less likely to admit forgetting, compared to truth-tellers (DePaulo et al., 2003). They also exhibit avoidance techniques, such as denial of

wrongdoing or simply skirting around the truth (Granhag & Hartwig, 2008). Lie-tellers' intentions are not just to avoid self-incrimination outright, but to also provide minimal information to limit strain on memory (i.e., less to remember) and fewer leads for investigators (Vrij, 2008).

Regardless of the differences between lie- and truth-tellers, the literature generally reflects the notion that the two groups are inherently different. However, there is select work which suggests that there might be more overlap than originally thought. Luke (2019) has demonstrated that evidence for deception cues which distinguish between veracity groups (e.g., plausibility, amount of details, nervousness) continues to be weak. Furthermore, there are countless countermeasures, such as rehearsal (e.g., Gawrylowicz et al., 2016; Vrij, 2000), that are used by lie-tellers to decrease differences on a variety of outcome measurements (e.g., reaction times, number of details, inconsistencies; Walczyk et al., 2009). It appears that many of the strategies that are used to induce differences between veracity groups may equally affect truth-tellers. In fact, Walczyk et al. (2013) outlined factors that may disproportionally add to the cognitive load of truth-tellers, including the comparative difficulty of retrieving old or decaying memories and simply recalling a rehearsed lie (Wixted, 2004). Evidently, there is some doubt that lie-tellers and truth-tellers are entirely distinct from one another.

1.2 Deception Detection

Overall, deception is difficult to detect (i.e., average accuracy is 54%; Bond & DePaulo, 2006). People from different professions (e.g., police officers, agents, lawyers, students, therapists, judges) are equally as skilled at detecting deception (Aamodt & Custer, 2006). In fact, the differences between individuals' abilities are minute (Bond &

DePaulo, 2008). Generally, people have higher rates of confidence in their skills to detect deception than their actual abilities warrant (DePaulo et al., 1997).

People also have a tendency to perceive statements as truthful (Bond & DePaulo, 2006). This truth-bias is speculated to be innate (Levine, 2014a). Although a lie-bias has been found in some contexts (e.g., when judging non-native speakers; Da Silva & Leach, 2013), observers do not typically have a tendency to assume that someone is lying. Thus, to overcome the truth-bias, elements of active decision-making might also influence the veracity judgments that are rendered.

1.3 Decision-Making

Across a multitude of fields, such as psychology, business, and mathematics, decision-making is perceived as a process driven by rational thought (Buchanan & O'Connell, 2006). Historically, however, theories in disciplines outside of psychology have failed to address the formulation stages of decision-making (Dillon, 1998). For example, normative theories, which focus on ideal decision-making, typically involve philosophical and mathematical considerations related to outcomes (Baron, 2004). Alternatively, psychologically-based notions of decision-making have primarily been concerned with descriptive decision-making (Dillon, 1998). These theories seek to explain or describe decision-making to justify subsequent behavior (e.g., Simon, 1955). For example, gambling researchers have advocated for the role of emotions in the formulation of decisions (e.g., Sanfey et al., 2003). Researchers have recently begun to develop prescriptive models of decision-making that are based on both normative theories and descriptive methods (Dillon, 1998).

1.3.1 Decision-Making in Law Enforcement Contexts

In general, our knowledge of decision-making in law enforcement contexts (i.e., how decisions are made in the field) is limited (Alpert et al., 2006). In fact, the majority of our understanding comes from official arrest reports and naturalistic observation (e.g., Black & Reiss, 1967). Decisions made under time pressures, for example, are more heavily influenced by negative, than positive, evidence (P. Wright, 1974). Alpert et al. (2006) have also found that contextual factors influence officers' decisions. Over the course of a typical shift, law enforcement officers formed many suspicions, but only investigated a proportion (i.e., 60%) of them further. The decision to follow up on a suspicion (or not) was based on specific factors related to job experience (e.g., suspect appearance and behavior, location and time of encounter) rather than evidence. Thus, law enforcement officers use both intuitive and deliberate components to make decisions (Brown & Daus, 2015).

Law enforcement decisions reflect ever-changing *processes* rather than fixed or predetermined outcomes. In Bonner's (2015) examination of decision-making in dispute encounters (i.e., situations in which a police officer witnessed or responded to a public complaint about unlawful behavior), law enforcement officers only used half of the information that they acquired. Officers also made several temporary decisions, prior to rendering a final decision, in more than half of the encounters (i.e., 56%). That is, officers altered their original decisions after new information was received. These findings speak to the importance of examining all elements of decision-making, especially timing, variability, and responsiveness to new information.

1.3.2 Decision-Making in Deception Detection

Perhaps because law enforcement officials' decision-making processes are complex, researchers have primarily attempted to study deception detection in the laboratory environment. Experiments typically take place over two phases: 1) deception production and 2) deception detection (DePaulo et al., 2003). First, witnesses[1] lie or tell the truth. For example, they are instructed to be deceptive or truthful about their beliefs or opinions (e.g., Leal et al., 2010), something that they witnessed (e.g., Elliott & Leach, 2016), or an event in which they participated (e.g., Evans et al., 2013; Porter & Yuille, 1996). Second, observers[2] view the accounts and make post hoc veracity decisions. These decisions can be made using a Likert scale, such as a 6-point scale (1 [definitely lying] to 6 [definitely telling the truth]; e.g., Leal et al., 2015). Most often, however, lie detection studies employ dichotomous measures (e.g., Brandt et al., 1980). Specifically, after viewing each account, observers choose either 'lie' or 'truth' (or variations, such as 'lying and 'telling the truth'; e.g., Leach & Da Silva, 2013).

These methodologies might constrain natural decision-making. First, findings in deception detection research are often based on a decision that is made after observers view the entire account (e.g., Leal et al., 2010). This approach renders researchers unable to determine the timing of observers' decisions, their initial decisions, and whether changes are made to those decisions over the course of the accounts. Second, observers are often required to render a decision about "whether the speaker was telling the truth"

---

[1] In the literature, the individual who is lying or telling the truth has been interchangeably referred to as the target, sender, witness, and interviewee (among others). For the sake of clarity, I will be referring to the person who is providing an account as the 'witness' throughout this dissertation.

[2] I will be using the term 'observer' (also referred to as listener, receiver, interviewer in the literature) to refer to the person who is making a veracity decision.

(Vrij, 1993, p.606). It is not clear whether observers are actually judging the veracity of the witness in general, the overall account, statements related to a particular event, or specific details provided. Third, there has been significant research on the types of details provided within accounts (e.g., auditory details, spatial details), but researchers have generally failed to examine the proportion of accurate and inaccurate details provided by witnesses. Witnesses are typically categorized as lie-tellers or truth-tellers based on their responses to a single critical question (e.g., "Did you cheat?"; Da Silva & Leach, 2013). However, observers' decisions might be affected by variations in the number of accurate and inaccurate details actually present in an account. Thus, in a series of studies, I explored each of these methodological gaps in the deception detection literature.

1.4 Research Aims

I examined the decision-making processes underlying deception detection. In Study 1, I investigated the strategies used by lie-tellers and truth-tellers through self-report, and by objectively assessing the contents of their accounts (i.e., accurate, inaccurate, and omitted information) across two interview stages. Next, I explored deception detection over time using a novel sliding scale in Study 2. In the final study, I analyzed how the framing of the deception detection task affected decision-making.

Chapter 2: Study 1

2.1 Introduction

Decades of research on deception and its detection have not yielded many definitive findings, aside from the consensus that the phenomena is complicated (Markowitz, 2020). Although there is no universal definition of deception, for a statement to be classified as deceptive it must be: 1) a false statement, or one that the lie-teller does not believe, and 2) made with an intent to deceive (for a full discussion, see Zuckerman et al., 1981). These conditions remain too fallible and narrow to be used reliably to detect deception (Carson, 2006; Mahon, 2016). The definition of deception is based on deeming specific statements as either lies or truths; however, a lie detection task requires a holistic judgment about deceptiveness (i.e., overall dishonesty across all statements). An average lie-teller's account might be made up of numerous lies and truths. In fact, self-reported 'good liars' have divulged that successful deception involves embedding false elements within truthful information (Verigin et al., 2019). Interestingly, it is not only lie-tellers who provide inaccurate information. Even truth-tellers have admitted to being dishonest throughout their accounts (Nahari et al., 2014); this appears to be an advantageous tactic when the truth would otherwise arouse suspicion (e.g., receiving exceptional test scores; Choshen-Hillel et al., 2020). Overall, the nuances of lie-telling and truth-telling might make it difficult – if not inappropriate – to characterize a person[3] or an entire account as deceptive.

[3] Throughout deception detection literature, individuals who take part in a study and are instructed to respond truthfully are named "truth-tellers", where people assigned to lie are often called "liar(s)" (instead of lie-*tellers*; Aamodt & Custer, 2006; Frank & Feeley, 2003; Lancaster et al., 2013)

Perhaps that is why some researchers have proposed that a person's overall deceptiveness should be judged on a continuum (e.g., Marsili, 2014; Street & Richardson, 2015). One such scale ranges from deceiving to telling the truth – including 'fuzzy lies' (Chisholm & Feehan, 1977) and 'graded-belief lies' (Carson, 2006) – and is based on the extent to which a person believes that he or she is deceiving (Marsili, 2014). Thus, the notion that sophisticated lie-tellers wield the truth strategically (Sutter, 2009) is logical when the intent to deceive becomes the central construct. On a continuum, dishonest accounts can include truthful information.

Yet, deception detection is typically examined as though lie-telling and truth-telling are completely distinct. One way to address this issue is to compare the information that each group chooses to provide or omit. I examined lie-tellers and truth-tellers' self-reported deception and honesty, and the types of information that they actually provided in their accounts (i.e., accurate details, inaccurate details, and omissions).

2.1.1 Relevant Deception Strategies

Lie-tellers and truth-tellers are strategic in their interview demeanors and responses. Both similarly engage in impression management, such as maintaining control of their nonverbal cues (e.g., Hartwig et al., 2010). They also adopt strategies – although different – related to information management. For example, lie-tellers may use omission-based deception due to their aversion to guilt (Battigalli et al., 2013). In fact, lie-tellers generally tend to prepare and carry out active approaches, such as denying guilt and feigning pleasantness, whereas truth-tellers rely on just recalling the event (Hartwig et al., 2007). Despite these strategic differences, observers struggle to identify deception (Bond & DePaulo, 2006).

This poor accuracy might be linked to deceivers' tactical use of the truth, otherwise known as paltering (Rogers et al., 2017). People rarely deceive (Abeler et al., 2019) and they do not like deceiving (Gneezy, 2005). According to the plausibility principle, when people do deceive, they tend to incorporate elements of the truth (Walczyk et al., 2014). Lie-tellers have self-reported strategically embedding truthful information into their otherwise deceitful accounts (Verigin et al., 2019). This selectivity has been observed in both individuals and groups of lie-tellers (Sutter, 2009). Providing accurate information is actually a clever way to omit information. For example, a suspect might choose to focus their account on harmless elements to appear sincere and cooperative while avoiding discussing critical or incriminating facts. In fact, although lie-tellers tend to provide fewer accurate details than truth-tellers, both groups provide more accurate than inaccurate information overall (Nahari, 2018).

Overall, there are three main strategies to deceive by manipulating the truth: paltering (i.e., strategically providing accurate information), omitting pertinent details, and purposefully providing inaccurate information (i.e., fabricating). The deception detection literature, to date, has predominantly focused on the latter. This approach might have led to an incomplete picture of the differences – and similarities – between lie-tellers and truth-tellers.

2.1.2 Analyzing Deception

Recently, researchers have highlighted the importance of examining verbal cues to deception, specifically in terms of the details provided by witnesses (e.g., Verschuere, Meijer, & Kleinberg in Nahari et al., 2019). Compared to behavioral cues, verbal cues may be easier to detect in real time because no special analyses or instruments (e.g., eye

trackers) are required. Furthermore, although there is no standard for how an honest person should behave, there is an expectation for people to maintain the basic maxims of conversation (Grice, 1989). The Information Manipulation Theory (McCornack, 1992) and the deception faucet metaphor (Markowitz, 2020) posit that deception can be identified through violations of effective communication, specifically Grice (1989) conversational tenets: quantity, quality, relevance, or manner. Three of the requirements – all except manner – are related to the details provided during a witness's account. Thus, it is important to examine details because they may reveal important clues to deception.

To date, the most common approach to examining details has been to focus on the first of Grice (1989) maxims: the quantity of information. Overall, lie-tellers tend to provide fewer details than truth-tellers (Amado et al., 2015, 2016; Masip et al., 2005; Vrij, 2008). Lie-tellers' accounts may lack in detail for strategic reasons, such as employing avoidant techniques (e.g., denial of wrongdoing; Granhag & Hartwig, 2008), minimizing the release of information that could be incriminating or inconsistent with evidence, and/or adhering to rehearsed responses to minimize cognitive demands. Moreover, both lie-telling and truth-telling activate working memory, but lie-tellers alone need to suppress incriminating truthful details (Vrij, Granhag, et al., 2010). Although there are apparent differences in information quantity between the two veracity groups, a more nuanced examination of the specific details provided is warranted.

There is a nascent focus on another of Grice's (1989) maxims: the quality of statements. A recent examination of the temporal positioning of deceptive statements revealed that statements followed by the truth contained more details than those followed by a lie (Verigin et al., 2020). However, the richness of lies did not appear to depend on

whether they had been embedded in honest or deceptive accounts. Thus, the quality of the entire account – not the specific lie – is impacted by the deceptive context.

The quality of details can also be conceptualized in terms of the verifiability of information (i.e., whether details were documented, could be corroborated, or were witnessed). Truth-tellers provide more verifiable details than lie-tellers overall (Palena et al., 2021). Moreover, truth-tellers provide even more verifiable information when they are told that their accounts will be checked, whereas lie-tellers continue to provide the same amount of details.

However, one vital component, in terms of any analysis of detail quality, that has been largely ignored in the literature is objective accuracy. The richness of an account or the mere presence of verifiable details does not guarantee that the information provided was correct. For example, a suspect could provide an extensive alibi about being home with a spouse while a crime was committed, but that does not mean that the alibi will be corroborated or that it is true. Forensic deception detection does not happen in vacuity; rather, it is a part of an investigation. Officers compare a suspect's account to the information or evidence that they have already gathered. Thus, a more comprehensive approach to assessing information quality would be to compare details directly against ground truth.

2.1.3 Interview Techniques

Interrogators use numerous techniques to encourage interviewees to reveal details (Kassin et al., 2007). In criminal justice contexts, that has included accusatorial approaches that recommend the use of pointed questions (i.e., direct questioning; Inbau et al., 2013). These close-ended requests to verify information do not result in

distinguishable differences between lie-tellers and truth-tellers (e.g., Tourangeau & Smith, 1996). Aggressive interrogation techniques have been known to elicit false confessions, as well (e.g., Kassin et al., 2010). Therefore, the focus has begun to shift away from accusatorial questioning to the use of information-gathering approaches (e.g., Meissner et al., 2015).

The cornerstone of information-gathering approaches is to allow an individual to speak and divulge information without interruption (e.g., Leins et al., 2013). Several researchers have recommended conducting multiphasic interviews, such as including a free recall portion followed by structured questions (e.g., Meissner et al., 2015). In one study, lie- and truth-tellers' repetitions (i.e., repeated details) were measured across three question protocols: free recall, sequential, and nonsequential (Deeb et al., 2017). Although truth-tellers provided more repetitions than lie-tellers across all conditions, differences between protocols were not statistically significant. However, the focus of the study was only on repetitions; question type might affect other types of details provided by lie-tellers and truth-tellers.

Indeed, interview questions do have an impact on speakers' rates of omissions. Work on the Strategic Use of Evidence (SUE) revealed that, compared to truth-tellers, lie-tellers omitted evidence more often during free recall and were more inaccurate when responding to structured questions (Hartwig et al., 2005). However, omissions and inaccuracies in that particular study were essentially the same concept (i.e., not mentioning a critical item was deemed an omission if it occurred during free recall, but an inconsistency if it occurred during structured questioning). Thus, there is a need to

more closely examine deceivers' accounts to find differences in the quality of details between lie-tellers and truth-tellers across interview protocols.

2.1.4 The Present Study

Research to date suggests that lie-tellers are not merely fabricating their entire accounts (e.g., Nahari, 2018; Sutter, 2009; Walczyk et al., 2014). I proposed a shift of focus from the quality of details to the objectively verified quality of details. To that end, I examined the rates at which lie-tellers and truth-tellers provided accurate and inaccurate details about an event using a systematic comparison to ground truth. Lie-tellers report leaving information out of their accounts (Hartwig et al., 2007), yet little is known about the frequency of these lies of omission. Therefore, I also explored witnesses' tendencies to omit details. Given best practice recommendations for multiphasic interviewing (e.g., Meissner et al., 2015), I also analyzed omissions and accuracy – operationalized as distinct constructs – in separate interview phases (i.e., free recall vs. structured questions).

2.1.4.1 Hypotheses.

2.1.4.1.1 Self-Reported Measures.

In accordance with the literature on cues to deception, I expected truth-tellers to provide more details than lie-tellers (DePaulo et al., 2003; Vrij, 2008). Based on previous work on lie-tellers' self-reported strategies (Nahari et al., 2014; Verigin et al., 2019), I hypothesized that both lie-tellers and truth-tellers would indicate that they embedded truthful and dishonest elements into their accounts. That is, I did not expect lie-tellers or truth-tellers to claim that they were exclusively deceptive or honest, respectively. However, I anticipated that witnesses self-reported proportions of honesty and deception

would coincide with their assigned condition. Specifically, I expected lie-tellers to report having lied for a larger proportion of their account than truth-tellers, and for truth-tellers to report having been honest for a larger proportion of their account than lie-tellers.

2.1.4.1.2 Objective Measures (i.e., Verified Details).

Across all three detail categories (i.e., accurate, inaccurate, and omitted details), I expected to find effects of veracity. I hypothesized that all witnesses (i.e., both lie-tellers and truth-tellers) would provide more information that was accurate than inaccurate or omitted (Nahari, 2018). However, I hypothesized that truth-tellers (vs. lie-tellers) would provide more accurate details and fewer inaccurate details, and omit fewer details.

I also expected to find effects of interview phase. I hypothesized that there would be more accurate and inaccurate details during structured questioning than free recall, whereas more omissions would be made during free recall than in response to structured questions (in keeping with Hartwig et al., 2005).

Finally, I conducted an exploratory analysis to determine whether veracity and interview phase would interact to affect any of the three types of verified details.

2.2 Method

2.2.1 Research Design

I used a 2 (Veracity: lie vs. truth) x 2 (Interview phase: free recall vs. structured-interview) mixed-factors design, with interview phase as a repeated measure. Specifically, participants were randomly assigned to either deceive or tell the truth during an interview that began with a free recall portion and ended with structured questions.

2.2.2 Participants

An *a priori* power analysis[4] (conducted in G*Power; Faul et al., 2007) revealed that 88 participants were required. I excluded participants who confessed (n = 15) and those who failed the manipulation check by incorrectly identifying their assigned condition (i.e., they said that they were assigned to tell the truth when they were supposed to deceive; n = 3). Overall, my final sample, after exclusions, consisted of 95 undergraduate students from Ontario Tech University who participated in exchange for course credit. The sample size was slightly larger than required because of the structure of the participant pool and the university's research ethics guidelines (i.e., people could sign up for sessions a week in advance and had to be permitted to participate once they had indicated their interest). Most of the participants were males (i.e., 56%) and the average age was 20.41 years (SD = 3.42). The participants self-identified as Arab/West Asian (7.4%), Black (6.3%), Chinese (3.2%), White (Caucasian; 31.6%), Hispanic (2.1%), Latin American (1.1%), South Asian (37.9%), South East Asian (7.4%), and Other (3.2).

2.2.3 Materials

2.2.3.1 Video Stimuli. Participants were presented with an innocuous or suspicious video (see Elliott & Leach, 2016 for a description of the Suspicious Event Paradigm). The focal point of the 30 second videos was a computer desk that was covered with personal belongings. As the camera zoomed into the background of the frame, the items on the wall (e.g., a map, newspaper articles, a calendar) became visible. In the innocuous video, participants saw a typical office setting with various items scattered on a desk

[4] To achieve 99% power when using a mixed factors ANOVA with 2 groups and 3 measurements to detect a medium effect size (Cohen's d = .50)

(e.g., a laptop, an empty package, books, pictures). The suspicious video featured the same office; however, some items were altered to arouse suspicion about a possible terrorist plot (e.g., a pipe bomb in the package, a bomb-making manual).

2.2.3.2 Demographics Questionnaire. Participants were asked to provide their demographic information including age, race, and gender.

2.2.3.3 Manipulation Questionnaire. Participants were asked to report their assigned condition by selecting whether they were instructed to deceive or tell the truth. Next, they were asked to select all the items that they remembered seeing in the video from a lengthy list of items that were either present in both conditions (e.g., a laptop), just one condition (e.g., a pipe bomb), or neither condition (e.g., a printer).

2.2.3.4 Post Experimental Questionnaire. Participants were asked to report the percentage of their account that was a lie and the percentage of their account that was the truth (from 0% to 100%). To assess motivation, participants were asked whether they were motivated to convince the experimenter that they were telling the truth (yes or no). Finally, participants were asked whether they thought that the experimenter believed them (yes or no).

2.2.3.5 Objective Details. I created two coding manuals – one for each veracity condition – that consisted of all possible objectively verified details contained in the innocuous and suspicious videos. For each item (e.g., a package), there were several descriptive details that could be provided (e.g., open, empty, cardboard, brown, rectangular, lid/flaps, label). Each detail could be categorized as accurate (i.e., described correctly), inaccurate (i.e., described incorrectly), or omitted (e.g., present in the video and selected by the participant in the manipulation check as being present, but not

mentioned in the participant's account). For example, if a participant in the suspicious video condition made the statement "I did not see anything because the box was closed", then there were three (underlined) details that represent an omission, an accurate detail, and an inaccuracy.

All of the participants' interviews were transcribed by two independent, trained raters. Next, the raters used the coding manual to analyze the transcripts. To establish reliability, both raters examined a random selection of transcripts (i.e., 20%; ICC = .89). Disagreements were resolved through discussion and the agreed upon scores were used. Once reliability was established, one rater coded the remainder of the transcripts.

2.2.4 Procedure

After participants arrived at the laboratory individually, the experimenter directed them to a small testing room. The experimenter, who was blind to condition, entered an anonymous identification number into MediaLab (Jarvis, 2018) which randomly assigned participants to watch either the innocuous or suspicious video on the laptop after she left the room. Those participants who saw the innocuous video were instructed to tell the truth about what they had seen (i.e., a typical office), whereas viewers of the suspicious video were told to deceive the experimenter and respond as though they had seen a typical office. After the video, an instruction on the screen notified participants that they had two minutes to prepare before the experimenter returned to conduct the interview. To motivate witnesses in both conditions, participants were told that they would be awarded $50 if they successfully convinced the experimenter that they were telling the truth. In fact, all witnesses were entered into a draw for the reward.

The experimenter came into the room, closed the door, and turned on the video camera. The interview began with a free recall question (i.e., "Tell me everything you saw in the video") and was followed with questions of increasing specificity (i.e., from open-ended questions to close-ended questions), such as, "What items did you see on the desk?", "What books were on the table?", "What date was marked on the calendar?", "Where was the bomb?" After the interview, participants were asked to complete the manipulation questionnaire truthfully (i.e., indicate what they had actually observed, regardless of their assigned condition). Finally, participants were debriefed and asked to provide consent for the use of their video interview in future studies (e.g., Studies 2 and 3).

2.3 Results

Overall, I collected 95 interviews (i.e., 52 truth-tellers, 43 lie-tellers) with an average length of 294.3 seconds (SD = 103.8 s).

2.3.1 Self-Reported Measures

The following analyses were conducted on Post-Experimental Questionnaire items.

2.3.1.1 Motivation. I used a chi-square test to analyze participants' motivation to convince the experimenter of their truthfulness. I found a small, but significant, relationship between veracity and motivation, χ^2 (1, N = 94) = 6.15, p = .013, ϕ = .28. That is, truth-tellers (90.2%) were more likely to report being motivated than lie-tellers (67.4%).

2.3.1.2 Perceived Convincingness. Next, I examined participants' perceived success in convincing the experimenter that they were telling the truth using a chi square test. I found a significant, medium-sized relationship between veracity and participants'

perceived success, χ^2 (1, N = 92) = 12.37, p < .001, ϕ = .39. Specifically, truth-tellers (76.9%) were more likely than lie-tellers (23.1%) to report that the experimenter believed them.

2.3.1.3 Perceived Veracity of Accounts. Finally, I conducted a MANOVA to examine the remaining items in the post-experimental questionnaire (Figure 1). I asked all participants how much of the account provided was true and how much was deceptive. As hypothesized, there was a significant main effect of veracity, V_s = .90, F(1, 94) = 399.06, p < .001, η^2 = .81, 95% CI [.74, .85]. I conducted pre-planned independent samples t-tests on each veracity group to investigate further. Compared to lie-tellers (M = 19.63, SD = 18.95), truth-tellers (M = 97.99, SD = 5.85) reported that a greater percentage of the account provided was true, t(94) = 28.26, p < .001, d = 5.82, 95% CI [4.89, 6.75]. Lie-tellers (M = 81.30, SD = 18.91) reported that a greater percentage of the account was deceptive than did truth-tellers (M = 3.59, SD = 14.34), t(94) = 22.76, p < .001, d = 4.69, 95% CI [3.90, 5.47].

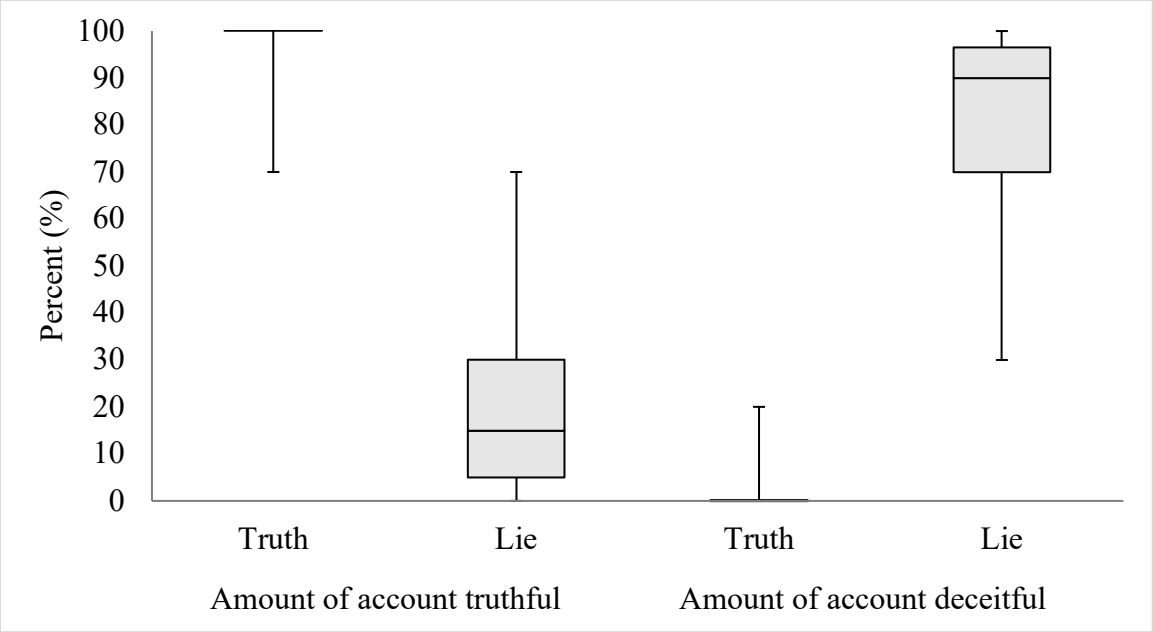
2.3.1.3.1. Honesty. The majority (i.e., 76.9%) of truth-tellers claimed that they did not deceive at all and that their account was entirely (i.e., 100%) honest; as expected, however, none of the lie-tellers (i.e., 0%) claimed to be completely honest. There was some variability in the percentage of truths that truth-tellers and lie-tellers admitted to telling: truth-tellers reported that no less than 70% of their account truthful, whereas lie-tellers reported that no more than 70% of their account was truthful.

2.3.1.3.2. Deception. Only 16.3% of lie-tellers reported that their accounts were completely fabricated (i.e., 100% deceptive), whereas 77% of truth-tellers claimed they provided no lies at all. Reported rates of deception had a ceiling effect, such that lie-

tellers stated that at least 30% of their accounts were deceptive. Truth-tellers claimed up to 20% of their accounts were fabricated.

Figure 1

Box plots of observers' self-reported proportions of truthful and deceitful information



Note. The boxplots range from 0% to 100%. Thus, boxes are not missing, but this is an indication that those quartiles and whiskers are all the same value. For example, based on truth-tellers' self-reported amounts of truthful information in their accounts: the 1st and 3rd quartiles, median, and maximum are all 100%.

2.3.2 Objective Measure - Verified Details

Although the number of accurate and inaccurate details that a participant could mention was theoretically infinite, I restricted my analysis of verified details to 12 categories: papers, cup with objects, tape, books, laptop, picture of Obama, calendar, map with pictures, articles, manual, package, and blueprints. These categories were selected because they were present in the manipulation questionnaire. Thus, I could verify that

participants actually observed and remembered the items, providing ground truth – particularly for omissions.

2.3.2.1 Accurate Details. Coders assigned a value of '1' to each item marked on the manipulation questionnaire and described correctly in the account. The accurate detail variable was calculated by averaging participant scores across the 12 detail categories.

2.3.2.2 Inaccurate Details. This variable was created by averaging across 12 categories using a similar method as above. That is, inaccurate details were assigned a value of '1' for each detail that was correctly reported in the manipulation questionnaire and incorrectly during participants' accounts.

2.3.2.3 Omitted Details. Omitted details were calculated by assigning a value of '1' to each detail that was reported in both the manipulation questionnaire and the participant's account, and a value of '0' to items that the participant indicated was present in the manipulation questionnaire, but was not mentioned in the account. Thus, lower scores represented a higher proportion of omissions. Finally, scores on all 12 categories were averaged to create the omitted details variable.

2.3.2.4. Analyses of Details. I conducted a Veracity (lie-tellers vs. truth-tellers) x Interview phase (free recall. vs. structured questions) repeated measures MANOVA on witnesses' use of details (i.e., accurate, inaccurate, omitted). I found significant main effects of veracity, $V_s = .68$, $F(1, 94) = 64.68$, $p < .001$, $\eta^2 = .41$, 95% CI [.26, .53], on the combined dependent variables. I examined univariate effects and found that veracity had an effect on accurate, $F(1, 94) = 55.53$, $p < .001$, $\eta^2 = .37$, 95% CI [.22, .49], and omitted details, $F(1, 94) = 171.43$, $p < .001$, $\eta^2 = .65$, 95% CI [.53, .72], but not on inaccurate details, $F(1, 94) = 2.18$, $p = .143$, $\eta^2 = .02$, 95% CI [.00, .11]. Post hoc pair-

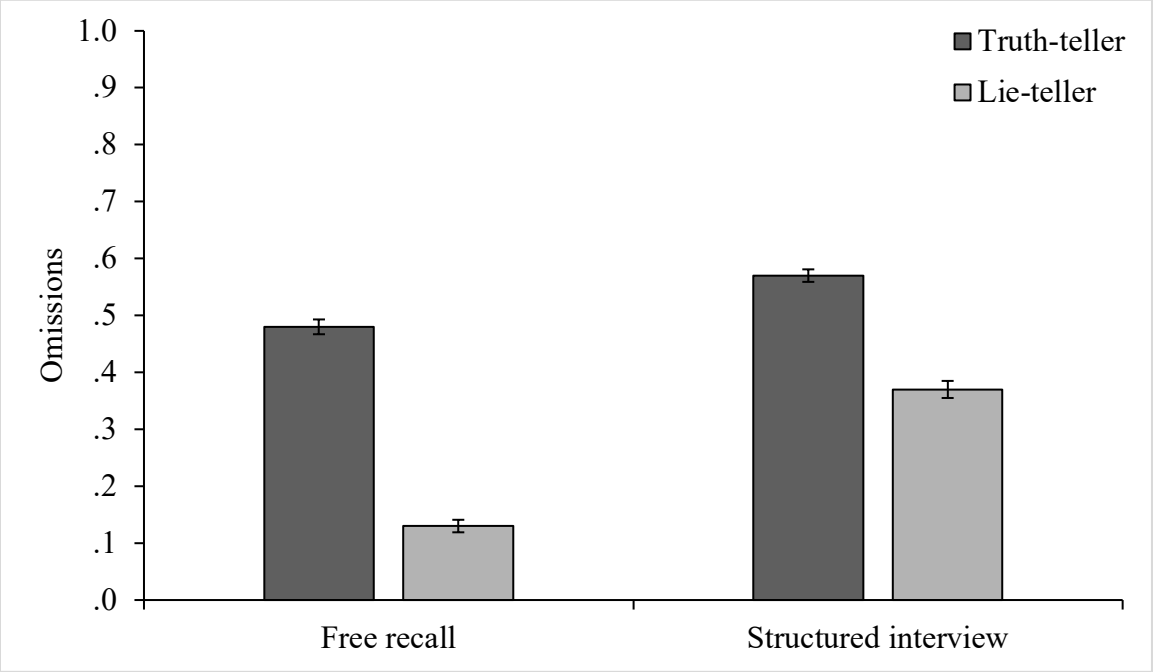
wise comparisons revealed that truth (M = 1.91, SD = .05) provided more accurate details than lie-tellers (M = 1.33, SD = .06), whereas lie-tellers (M = .25, SD = .02) omitted more details in their accounts than truth-tellers (M = .53, SD = .01).

I also found a significant main effect of interview phase, V_s = .71, $F(1, 94) = 73.33$, $p < .001$, $\eta^2 = .44$, 95% CI [.29, .55] on the combined dependent variables. Univariate effects revealed that interview phase had an effect on accurate details, $F(1, 94) = 58.95$, $p < .001$, $\eta^2 = .39$, 95% CI [.23, .51], and omitted details, $F(1, 94) = 160.66$, $p < .001$, $\eta^2 = .63$, 95% CI [.51, .71], but not on inaccurate details, $F(1, 94) = 0.12$, $p = .735$, $\eta^2 = .00$, 95% CI [.00, .05]. Specifically, witnesses provided more accurate details during structured questioning (M = 1.82, SD = .51) than during free recall (M = 1.47, SD = .53), but witnesses made more omissions during the free recall portion of the interview (M = .32, SD = .21) than during structured questioning (M = .48, SD = .16).

Finally, I found a significant interaction between interview phase and veracity, V_s = .28, $F(1, 94) = 11.67$, $p < .001$, $\eta^2 = .11$, 95% CI [.02, .23]. Univariate analyses revealed that interview phase and veracity interacted with omitted details, $F(1, 94) = 33.54$, $p < .001$, $\eta^2 = .27$, 95% CI [.12, .39], but not accurate details, $F(1, 94) = 1.16$, $p = .285$, $\eta^2 = .01$, 95% CI [.00, .09], nor inaccurate details, $F(1, 94) = 1.55$, $p = .216$, $\eta^2 = .02$, 95% CI [.00, .10]. Post hoc independent samples t-tests revealed that, during the free recall phase of the interview, lie-tellers provided more omissions than truth-tellers, $t(94) = 15.93$, $p < .001$, d = 3.28, 95% CI [2.66, 3.90]. When answering structured questions, lie-tellers also omitted more details than truth-tellers, $t(94) = 7.68$, $p < .001$, d = 1.58, 95% CI [1.12, 2.04], but to a lesser extent.

Figure 2

Lie-tellers' and truth-tellers' omissions during free recall and structured questioning



2.4 Discussion

I examined lie-tellers' and truth-tellers' perceptions of the extent to which they were actually deceitful or honest during their accounts. Lie-tellers reported providing accounts that were largely deceptive, whereas truth-tellers' accounts were reported to be mostly truthful. However, as I hypothesized, both groups indicated that they incorporated truthful and deceptive elements into their accounts. Recall that lie-tellers self-reported that about 20% of their accounts were accurate. Perhaps they strategically provided a non-trivial amount of truthful information – a strategy known as paltering – to ensure believability or to lessen cognitive load (Nahari et al., 2014; Vrij, 2008). Furthermore, approximately 23% of truth-tellers said that a portion of their accounts were deceitful. Although it is unclear why truth-tellers would deceive during interviews, one explanation

is that they were also being strategic: they might have been deceptive when they believed that accurately reporting certain details would appear incriminating and arouse suspicion.

To assess how participants' perceptions reflected reality, I analyzed their accounts to determine the actual amount of information that was accurate, inaccurate, and omitted. As hypothesized, truth-tellers provided more accurate information than lie-tellers. These results were unsurprising as they replicated previous research (e.g., Hartwig et al., 2005; Nahari, 2018) and were in keeping with traditional notions of lie-telling and truth-telling. However, they may not be practically meaningful. Upon closer inspection of the data, the difference between the groups was less than a single detail. This finding suggests that, similar to other cues to deception (see Hartwig & Bond, 2011), there was a distinction between lie-telling and truth-tellers, but it was so minute that it may not be perceptible to investigators. Consequently, deception detection accuracy may be improved by increasing the differences between the amount of accurate verified details that lie-tellers and truth-tellers provide.

Similar to the findings on accurate details, there were no significant differences between lie-tellers' and truth-tellers' accounts in terms of inaccuracies. The lack of differences between groups cannot be attributed to a dearth of incorrect responses, however. Both groups did provide inaccurate information (at a lower rate than accurate information and higher than omissions). Thus, verifying information by only checking for explicit inconsistencies between an account and evidence, as is the current practice (e.g., Deeb et al., 2018), will not be diagnostic and may lead to incorrect conclusions.

I had also hypothesized that truth-tellers would rarely provide inaccuracies (as suggested by Nahari, 2018), but had not anticipated that lie-tellers would follow suit. This

finding contradicts the common notion that lie-tellers' statements are riddled with false information. When reviewed holistically, the distinction between veracity groups is not that lie-tellers provide more falsifiable details, but that truth-tellers provide more accurate details. This conclusion fits with the recent meta-analytic finding that truth-tellers provide more verifiable details than lie-tellers (Palena et al., 2021).

The findings supported my hypothesis that lie-tellers would omit more information than truth-tellers. In fact, my research suggests that lie-tellers tell lies of omission (i.e., 75% of details) rather than commission. This finding echoes lie-tellers' self-reported behavior. Specifically, when lie-tellers are given the opportunity to speak freely (i.e., during the free recall phase of an interview), they would rather not tell lies of commission (Leins et al., 2013). Omitting information might be lie-tellers' strategy because inaccurate information could be verified or falsified throughout an investigation. Thus, investigators may wish to pay closer attention to interviewees' omissions (i.e., examine what they *do not* say as opposed to what they *do* say). Importantly, as in this study, omissions are only verifiable if they are compared to ground truth; therefore, investigators need to compare accounts to evidence and play close attention to critical omissions. Caution is needed, however, because the omission effect is novel, and more research is needed to fully understand its scope before it is implemented.

Finally, I found that interview phases differentially affected the presence and accuracy of details in participants' accounts. As expected, and in line with previous findings (Hartwig et al., 2005), witnesses provided more accurate information during structured questioning than during free recall. The finding is not surprising given that the structured questions performed exactly as expected (i.e., elicited more information).

Furthermore, my data provides further support for the notion that interviewers should not rely on free recall alone, as witnesses do not know how much information should be provided (this is akin to the logic behind the model statement approach; e.g., Vrij et al., 2020).

Although accurate information varied based on questioning type, structured questions did not highlight the differences between lie-tellers' and truth-tellers' use of inaccuracies. Rather, it was the difference in omissions provided by lie-tellers and truth-tellers that depended on the interview phase. Specifically, I found that lie-tellers and truth-tellers omitted more details in the free recall phase than in the structured questions phase. Given that the difference between lie-tellers and truth-tellers was more substantial during free recall – with lie-tellers omitting 87% of details – interviewers might consider paying closer attention to that phase of the interview. Of course, the use of omissions to detect deception depends on knowing that information is being omitted (i.e., having ground truth or evidence). Interestingly, Hartwig et al. (2005) reported that observers tend to assume that someone is being deceptive if they omit information during free recall compared to during structured questioning. Although observers' intuition is correct (i.e., lie-tellers do omit more than truth-tellers during free recall), my data suggest that *both* lie-tellers and truth-tellers are more likely to omit information during free recall than when they are asked structured questions. The problem is that omission during free recall is not a Pinocchio's nose, as observers expect. Thus, observers should not rely on it as their sole decision-making criterion to discern between lie-tellers and truth-tellers.

2.4.1 Limitations and Future Directions

My results may have been impacted by the paradigm that I used to elicit deceptive behavior. Although the suspicious event paradigm has been used successfully in previous studies (e.g., Elliott & Leach, 2016; Woolridge et al., 2020), the task may be more conducive to instances of omissions than other approaches. Specifically, it was likely easier for witnesses to omit critical details altogether instead of falsifying information (e.g., denying presence of a bomb instead of fabricating information about a microphone). Thus, it is unknown whether the nature of the task or witnesses' deliberate strategies produced the tendency to omit, instead of fabricate, details. Upon reviewing witnesses' responses, however, I did find examples of lies of commission. Although it may not have been the preferred deception strategy, several commissions were elicited by the chosen paradigm. Nonetheless, another study should be conducted to examine whether the findings will be replicated with different conditions or paradigms (e.g., a mission paradigm; Vrij et al., 2010), and generalize to other types of deception.

I set out to record any details that witnesses mentioned and compare them to ground truth, which was originally operationalized as anything appearing in the stimulus video. However, I quickly ran into the difficulty of analyzing an infinite number of details. It was impossible to preemptively account for all possible details that a participant might mention, regardless of their relevance (e.g., the color of the lights), and ensure that they had been attended to by all participants. Given that one of the variables (i.e., omissions) was reliant on the number of items that were identified on the manipulation questionnaire, I elected to create the variables by focusing on the details pertaining only to those items. Thus, my results are not reflective of all of the details that were – or could

have been – mentioned. My original operationalization of ground truth was abandoned, and a new definition was adopted: ground truth encompassed participants' responses to the manipulation questionnaire (i.e., items reported) and the coding manual based on the stimuli videos. Future researchers may achieve ground truth differently to assess central details, as well as peripheral details. One way to establish ground truth might be to create a manipulation check that is more extensive (i.e., list all possible items present in the video and their descriptors). Alternatively, participants could be made to review the footage of their own accounts and point out where they were telling the truth or deceiving. Regardless, the method used in this study was a valid approach and a good first step in the analysis of witness accounts.

One might argue that a final limitation was the lack of counterbalancing of interview question types (i.e., free recall vs. structured questions). However, it is unusual to ask direct questions prior to free recall; doing so would run counter to most interviewing techniques. Furthermore, the approach used in this study followed best practices (e.g., Meissner et al., 2015). Thus, perhaps future research should examine other investigative interviewing procedures, such as the PEACE interview model, to determine whether techniques with a focus on information gathering differently impact witnesses' accounts.

2.5 Implications and Conclusions

I examined witnesses' accounts from a novel perspective: I focused on the objectively verified quality, rather than the quantity, of three distinct types of information. Truth-tellers were more honest than lie-tellers overall. However, witnesses' self-reports and objective analyses of their accounts revealed that both provided incorporated truthful *and* deceptive information. They also omitted information at

unexpectedly high rates. On the whole, lie-tellers omitted information to a greater extent, suggesting that lie-tellers might prefer this strategy to lies of commission. Thus, there is an opportunity to examine the use of omissions as a method of deception detection. Given that the omission effect is novel, and its practical use is unknown, further research is needed. Additional work is also required to determine how accounts differ across multiphasic interviews. Here, free recall yielded less accurate details, but more differences between lie-tellers and truth-tellers in terms of omissions. Changes in witnesses' strategies, and ways in which deception is revealed across questions, merits further study.

Overall, my findings challenge conventional depictions of lie-tellers as purely deceptive and truth-tellers as exclusively honest. Throughout the study, there was evidence of the ways in which the two veracity groups differed from one another – in terms of the quality of details – which was contradictory to expectations. Thus, relying on incorrect conceptualizations of lie-tellers and truth-tellers might contribute to poor deception detection rates.

Chapter 3: Study 2

3.1 Introduction

Deception detection has now been examined for several decades (e.g. (Bond & DePaulo, 2006; Chisholm & Feehan, 1977; Ekman, 1985; Vrij, 2008). Over time, fundamental "facts" have been cemented in the literature. First, the average accuracy rate is about as good as a guess (i.e., 54%; Bond & DePaulo, 2006). Although attempts have been made to shift the accuracy rate (e.g., Vrij et al., 2010), the improvements have been modest. Second, there is a truth bias: people tend to assume that others are telling the truth (e.g., Bond & DePaulo, 2006; Levine, 2014; Vrij, 2008; Zuckerman et al., 1981). However, these seminal findings are primarily predicated on a single approach to measure deception: dichotomous post hoc decision-making. Anecdotal evidence suggests that restricting judgments to the end of an account does not mirror how people make decisions. I examined whether observers' natural deception detection abilities have been obscured by this nearly standard methodology.

3.1.1 Static Decision-Making

Although the decision-making processes leading to a veracity judgment are decidedly complex, researchers have attempted to study them using relatively uniform and simplistic measures. Post hoc decisions are a staple of measurement in experimental psychology (Shanks, 2017). In deception research, observers render a decision after viewing a witness providing a statement or being interviewed (e.g., Hartwig et al., 2005). These decisions can be made using rating scales (e.g., 1 [definitely lying] to 6 [definitely telling the truth]; Leal et al., 2015), but researchers typically use dichotomous forced-choice measures (e.g., asking participants to identify the statement as a 'lie' or 'truth'; Bond & DePaulo, 2006). Most notably, there is a distinct delay – hence *post* hoc – and

change in task from watching a witness's account to rendering a decision. This approach suggests decision-making is static: there is an inherent assumption that people reach a single conclusion and do not change their mind, and that a single judgment can represent the entire decision-making process. Thus, our current measures may constrain or not fully capture natural decision-making.

3.1.2 Indirect versus Direct Decision-Making

Whether deception detection is best performed directly or indirectly is a hotly debated topic in psychology, with each group claiming superior accuracy (e.g., Levine, 2019; Sporer & Ulatowska, 2021). Proponents of direct deception detection claim that deliberate and effortful thought leads to higher accuracy than intuitive or less than fully conscious cognitive processing (Levine, 2019). Conversely, one of the arguments in favor of indirect decision-making is that accuracy is relatively poor when people are explicitly told to detect deception (Bond & DePaulo, 2006). The distinction between the two opposing concepts is that direct deception detection requires conscious mental effort (e.g., attention, memory), whereas (unknown) subconscious processes are required to detect deception indirectly (Sporer & Ulatowska, 2021). For this reason, there are many different terms used for direct and indirect deception detection, such as unconscious or subconscious versus conscious, explicit versus implicit, and deliberate versus intuitive (e.g., Albrechtsen et al., 2009; Ask et al., 2013; Street & Richardson, 2015). Overall, deception detection is typically categorized as direct when an observer is explicitly told to judge veracity (e.g., "Is the person lying or telling the truth?"), whereas indirect deception detection occurs when an observer assesses veracity through impressions of, or feelings

about, a witness (e.g., concepts associated with deception, such as *truthfulness* and *honesty*).

Our lack of understanding about the decision-making process perpetuates this debate. Champions of direct deception detection call into question whether natural (i.e., innate) decision-making exists (Street & Richardson, 2015c). The cruces of their debate are interpretation and semantics: *indirect* judgment versus judging *indirectly*. Judging indirectly refers to the typical way that deception tasks are performed: the observer is expected to determine whether a witness is lying without further direction. Indirect judgment, however, involves specific instructions for observers to use indirect cues of deception (e.g., thinking hard, nervousness) to determine veracity. After conducting two experiments, they proposed an alternative explanation for observers' higher accuracy when receiving indirect (vs. direct) prompts: indirect instructions focus observers' attention on practical cues compared to the general directions associated with the deception detection task. In addition, they suggested that current examinations of indirect decision-making did not uncover subconscious processes – the very processes that they are meant to test. Alternatively, indirect decision-making enthusiasts claim that accuracy improvements made through training may be external representations of indirect abilities (ten Brinke et al., 2014). Ultimately, the debate comes down to the fact that the two concepts are not mutually exclusive – one concept cannot be examined in isolation from the other – and cannot be resolved using the current methods.

I did not strive to resolve the debate about the (in)directness of decision-making in this study; rather, I have presented a brief overview of the issue to illustrate that understanding the decision-making process is vital in deception detection research. Our

knowledge is constrained by the measures that we use to examine deception detection and should serve as motivation to explore alternative approaches. One of the ways to push the field forward is to shift from static, dichotomous assessments to dynamic (i.e., continuous) examinations of decision-making.

3.1.3 Decision-Making Over Time

Rather than exclusively measuring deception detection at the end of an account, it is important to understand decision-making *during* an interrogation. Overall, the literature on deception detection decisions over time is nascent and it is unclear whether exposure length affects decision-making. Ambady and Rosenthal (1992) coined the term 'thin slicing' in reference to short observations (i.e., under 5 min). Their seminal meta-analysis revealed that decisions based on observations between 30 s to 5 min in length were equally accurate, regardless of the type of stimuli presented (e.g., face, speech), testing location (i.e., the laboratory vs. the field), or subject matter (e.g., clinical decisions, deception detection). In a subsequent study, they found no difference in student evaluations made after an entire semester and after only 30 s of exposure to a professor (Ambady & Rosenthal, 1993). They concluded that accurate decisions can be made after as little as 30 s and remain constant over time – or, said another way, there were no improvements in accuracy with longer exposure times. In a more recent examination of deception detection performance using very thin slices, however, observers who viewed 15 s clips were more accurate than those who saw the videos in their entirety (i.e., about 3 min; Albrechtsen et al., 2009). Although these studies examined the nature of exposure length and decision-making, the findings may have been mixed because the lengths of

exposure varied greatly (i.e., an average of 137 s versus 15 s). Thus, the association between time and decision-making in deception detection remains unclear.

Timing can also be a vital component of decision-making (e.g., Shanks, 2017). Typically, study participants are asked to provide veracity judgments at the end of the video clip regardless of its length. In one of the first studies of its kind, police officers and undergraduate students viewed two accounts provided by a single witness, and then indicated whether she was being deceptive or telling the truth in each account as well as when they had reached that conclusion (Masip et al., 2003). Analyses of post hoc self-reports revealed that decisions made at the beginning of interviews were more accurate than decisions made at the middle or end. In addition, decisions about deceptive accounts were more likely to have been reported as being reached early, whereas decisions about truthful accounts were unrelated to timing.

The same authors conducted two follow-up studies to further examine timing effects (Masip et al., 2006, 2009). Seated in the same room, groups of 12-14 observers (who were asked not to speak or make noise during the experiment) watched videos on a television and made veracity decisions. They watched 12 separate videos of people providing responses to three prompts (e.g., "Describe in detail what the man with a moustache did"). The experimenter paused the video after each prompt and observers recorded a veracity judgment. They were also asked when they solidified their decision (i.e., following the witness' first, second, or third responses). Contrary to Masip et al. (2003), there was no impact of time on accuracy (Masip et al., 2006) nor did observers' accuracy increase over time (from the first video to the last; Masip et al., 2009). However, this latter effect was largely impacted by observers' judgments of truthful

statements because accurate detection of truth-tellers decreased over time, whereas the detection of lie-tellers increased over time.

Although this work offers insight into the role of timing in decisions, it is not without its limitations. Observers were asked to retrospectively estimate when they made their decisions; thus, it is possible that observers forgot or changed their answers. Despite efforts to examine timing and decision-making, the methodology could not fully capture decisions as they happened. Finally, the studies compared lie accuracy, truth accuracy, and their proportions, but lacked examinations of observers' ability to discriminate between lie-tellers and truth-tellers, and their tendencies to identify witnesses as either deceitful or honest (i.e., bias).

Indeed, there have been only a few published articles on how deception detection decisions change over time in terms of discrimination and response bias. Leins et al. (2017) conducted a study in which signal detection was examined using experienced observers (i.e., military or law enforcement personnel). Observers watched two videos – one in English and one in Arabic – and made real-time veracity judgments using a joystick to move a vertical onscreen scale (i.e., up to indicate 'lie' and down for 'truth'). The onscreen scale was color coded: the top section was red, the middle section was yellow, and the bottom section was green. Instead of recording every movement of the joystick, a program recorded the scale position every second. Twenty-five percent of the judgments were deemed 'uncertain' as the responses fell in the yellow zone. These were discarded because they could not be factored into signal detection analyses. Overall, observers' ability to discriminate between lie-tellers and truth-tellers decreased over time

(i.e., between the first and fourth quarters of each interview), whereas bias remained the same throughout.

As with discrimination, research on response tendencies has not yielded a distinct finding. In two studies, observers made decisions using computer keys labelled 'lie,' 'truth,' and 'unsure' (Street & Richardson, 2015a, 2015b). As they viewed each witness's account, they held down the key that corresponded with their decision until they changed their mind (e.g., from 'lie' to 'truth'). For analyses, the data were split into time intervals made up of five equal sections (which varied in length with each clip). Overall, observers did not make their decisions immediately: on average, there was a 3.64 s delay before the first decision was provided. Although the proportion of truth judgments remained stable when observers viewed deceptive accounts, there was an early increase and eventual plateau (i.e., increase between times 1, 2, 3, but not 4 and 5) when observers made decisions about truthful accounts. Observers' use of the 'unsure' option, however, decreased over time. Because analyses were only conducted on truth and uncertainty judgments, effects on lie-bias remain unknown. Moreover, time segments and the corresponding content were not equal across stimuli because participants spoke for different amounts of time (e.g., truthful accounts lasted between 16 s to 91 s and deceptive accounts ranged from 10 s to 55 s), likely resulting in data loss. However, this work does offer insight into the decision-making processes underlying deception detection. Most importantly, the findings here, and in the previously reviewed studies, demonstrated that people change their minds over the course of an account.

Street and Richardson's (2015) work also suggests that the response options impact responding. Observers made truth judgments early, but the addition of the

'unsure' option eliminated this timing effect. The central premise of Truth-Default Theory (Levine, 2014a) is that when people are uncertain and are forced to choose between calling someone a lie-teller or truth-teller, they default to assuming that people are honest. Thus, dichotomous scales – which are typical in the field of deception detection – might artificially elicit a truth-bias that is remedied by the presence of an alternate option. Of note, however, is that even in Street and Richardson's (2015a, 2015b) studies, observers were still forced to make categorical decisions (i.e., they were restricted to choosing between 'lie,' 'truth, and' 'unsure').

The above review highlights that the way in which decision-making has been examined in the deception detection literature may have contributed to artificially producing the results that we deem "facts". Specifically, observers' poor deception detection performances and response biases have typically been determined by assessing *post hoc* judgments. However, anecdotal and psychological evidence both within and outside of deception detection indicates that decision-making does not happen at a single final moment (e.g., Krix et al., 2014) and that people change their minds (e.g., Street & Richardson, 2015). A continuous measurement might be more appropriate to evaluate observers' ability to discriminate between lie-tellers and truth-tellers, and to assess their decision bias.

3.1.4 The Present Study

In the current study, I examined an alternative to static decision-making: dynamic deception detection. This was a novel approach whereby deception detectors rendered moment-to-moment decisions using a proprietary sliding scale (i.e., the Dynamic Decision Scale; https://osf.io/jqckr/?view_only=7e899750429 24a7e8f454877f9df7775)

as they viewed lie-tellers and truth-tellers. The scale was presented while the account was being shown, started at a neutral position (i.e., with the cursor in the middle), and included only veracity labels (i.e., 'lie' and 'truth') on either end. In previous research, decision times were analyzed using equally spaced time intervals (Masip et al., 2003; Street & Richardson, 2015b). Restricting decisions incrementally may have obscured their true distribution. Here, I captured observers' responses in terms of decision time (in milliseconds) as the Dynamic Decision Scale was used. I sought to compare observers' abilities to discriminate between lie-tellers and truth-tellers, and their corresponding decision biases across three conditions: control (i.e., dichotomous, post hoc), dynamic (i.e., continuous throughout an account), and combined. The goal was to examine the (dis)advantages of dynamic judgments by comparing them to a traditional dichotomous approach. I also included a combined condition to directly compare the unique experiences of observers who rendered both a dichotomous and dynamic decision.

3.1.4.1 Hypotheses.

3.1.4.1.1 Discrimination. Based on previous deception detection research, I anticipated that observers in the control condition would be sensitive to the differences between lie-tellers and truth-tellers, and be able to discriminate between the two groups (e.g., Bond & DePaulo, 2006). Given that this was the first study to analyze dynamic deception detection decisions, I did not have specific hypotheses in terms of how it related to a traditional measure. Ultimately, I sought to explore the effects (if any) of the dynamic condition.

3.1.4.1.2 Bias. In typical deception detection studies (e.g., Bond & DePaulo, 2006), observers exhibit a truth-bias; therefore, I expected that the responses of observers in the

control condition would be consistent with a truth-bias. Research on response bias as it relates to time, however, was limited and problematically defined (e.g., early responses categorized as indicators of bias; Street & Richardson, 2015). Two competing findings were derived from the literature: Masip et al. (2003) found that lie judgments were made early, whereas Street and Richardson (2015) found that observers' initial decisions tended to be 'truth' which was eliminated when they were given an 'unsure' option (Street & Richardson, 2015b). Given that my scale provided observers with an implied 'unsure' option – the midpoint – I hypothesized that observers in the dynamic condition would not possess a truth-bias.

3.1.4.1.3 Decisions Over Time (Combined Condition Only). There were competing findings in the literature relating to deception detection decreasing (Leins et al., 2017; Masip et al., 2003), increasing (Masip et al., 2009), and not changing (Masip et al., 2006) over time. In a study with continuous judgments (as opposed to post hoc decisions) and one interviewee, Leins et al. (2017) found a decrease in performance between the first and last quarters of each interview. I extended this study by using multiple witnesses to examine my hypothesis that discrimination between lie-tellers and truth-tellers diminished over time (i.e., in the dynamic condition). Furthermore, based on Leins et al. (2017), I predicted that bias would not change over time (i.e., between the first and last decisions).

I also expected observers' first and final decisions to differ. Based on Albrechtsen et al. (2009) finding that very thin slices yielded higher accuracy than viewing an entire account, I hypothesized that first decisions in the combined condition would be more accurate than final dichotomous decisions.

3.1.4.1.4 Time to Decision. I expected observers to render their first decisions within the first 10% of the average video length (i.e. early), in keeping with Street and Richardson (2015) finding.

3.2 Method

3.2.1 Research Design

I used a 2 (Veracity: lie vs. truth) x 3 (Response Condition: control vs. dynamic vs. combined) mixed-factors design, with veracity as a repeated measure. Specifically, observers were randomly assigned to view videos of lie-tellers and truth-tellers. They were asked to detect deception either at the end of each interview (i.e., *control condition*), throughout each video (i.e., *dynamic condition*), or both (i.e., *combined condition*).

3.2.2 Participants

An a priori power analysis[5] using G*Power (Faul et al., 2007) revealed that 120 observers were required. There were no exclusion criteria based on age, gender, or race; however, 22 participants were excluded because they either did not follow instructions (e.g., did not make a judgment) or did not finish the experiment. The final sample consisted of 120 undergraduate students from Ontario Tech University who participated in exchange for course credit. The sample was made up of participants who self-identified as female (59.2%) and male (40.8%). Their average age was 20.54 years (SD = 3.05), and their self-identified races were as follows: White (34.2%), South Asian (24.2%), Black (10.8%), Arab/West Asian (7.5%), South East Asian (6.7%), Chinese (5.0%), Filipino (5.0%), Other (4.2%), Latin American (1.7%), and Hispanic (.8%).

[5] To achieve 80% power when using a repeated measures between-factors ANOVA with 3 groups and 2 measurements to detect a medium effect size (Cohen's d = .50)

3.2.3 Materials

Video Stimuli. The stimuli featured a selected subset of 20 (from 95 participants in Study 1) with an average length of 265.1 s (*SD* = 72.1) was for this study. The videos of truth-tellers ranged from 158 s to 404 s (*M* = 284.5, *SD* = 73.3), whereas videos of lie-tellers ranged from 155 s to 367 s (*M* = 245.7, *SD* = 69.1). Whenever possible, the videos were yoked for gender, age, race, and video length. Half of the witnesses were female, and the average age of the witnesses was 19.35 years (*SD* = 1.57). The witnesses self-identified as South East Asian (25%), Black (20%), Hispanic (20%), White (15%), Chinese (10%), and Other (10%).

Lie Detection Task. Observers were assigned to one of three lie detection task conditions: *control*, *dynamic*, or *combined*. As is typical in deception detection studies (e.g., DePaulo et al., 2003), observers in the *control* condition completed a static dichotomous decision after each video by identifying whether the witness was being deceptive or telling the truth. In the *dynamic* condition, observers used the computer's mouse to click and slide a cursor along the onscreen scale; the decision was recorded (in milliseconds) when the mouse click was released. At the top of the screen, observers saw the question "Is the participant in the video lying or telling the truth?" The scale was located below the video on the computer screen and was labeled with 'Lie' or 'Truth' on opposite ends, with the labels counterbalanced to avoid left side bias (e.g., Mathews, 1929). There were no numbers or dashes along the length of the scale, and the initial position of the cursor was in the middle of the scale. Observers were given a short tutorial on using the sliding scale that included an instructional video and an interactive manipulation test. The tutorial featured me describing the task and then instructing

participants to slide the cursor to the right, then left, half of the scale to practice registering a decision. As they completed their tasks, the program verified their selection and returned a popup with either a message stating 'Try again' or 'Task completed.' Finally, observers in the *combined* condition were asked to provide their deception judgments throughout the video using the scale and then to render a static dichotomous decision at the end of the video.

Demographics Questionnaire. Observers self-reported demographic information including gender, age, and race.

3.2.4 Procedure

After arriving at the laboratory and providing their consent to participate in the study, observers were seated individually at computers and asked to wear headphones throughout the experiment. Each observer was randomly assigned to one of the three conditions, but they all saw the same 20 videos of witnesses either being deceptive (n = 10) or telling the truth (n = 10) about what they had witnessed. Witness order was random and counterbalanced to avoid order effects. Based on their assigned conditions, observers either made veracity decisions while watching a witness (dynamic), a static dichotomous decision after the video (control), or both (combined). Regardless of condition, as soon as participants clicked 'continue' to begin the study or after rendering a decision, the stimulus video would automatically play (i.e., participants did not need to start the video). Finally, observers completed a demographic questionnaire and were then debriefed. The entire session was typically completed in about an hour and a half.

3.3 Results

I performed data cleaning by looking for extreme outliers (e.g., no overall accuracy, missing data) and subsequently reviewing individual responses. There were several

reasons for which participants were excluded. I omitted any observers who had a response rate lower than 90% (i.e., at least 18/20 trials completed) to allow for a learning curve on the new task. Some observers did not make timely decisions prior to the end of the first couple of videos (i.e., did not render a decision prior to the end of the clip) but they completed all successive trials, for example. In total, I excluded 22 witnesses with incomplete data: 7 participants did not follow instructions (i.e., failed to complete enough trials), and 15 participants did not finish the study (i.e., they withdrew from the study prior to completion[6]).

Preliminary analyses indicated that there were no effects of race, age, or gender on the outcome variables, nor were there notable differences based on the distribution of demographic characteristics across conditions. Thus, analyses were collapsed across these demographic conditions. I also collapsed all analyses across the *control* and *dynamic* conditions[7]. I ran a series of independent-samples *t*-tests to compare accuracy (i.e., overall, lie, and truth accuracy) rates across dichotomous responses in the *control* and the *combined* conditions, and the same type of tests to compare dynamic responses in the *combined* and *dynamic* conditions. The analyses revealed that there were no significant differences between the same response types across conditions (all *p*s > .097). Therefore, I used the collapsed conditions in the analyses below, except where a direct comparison in the *combined* condition was conducted, to form two groups: *control* (n = 74) and *dynamic* (n = 80).

---

[6] Anecdotal evidence suggests that this was because they had scheduled other activities during the experimental session.
[7] Collapsing conditions violated the assumption of independence associated with a one-way ANOVA.

When nonsignificant effects were observed, I provided the Bayes Factor (BF_{01}; calculated with a Cauchy prior of .707 using Rouder et al., 2009) to explain the strength of the evidence for a null effect. This measure can help us understand the relative strength of the results as evidence for two opposing theories (i.e., null and alternative hypotheses). In recent years, Bayes factors have been used to better explain the meaning behind non-significant findings, gaining favor over simply using *p*-values or effect sizes (Dienes, 2014).

3.3.1 Deception Detection

I calculated observers' average first, final, and overall decisions. The first decision could only be assessed in the *dynamic* condition. The final decision was specified as the last selection on the scale in the *dynamic* condition and dichotomous selection in the *control* condition. Overall decisions in the *dynamic* condition were calculated by averaging all scale decisions made by each observer for each witness, whereas overall decisions in the *control* condition were calculated by averaging all dichotomous decisions.

The sliding scale did not include numerical labels; however, it was coded from 0 (i.e., 'Lie') to 100 (i.e., 'Truth') with each unit corresponding to a precise distance on the scale. To simplify analyses, I categorized any decisions between 0 and 49 on the scale as deception judgments, 50 as 'unsure' or missing data, and truth judgments as decisions ranging from 51 to 100. These numbers were reverse coded when the category labels were counterbalanced. I assigned a '0' to each incorrect decision and a '1' to each correct decision, and then averaged observers' scores.

3.3.1.1 Comparing Conditions. I used Signal Detection Theory (SDT; Green & Swets, 1966) and Stanislaw and Todorov's (1999) formulas to explore the effects of condition on observers' discriminations between lie-tellers and truth-tellers (i.e., $d' = Z_{Hit} - Z_{False\ Alarm}$) and their decision biases (i.e., $c = -[(Z_{Hit} + Z_{False\ Alarm})]/2$). Observers' decisions were categorized as 'hits' if they correctly identified lie-tellers, whereas incorrectly classifying truth-tellers as lie-tellers was categorized as 'false alarms.' I applied standard corrections [i.e., .5/signal or noise or (signal or noise - .5)/signal or noise] to all extreme values (i.e., 0s and 1s). The stimuli were evenly distributed for each veracity condition: 10 noise and 10 signals. Thus, hit rates and false alarms of 0 were changed to .05, whereas hit rates and false alarms of 1 were changed to .95. To compare conditions, I used the calculated discrimination and bias values in the *control* and *dynamic* conditions. The results of the SDT analyses are provided in Table 1.

Table 1

Results of one-sample t-tests comparing observers' discrimination (d') and bias (c) scores to 0

		Discrimination (*d'*)				Bias (*c*)			
Condition	*n*	*t*	*p*	*d*	95% CI	*t*	*p*	*d*	95% CI
Control	40	9.75	<.001	1.13	[.84, 1.42]	3.96	<.001	.46	[.22, .70]
Dynamic	46	1.94	.056	.22	[-.01, .44]	.40	.692	.05	[-.18, .26]

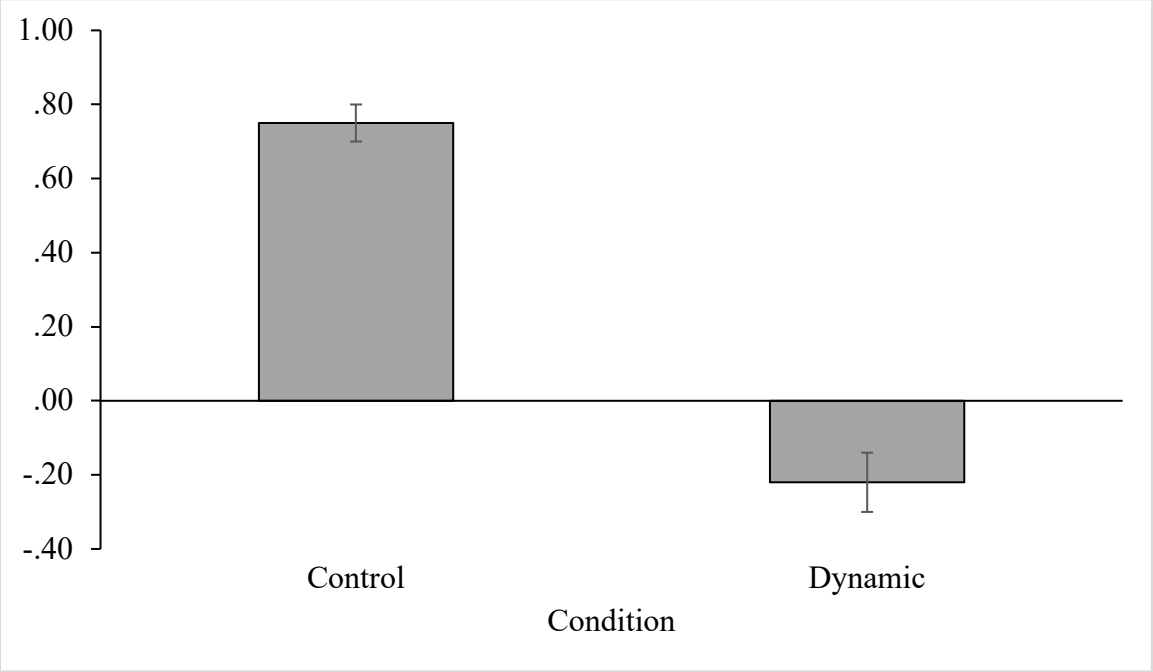
Note. CI = confidence interval.

3.3.1.1.1 Discrimination (d'). I conducted a one-way ANOVA on the effect of condition on observers' abilities to discriminate between lie-tellers and truth-tellers (results in Figure 3). I found a main effect of condition, $F(1, 153) = 48.92$, $p < .001$, $\eta_p^2 = .24$, 95% CI, [.13, .35]. Specifically, observers in the *control* condition were better at

discriminating between lie-tellers and truth-tellers than observers in the *dynamic* condition. I then used one-sample *t*-tests to compare observers' discrimination scores to "0" (i.e., no discrimination). Observers were able to discriminate between lie-tellers and truth-tellers in the *control*, but not the *dynamic*, condition. The Bayes factor for observers' discrimination in the *dynamic* condition ($BF_{01} = 1.37$) implies that there is no evidence for the null or alternative hypothesis.

Figure 3

Observers' discrimination (d') in the control and dynamic conditions



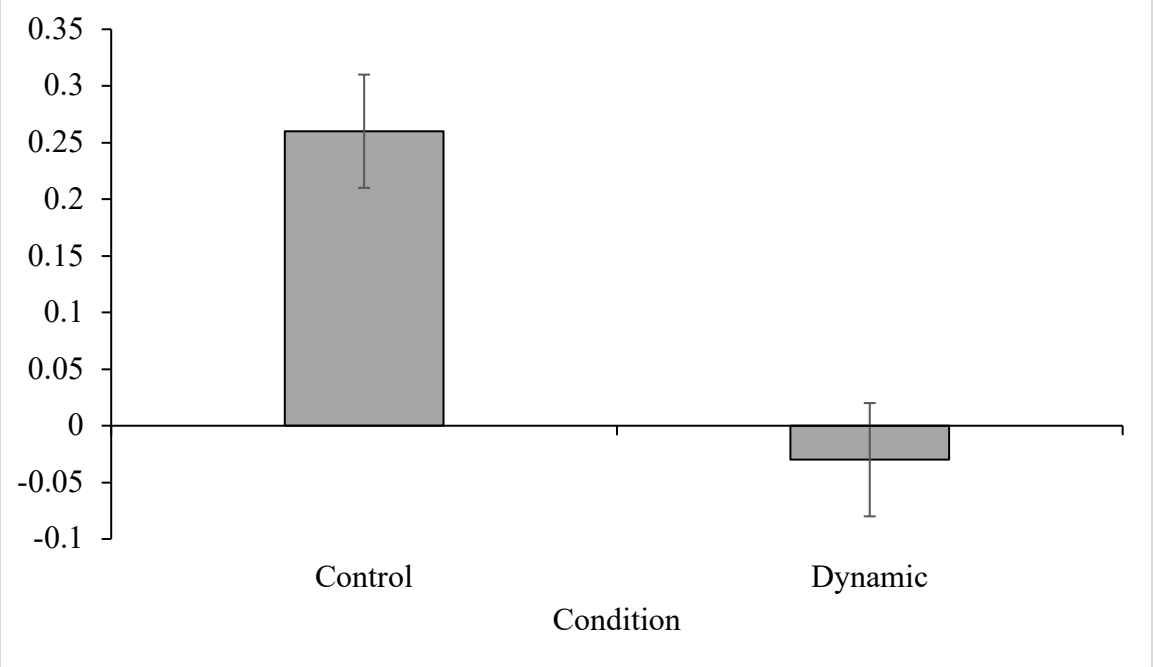
3.3.1.1.2 Bias (c). Next, I conducted a one-way ANOVA on the effect of condition on observers' decision bias (results in Figure 4). There was an effect of condition, $F(1, 153) = 8.91$, $p = .003$, $\eta_p^2 = .06$, 95% CI [.01, .14], such that observers in the *control* condition were more likely to exhibit a bias than observers in the *dynamic* condition. I used one-sample *t*-tests to compare observers' decision bias to "0" (i.e., no bias). Observers exhibited a truth-bias in the *control* condition, but no response bias in the

dynamic condition (BF_{01} = 7.52). The large Bayes factor implies that these findings are substantial evidence for observers' lack of response bias in the *dynamic* condition.

I also looked at bias in terms of observers' tendency to choose 'lie-teller' or 'truth-teller' as their first decisions. I calculated the mode of the first decision for each witness. The findings revealed that in 10 of the 20 videos, observers tended to judge witnesses as lie-tellers in their initial decisions. However, this was closely followed by 8 videos in which the most common first decisions were 'truth-teller', and 2 videos where the proportion of initial 'lie-teller' and 'truth-teller' judgments were the same.

Figure 4

Observers' response bias (c) in the control and dynamic conditions



3.3.1.2 Combined Condition Only. The *combined* condition was unique because observers made both moment-to-moment and a post hoc decision.

I conducted a paired-samples *t*-test to compare observers' discrimination between lie-tellers and truth-tellers in their overall moment-to-moment decisions and post hoc

selections. Observers' ability to discriminate was significantly better when they rendered post hoc decisions (M = .83, SD = .69) compared to moment-to-moment decisions (M = -.43, SD = .97), $t(34) = 5.03$, $p < .001$, d = .86, 95% CI [.46, 1.25]. A comparison of observers' discrimination scores to "0", I revealed that they were able to discriminate between lie-tellers and truth-tellers when making post hoc decisions, $t(33) = 6.99$, $p <$ .001, d = 1.20, 95% CI [.75, 1.64], and moment-to-moment decisions, $t(33) = -2.62$, $p =$ .013, d = -.45, 95% CI [-.80, -.09].

Next, I compared observers' response biases in their moment-to-moment and post hoc decisions. A paired-samples t-test revealed that there was a significant difference between observers' moment-to-moment (M = .32, SD = .51) and post hoc decisions (M = -.17, SD = .66), $t(33) = 3.16$, $p = .003$, d = .54, 95% CI [.18, .90]. I conducted one-sample t-tests to compare observers' bias scores to "0" and found that their post hoc decisions were truth-biased, $t(33) = 3.67$, $p < .001$, d = .63, 95% CI [.26, .99], but their moment-to-moment decisions were without bias (BF_{01} =.37), $t(33) = -1.64$, $p = .110$, d = -.28, 95% CI [-.62, .06]. The Bayes factor indicates that the evidence is weak (i.e., no strong support for null or alternative hypothesis).

3.3.1.2.1 Decisions Over Time. Observers' decisions ranged from 0 decisions up to 200 decisions during a single account. I examined discrimination over time by comparing observers' first scale decisions, last scale decisions, and final dichotomous decisions. A repeated-measures ANOVA revealed that there was a significant difference between observers' first (M = -.31, SD = .90), last (M = -.41, SD = 1.12), and final decisions (M = .83, SD = .69), $F(1, 33) = 24.17$, $p < .001$, $\eta_p^2 = .42$, 95% CI [.16, .60]. Follow up paired-samples t-tests indicated that observers were better able to distinguish between lie-tellers

and truth-tellers when making their final dichotomous decisions compared to their first scale decisions, $t(33) = 4.92$, $p < .001$, $d = .84$, 95% CI [.45, 1.23], and their last scale decisions, $t(33) = 4.59$, $p < .001$, $d = .79$, 95% CI [.40, 1.17]. In contrast, there was no difference in observers' sensitivity to veracity between their first and last scale decisions, $t(33) = .67$, $p = .507$, $d = .12$, 95% CI [-.22, .45], $BF_{01} = .81$. The Bayes factor indicates no support for either a null or alternative hypothesis.

I also conducted a repeated-measures ANOVA to analyze bias over time. I found a significant effect of observers' bias on their first ($M = -.26$, $SD = .89$), last ($M = -.11$, $SD = .69$), and final decisions ($M = .32$, $SD = .51$), $F(1, 33) = 5.03$, $p = .013$, $\eta_p^2 = .24$, 95% CI [.00, .34]. I performed paired-samples *t*-tests to determine the direction of the effects. Bias was similar across first and last scale decisions (i.e., there was a tendency to indicate that the witness was a lie-teller), $t(33) = 1.58$, $p = .123$, $d = .27$, 95% CI [.61, .07], $BF_{01} = .28$. However, bias differed when comparing final dichotomous decisions (i.e., truth-biased) to first scale decisions, $t(33) = 3.21$, $p = .003$, $d = .55$, 95% CI [.91, .19], and their last scale decisions, $t(33) = 2.56$, $p = .015$, $d = .44$, 95% CI [.79, .08]. The Bayes factor associated with observers' bias in their first and last scale decisions indicates strong support for a lack of bias.

3.3.2 Time to Decision

Analyses pertaining to observers' decision times were only conducted on the *dynamic* and *combined* conditions because they could vary over time. Overall, the average time to a first decision was 65.62 s ($SD = 39.49$), and ranged from 10.67 s to 188.40 s. I found a significant difference when I tested the hypothesis that first decisions would be made within the first 10% of the average length of a video (i.e., 26.51 s) and the

actual average time to first decision (i.e., 65.62 s), $t(79) = 8.86$, $p < .001$, $d = .08$, 95% CI [-.14, .30]. That is, on average, first decisions were made much later than 10% of the way through an account. I conducted a paired-samples *t*-test on the average amount of time that elapsed before observers made their initial decisions (i.e., moved the curser from the default starting point) about each veracity group. I found no significant differences, however, between the average amount of time that observers took to make their first decisions while watching truth-tellers ($M = 66.47$ s, $SD = 44.57$) versus lie-tellers ($M = 64.77$ s, $SD = 36.76$), $t(79) = .73$, $p = .469$, $d = .08$, 95% CI [-.14, .30], $BF_{01} = -.73$. The associated Bayes factor indicates strong support for the lack of difference in time to first decisions between veracity conditions.

3.4 Discussion

In this study, I used a novel approach to assess deception detection and examine two seminal findings in the literature: the almost chance-level ability to discriminate between lie-tellers and truth-tellers and the presence of a truth-bias. I compared dynamic judgments to the traditional dichotomous approach first. Recall that I did not have a specific hypothesis for the impact of *dynamic* decision-making on observers' discrimination; however, I did anticipate that observers in the *control* condition would be sensitive to the differences between lie-tellers and truth-tellers (e.g., Bond & DePaulo, 2006). Not only did I find that observers in the *control* condition were able to discriminate between lie-tellers and truth-tellers, but they were also more sensitive than observers in the *dynamic* condition. In fact, when I compared discrimination rates to 0 (i.e., no sensitivity), observers who used the scale to make their deception judgments were ineffective at discerning between lie-tellers and truth-tellers. This is surprising given

the assertion that passive observation (akin to what was present in the post hoc control condition) should be inferior to active deception detection (Levine, 2014b).

Nonetheless, there are several potential explanations for observers' poorer performance when they used the scale to render their decisions. In reviewing the data closely, the effect is likely most impacted by the high rate at which observers in the *control* condition were detecting truth-tellers (i.e., 71% accuracy) compared to lie-tellers in the same condition (54%), and truth-tellers (46%) and lie-tellers (47%) in the *dynamic* condition. Most of these values were expected because we know that truth (vs. lie) accuracy tends to be higher and deception detection rates are around chance (Bond & DePaulo, 2006). However, I did not anticipate that observers would be unsuccessful at detecting truth-tellers in the *dynamic* condition. One explanation is that deliberate, concurrent, deception detection is more difficult than quick, post hoc judgments. For example, in a study of people's previous experience with deception, most people reported detecting deception long after it happened, with only 2% declaring that they made judgments in real time (Park et al., 2002).

We can also look to the memory literature for an explanation. The observers' added task of continuously attending to, and reporting deception judgments, in the *dynamic* condition likely led to extraneous cognitive load. Cognitive load theory is based on the capacity constraints of working memory (Beckmann, 2010). Specifically, individuals experience cognitive load when they are performing a task that causes them to exceed the limits of their working memory. In the deception detection literature, cognitive load is often discussed in terms of the deceiver (e.g., Vrij, 2008). However, it is plausible that observers are just as likely to experience cognitive load during a complex deception

detection task because they perform many of the same activities (e.g., monitoring behaviors, remembering details). In fact, observers' memory for witnesses was poorer when they were tasked with performing a concurrent task while also rendering veracity judgments (Hanway et al., 2021). Not only were participants in the *dynamic* condition judging veracity, but they were doing so continuously and likely overloading their working memory capabilities.

It is also important to note that the inability to discriminate between lie-tellers and truth-tellers was not driven by a lie-bias in the *dynamic* condition. In keeping with previous research, I expected that observers in the *control* condition would hold a truth-bias (Bond & DePaulo, 2006) and that observers in the *dynamic* condition would not (Street & Richardson, 2015b). To align with the literature, I assessed bias in two ways: overall response bias and observers' tendencies to choose 'truth-teller' as their first decisions. There was a difference between conditions such that response biases were higher in the *control* versus *dynamic* condition. As predicted, only observers in the *control* condition exhibited an overall truth-bias. However, observers' first decisions in the *dynamic* condition were most likely to be 'lie.' I was not entirely surprised that the use of the scale eliminated a near-standard finding in the field of deception (i.e., truth-bias). Although the basis for Truth-Default Theory is that people tend to believe others, they do adapt to different situations accordingly (Levine, 2014a). If people expect that others are telling the truth, then I should have found that the most common selection was 'truth'. Levine's (2014) theory, however, leaves room for variations as it posits that the tendency to believe others is adaptive. It appears that making holistic decisions about deception (such as in the *control* condition) allows observers to utilize their truth-bias;

however, observers may be more guarded and do not trust their bias when they are asked to make continuous decisions about a witness (as in the *dynamic* condition).

The tendency to initially assume that an individual is deceiving might also be related to suspicion. Specifically, people can anticipate that others will be honest in most situations, but there are contexts in which deception might be expected. Thus, another way to understand my findings is to see them as a representation of people's beliefs about lying and how they change with decision modality. In another study, observers who were suspicious – defined as expecting to make veracity judgments – had increased rates of lie-bias (Pickel et al., 2013). It is possible that the scale evoked suspicion given that observers anticipated continuing to perform the deception detection task throughout the video. For example, instead of making a holistic judgment, as is typical of a dichotomous deception detection task, observers may have scrutinized each sentence or behavior because the scale task implied that deception would appear at various points during the account. I did not explicitly examine this notion; however, it should be explored in future research.

I also explored the timing of decision-making. Based on Street and Richardson's (2015) findings, I hypothesized that initial decisions would be made within the first 10% percent of the average video length (i.e., early in the video). I found, however, that there was a significant difference between when I expected them to make their first decision (i.e., 10%) and the actual average time that observers took to make their first decisions (i.e., 24.8%). In fact, observers' first decisions were rendered between 4.0% and 71.1% of the way through videos. The discrepancy in results between my study and Street and Richardson's (2015) might be explained by the differences in video lengths. Most

notably, the video stimuli used in their study were significantly shorter than those used here (10 s – 91 s versus 155 s – 404 s, respectively). Arguably, observers' average initial decisions in both studies were made 'early on' in the video (i.e., within the first quarter of the video). Observers made these early decisions whether they were judging truth-tellers or lie-tellers. Thus, these early decisions do not seem to be related to the notion that lie-tellers would stand out or that the truth would shine through an account.

I also tested whether, in keeping with Leins et al.'s (2017) findings, observers' sensitivity decreased throughout the interview (i.e., between the first and last quarter). Contrary to that study, and my hypothesis, I did not find support for diminished sensitivity: observers' discrimination was equally poor in both their first and last scale decisions. One of the explanations for this finding is that observers' performance in moment-to-moment decisions should not technically improve unless they receive feedback or instruction (e.g., Meissner & Kassin, 2002). Furthermore, similar performance in moment-to-moment decisions might imply that observers are truly judging an account sentence by sentence and, thus, a decision made at the beginning should be independent of a final decision.

For this reason, my most intriguing findings are perhaps those related to the *combined* condition. Observers were better able to discriminate between lie-tellers and truth-tellers when they rendered their post hoc dichotomous decisions compared to their scale decisions. The accuracy of their final scale decisions was also poorer than their final dichotomous decisions. It is possible that the dichotomous decisions were holistic, whereas the dynamic decisions were based on specific moments (e.g., details, witness reaction or movement). There is even an argument to be made for the presence of

deliberate versus intuitive processing. The use of the scale for deception detection may have encouraged deliberate processing, and the post hoc dichotomous decision intuition. Specifically, akin to direct processing, moment-to-moment decisions likely require immediate attention and assessment of evidence at face value (Levine, 2019). Post hoc decisions, however, are similar to indirect decisions as observers are required to make holistic judgments after viewing all of the evidence (Sporer & Ulatowska, 2021). In fact, observers might rely on their intuitive or subconscious knowledge, as it compares to the cognitive schema of a 'lie', to render a decision (ten Brinke et al., 2014). Although my findings may be explained using research on direct versus indirect deception detection, I did not examine these concepts in this dissertation. Thus, the use of direct versus indirect decision-making in deception detection remains an empirical question.

3.4.1 Limitations and Future Directions

Most of the limitations in this study stem from the novelty of the approach. Not only did I change the time at which decisions were rendered, but also the method by which they were retrieved. Specifically, I employed two different measures of decision-making: a scale and a dichotomous measure. This change in modality, however, was necessary to reconceptualize the way in which deception detection decisions were made based on the growing research into how decisions change over time (Albrechtsen et al., 2009; Leins et al., 2017; Street & Richardson, 2015b). Because deception detection research often focuses on the prescribed methods of the judgment (e.g., cue use; Evans et al., 2013), we do not yet know the direct impact of information throughout an account on a decision (whether it is moment-to-moment or holistic). Thus, future researchers should

examine not only dynamic decision-making, but also the direct impact of incoming information on corresponding decisions.

This study was conducted to examine the decision-making process during deception detection. But I did not collect metrics and feedback directly from observers about their experiences completing the task. Future researchers might consider incorporating Ericsson and Simon's (1980) “think aloud” protocol to understand the reasoning behind observers’ decisions. That would allow us to further understand the decision-making process, such as when decisions changed polarity (i.e., from ‘lie’ to ‘truth’ and vice versa). However, caution should be used in creating a task where cognitive resources will be depleted, as deception detection performance may suffer further (Hanway et al., 2021).

Finally, I sought to explore the dynamic nature of decision-making. I did not, however, examine the data dynamically. There are countless challenges with data in a study where an infinite amount of data points can be recorded for each participant. The most germane challenge to the analyses was the difficulty in assessing data that was relevant and applicable to previous findings in the literature. The few studies (e.g., Leins et al., 2017; Street & Richardson, 2015) in which veracity judgments were made moment-to-moment assessed the data dichotomously and/or through the division of time into increments. Future studies should attempt to implement methods outside of the deception detection literature to be able to assess the data dynamically (e.g., Drift Diffusion Model; Ratcliff, 1978) and with greater precision. My approach, however, is a first step in the exploration of the dynamic nature of decision-making. Although there may be shortcomings associated with the specific methodology employed, this work constitutes a much-needed reconsideration of traditional deception detection.

3.4.2 Implications

The results of this study make us question what we "know" about deception detection. The current literature is predicated on the principle that the decision-making process can be represented by a single post hoc decision. Not only do people change their minds, but their tendency to do so actually undermines their ability to discriminate. This novel scale and method help to move the field of deception detection forward as they call into question seminal findings. Overall, these results demonstrate that previous deception research, which only focused on observers' final decisions, had failed to assess how people truly make decisions. I found that people are making decisions *throughout* an account and their decisions are *poor*, regardless of whether witnesses are deceiving or telling the truth.

My findings also have implications for deception detection research. The goal of this field of forensic psychology is to recommend best practices in contexts where important veracity judgments are made (e.g., borders, courtrooms, police stations). My findings suggest that decision-making in deception detection occurs differently than previously represented in research; instead of observers making a single definitive decision, their decisions change over time. Thus, the way in which veracity decisions have been assessed in research does not match observers' natural patterns of decision-making. My methodology for gathering decisions (i.e., moment-to-moment responses) likely more closely mimics the way in which law enforcement personnel judge deception, for example. On the whole, my results call into question previous research on deception detection. Given that my findings show a different pattern of decision-making (i.e., poor

discriminability, absence of truth-bias), it is possible that previous assessments overestimated observers' performance and misrepresented biases.

In my analyses, I found that people made similar initial and final decisions even though they might have changed their minds throughout the decision-making process (i.e., as they listened to the account). In the test-taking literature, psychologists have found that it is common for people to change their mind (i.e., alter their responses; (Aryadoust, 2020). Interestingly, they only change a small proportion of their answers, but when they do, the answers tend to be changed from incorrect-to-correct (Benjamin et al., 1984). Although I found that observers changed their decisions throughout an account, their final decisions were often similar to their last. Perhaps because observers had no knowledge of ground truth, as test takers do (i.e., they are tested on recall of known subjects), their decision-making follows a different pattern. Instead of changing their decisions from incorrect-to-correct, observers in my study changed their decisions from initial-to-alternate-to-initial – regardless of the accuracy of that decision. Some potential avenues for exploration include the presence of tunnel vision and confirmation bias (e.g., Elaad, 2022).

3.5 Conclusion

The results of this study contribute to, and enhance, our understanding of decision-making in deception detection. To date, deception detection has been regarded like a trial verdict – a person is either deemed guilty (i.e., deceiving) or not guilty (i.e., telling the truth); however, real-world interviewing is more fluid, and a final decision is rarely recorded. My findings uncovered that, as with real-world interviewing, people change their minds during decision-making. Also, these results indicate that accuracy and bias might be overestimated in current studies. This challenges the prevailing approach to

deception detection research. Thus, we may need to re-examine the “fundamental facts” in deception detection.

Chapter 4: Study 3

4.1 Introduction

Determining whether an individual is deceptive or honest is central to deception detection research. Although all deception detection questions are implicitly considered equivalent, there is considerable variability in terms of the depth and clarity of phrasing of the veracity task given to participants (e.g., O'Sullivan & Ekman, 2004). As survey researchers are keenly aware, even minor changes in the format or wording of a question may affect responses (e.g., Payne, 1951). Furthermore, similar to other areas in psychology (for full discussion see Wiggins & Christopherson, 2019), deception researchers' tendency to omit a full description of their task in their articles may threaten the replicability of findings. One of the ways to address this lack of transparency is to rule out whether there is an effect of question phrasing on deception detection. In this study, I explored how one aspect of the veracity judgment task - the focus of the decision (i.e., the person, details, or the event) - may affect observers' performance.

4.1.1 The Nature of Survey Questions

There is a rich body of literature on survey question formulation and the importance of considering relevant cognitive components (for reviews see Sudman et al., 2010; Tourangeau et al., 2000). The consensus is that responding to a survey question can be broken into several cognitive tasks. First, a respondent must comprehend the question and its contents (Clark & Schober, 1992); thus, a question should be written in simple language that is free of complex terms. The meaning of the question needs to be clear to respondents so that they understand the information that is required in response. For example, asking a question such as, "What did you see?" is far too general and may elicit an unexpected answer (e.g., a description of the respondent's physical surroundings, not

the task-specific stimuli). According to Grice's (1989) maxims of conversation, conversational partners need to make relevant contributions by using contextual information (relational maxim) and determining which information is required (quantity maxim). Ultimately, a respondent needs to comprehend not only the words used, but also the intended meaning of the question.

Once the question is understood, the respondent must recall relevant information. Importantly, the information that is recalled is likely intertwined with the context in which it is asked (e.g., event during the experiment vs. events earlier that day; Bodenhausen & Wyer, 1987). In a deception detection task, however, participants are required to complete an evaluative judgment about a stranger. One way in which they can decide whether someone is deceiving is by holistically assessing them against a stereotypical representation of a lie-teller. The retrieval of information here is complex for several reasons. First, they must continually compare their prototype of a lie-teller and any relevant deception cues to the stimuli provided (i.e., the current target of the decision). Next, observers must maintain their decisions and all relevant information in memory as they typically make veracity decisions *after* viewing an entire account. Thus, the task of deception detection requires memory retrieval.

When the respondent comes to a decision, that response must be communicated. In some situations, respondents can freely convey their responses (e.g., during free recall). Most often, however, they have to alter their responses to fit the required format (e.g., multiple choice questions, rating scale). Respondents may also change their responses because of social influence or desirability. This is more likely to occur in face-to-face contexts than on a survey (Schwarz, 2008), however.

On the whole, there are important factors to consider when designing a survey. One of these factors has been extensively analyzed by psychologists: question phrasing. Across domains – from eyewitness memory to investigative interviewing – there has been an interest in the ways in which question phrasing can be used to guide responses (e.g., Aldridge & Cameron, 1999; Loftus & Zanni, 1975; Rogers et al., 2017). In their seminal work, Loftus and Palmer (1974) found that a mere change in a single word in the question, "About how fast were the cars going when they ___ each other?" (p. 586; e.g., hit vs. bumped vs. smashed) influenced the subsequent speed estimates. Not only does question phrasing cause participants to alter their responses, but it may even impact their memories for an event (i.e., more participants reported seeing broken glass when 'smashed' was used). In fact, there is an entire area of forensic psychology dedicated to examining the effect of question phrasing on memory (i.e., the misinformation effect; e.g., Gabbert et al., 2012; Sharman & Powell, 2012).

4.1.2 Deception Detection Tasks in the Literature

Deception detection researchers have only recently begun to study phrasing effects. This exploration, however, has largely been focused on the questions posed to witnesses (e.g., Hartwig et al., 2005; Lancaster et al., 2013; Minson et al., 2018). For example, researchers have assessed differences between lie- and truth-telling witnesses' responses in interviews featuring free recall versus specific questions (Deeb et al., 2017). Even in seminal meta-analyses on deception detection (e.g., Bond & DePaulo, 2006; DePaulo et al., 2003; Hartwig & Bond, 2011; Palena et al., 2021), there have not been any analyses regarding question phrasing. The questions asked of *observers* have received comparably little empirical attention.

Question phrasing seems to impact decision-making, but it has not been examined explicitly (Reed, 1980). There have been some studies in which observers were asked to focus on a specific aspect of the interview. Most notably, observers have been asked to determine whether the *person* speaking was a lie-teller or truth-teller (e.g., Leal et al., 2010; Vrij, 1993). When judging lie-telling witnesses, laypeople and law enforcement personnel's accuracy was below chance (i.e., 49% and 46%, respectively). In other studies, observers have been tasked with attending to an account about an *event* or *scene* and reporting whether it was 'true' or 'false' using a tool to assess memories [i.e., Reality Monitoring (RM); Johnson & Raye, 1981]. Results from a meta-analysis on RM suggested strong discriminability between veracity groups, but only when witnesses experienced events that they described rather than events they witnessed or viewed on video (Masip et al., 2005). Observers have also been asked to pay attention to specific cues when making their judgments (e.g., analyze the presence of details; Evans et al., 2013). When observers used *details*, amongst other cues to deception, they were better able to discriminate between lie- and truth-tellers than observers who provided a holistic veracity judgment. In another study, observers were found to be less accurate when asked to determine whether a person was being deceptive or telling the truth than when they were asked to monitor changes in a witness's behavior and speech (i.e., cues; Hart et al., 2009). Although few studies report the focus of their deception detection question, there is clear variability in question phrasing. Furthermore, there are differences in results across these studies; however, question phrasing – namely, the objects of decision-making – has not been explicitly examined in deception detection research.

4.1.3 The Present Study

In the deception detection literature, the exact phrasing of the veracity task is either not provided or researchers indicate that observers were asked to determine whether the *person* was deceiving or telling the truth (e.g., Leal et al., 2010). Deception detection decisions in the laboratory and the field, however, are not always about the person's credibility as a whole. As evidenced by Study 2, time-sensitive (i.e., dynamic) decision-making might differ from holistic judgments (i.e., a post hoc deception task). We know that in other domains, a single word may change a person's response (e.g., Loftus & Palmer, 1974), but it is unknown whether the same is true in deception detection. Thus far, deception researchers have used numerous terms as the focus of the veracity question interchangeably: person (e.g., Leal et al., 2010), event/scene (e.g., Wright et al., 2001) or details (e.g., Evans et al., 2013). It appears as though there are differences in decision-makers' performance across studies, but it is unclear whether this can be attributed to question phrasing as that has not be explicitly examined. Thus, the goal of the present study was to determine how these types of differences in question phrasing affect deception detection.

I conducted a typical deception detection study in which observers were asked to view witnesses who were lying and telling the truth. However, I randomly assigned observers to focus on either the person, details, or event when detecting deception (emulating phrasing that has been used in the literature). Thus, I explored whether there was empirical support for researchers' current practice of implicitly regarding all types of questions as equivalent.

4.1.2.1 Hypotheses. Because details are known to be diagnostic of deception (e.g., Vrij, 2008), I hypothesized that observers would be better able to discriminate between lie- and truth-tellers when they focused on details compared to when they focused on a person (i.e., the witness) or the event described.

I also expected that observers would exhibit a truth-bias (i.e., tendency to assume people are telling the truth; Bond & DePaulo, 2006). It was unknown whether an observer would be (more or less) reluctant to accuse a person of being a liar than lying about a detail or event. Thus, analyses pertaining to bias were exploratory.

Finally, I anticipated that observers would be more confident in their judgments of truth-tellers than lie-tellers (e.g., Masip et al., 2006). I also conducted an exploratory analysis of the effect of phrasing on confidence, as this had not been examined in previous literature.

4.2 Method

4.2.1 Research Design

I used a 2 (Veracity: lie vs. truth) x 3 (Question phrasing: individual vs. detail vs. event) mixed-factors design, with veracity as a within-subjects variable. Specifically, observers viewed videos of both lie-tellers and truth-tellers, and then were prompted with one of three randomly assigned question phrasings.

4.2.2 Participants

An a priori power analysis[8] (conducted using G*Power; Faul et al., 2007) revealed that 251 observers were required. After excluding 31 participants for failing to complete

[8] To achieve 95% power when using a repeated measures ANOVA with 3 groups and 2 measurements to detect a medium effect size (Cohen's d = .50)

the study in a timely manner (i.e., they did not complete the study or took longer than 90 minutes to complete the study), the final sample consisted of 255 participants. Again, the sample size was slightly larger than required due to the nature of the participant pool: once observers signed up for the study, they had to be allowed to participate. The majority of observers were females (N = 147, 57.6%) and the average age was 20.74 years (SD = 4.26). Observers self-identified as Arab (9.0%), Black (8.2%), Chinese (4.3%), Filipino (2.0%), Latin American (1.2%), South Asian (30.2%), South East Asian (6.3%), White (31.4%), and Other (7.1%). They completed the study in exchange for bonus course credit.

4.2.3 Materials

4.2.3.1 Video Footage. A subset of 10 videos (i.e., 5 truth-tellers, 5 lie-tellers) from Study 2 were used in this study. The average length of a video was 243.50 s (SD = 56.61). The average age of witnesses in the videos was 19 (SD = .94), gender was equally represented (i.e., there were 5 males and 5 females), and witnesses self-identified as the following races: Arab (20%), Black (10%), South Asian (10%), South East Asian (20%), White (30%), and Other (10%).

4.2.3.2 Deception Detection and Confidence Task. The deception detection task was phrased in one of three ways. Observers were presented with one of the following statements to complete: “The details provided were:”, “The person was telling:”, The event described was:”. The two response options – “a lie” or “the truth” – were identical across conditions. Finally, all observers were asked to rate their confidence in their decisions, from 0% to 100%.

4.2.3.3 Demographics Questionnaire. Observers were asked to provide demographic information, including their age, gender, and race.

4.2.4 Procedure

Observers accessed the Qualtrics study online through the University's participant pool (Sona Systems, n.d.; https://www.sona-systems.com). First, they read the consent form and acknowledged their agreement to participate. Then, they were randomly assigned to receive one of the three phrasings of the deception task. In all conditions, observers were shown a series of 10 video clips of witnesses either being deceptive or telling the truth about what they witnessed. Observers completed the Deception Detection and Confidence Task after each video. At the end of the session, observers were asked to provide demographic information (i.e., age, gender, ethnicity). All of the observers were debriefed at the end of the study. On average, participants spent 60 minutes completing the study.

4.3 Results

Preliminary analyses incorporating video order and demographic information (i.e., gender, ethnicity, age) were conducted to eliminate potential covariates. The following analyses were collapsed across these variables because there were no significant effects (all ps $\geq$.123).

For all nonsignificant statistical results below, I calculated a Bayes Factor (BF_{01}; calculated with a Cauchy prior of .707 using Rouder et al., 2009). This measure helps researchers understand the relative strength of the results as evidence for two opposing theories (i.e., null and alternative hypotheses; Dienes, 2014).

4.3.1 Signal Detection Theory

As in Study 2, I examined the data using Signal Detection Theory (SDT; Green & Swets, 1966). Again, I used Stanislaw and Todorov's (1999) formulas for discrimination (i.e., $d' = Z_{Hit} - Z_{False\ Alarm}$) and decision bias (i.e., $c = -[(Z_{Hit} + Z_{False\ Alarm})]/2$). When observers correctly identified lie-tellers, their decisions were categorized as 'hits'; decisions to incorrectly classify truth-tellers as lie-tellers were categorized as 'false alarms.' Standard corrections were applied [i.e., .5/signal or noise or (signal or noise - .5)/signal or noise] to all extreme values (i.e., 0s and 1s). There was an even distribution across both veracity conditions: 5 noise and 5 signal stimuli. Thus, hit rates and false alarms of 0 were changed to .1, whereas hit rates and false alarms of 1 were changed to .9.

4.3.1.1 Discrimination (*d'*). I conducted a one-way ANOVA on observers' ability to discriminate between lie-tellers and truth-tellers. There were no differences in sensitivity to veracity between the person (M = .50, SD = .77), event (M = .46, SD = .71), or details conditions (M = .54, SD = .82), $F(2, 254) = .27$, $p = .764$, $\eta^2 = .00$, 95% CI [.00, .02], $BF_{01} = -1.22$. The corresponding Bayes factor indicated that there was strong evidence for the null: there was no effect of question phrasing on discrimination. Next, I used one-sample *t*-tests to compare observers' discrimination scores to "0" (i.e., no discrimination). Observers were able to discriminate between lie-tellers and truth-tellers in all conditions (Table 2).

4.3.1.2 Bias (*c*). I also conducted a one-way ANOVA on observers' biases. There were no significant differences between the person (M = .33, SD = .53), event (M = .18, SD = .61), and details conditions (M = .28, SD = .43), $F(2, 254) = 1.78$, $p = .171$, $\eta^2 =$

.01, 95% CI [.00, .05], BF_{01} = -.28. According to the Bayes factor, these results provide strong evidence for the lack of differences in bias between condition. I did, however, find that there was a truth-bias across all conditions when I compared observers' bias scores to "0" (i.e., no bias; Table 2).

Table 2

Results of one-sample t-tests comparing observers' discrimination (d') and bias (c) scores to 0

		Discrimination (*d'*)				Bias (*c*)			
Condition	*n*	*t*	*p*	*D*	95% CI	*t*	*p*	*d*	95% CI
Person	85	5.99	<.001	.65	[.41, .88]	5.71	<.001	.62	[.39, .85]
Event	85	5.93	<.001	.64	[.41, .88]	2.73	.008	.30	[.08, .51]
Detail	85	6.06	<.001	.66	[.42, .89]	6.02	<.001	.65	[.42, .89]

Note. CI = confidence interval.

4.3.2 Confidence

I conducted a Veracity x Question phrasing mixed-factors ANOVA on observers' confidence. I found a main effect of veracity, $F(2, 254) = 11.03$, $p = .001$, $\eta^2 = .08$, 95% CI [.03, .15]. Observers were more confident when they judged truth-tellers (M = 72.76, SD = 13.12) compared to lie-tellers (M = 70.48, SD = 12.78). I did not find an effect of question phrasing, $F(2, 254) = .08$, $p = .923$, $\eta^2 = .00$, 95% CI [.00, .01], BF_{01} = -1.44, nor did I find an interaction between veracity and question phrasing, $F(2, 254) = .21$, $p = .813$, $\eta^2 = .00$, 95% CI [.00, .02], BF_{01} = -1.29. The corresponding Bayes factors indicate strong support against the impact of question phrasing or the interaction between question phrasing and veracity on confidence. That is, there is strong evidence that phrasing did not affect confidence.

4.4 Discussion

In this study, I set out to examine the effect of question phrasing on observers' discrimination and confidence. Overall, my hypotheses related to standard findings in the deception detection literature were supported (Bond & DePaulo, 2006; Masip et al., 2006): there was higher confidence with judgments of truth-tellers than lie-tellers, observer sensitivity to veracity, and a prevalent truth-bias.

I did not, however, find support for my main hypothesis. I expected observers in the detail-focused condition to be better able to discriminate between lie-tellers and truth-tellers than observers in the event or person conditions. However, there were no differences between conditions. One way to interpret this non-significant result is to assess the associated Bayes factor. Given that the Bayes factor in the current study was close to zero (i.e., -1.22), I can conclude that there was strong evidence for the null hypothesis: changing the focus of the deception detection task did not affect observers' sensitivity to veracity. It is important to note that observers in all conditions proved to be able to discriminate between lie-tellers and truth-tellers. This finding implies that it is not that observers were insensitive to the differences between lie-tellers and truth-tellers in general, but rather that a change in question phrasing did not impact their ability to discriminate between the groups.

In keeping with Bond and DePaulo (2008), I expected that observers would exhibit a truth-bias. I found support for this hypothesis across all conditions. That is, regardless of question phrasing, observers tended to assume that people were telling the truth. In terms of differences in bias between conditions, I did not have a specific hypothesis. I found no difference in response bias between the conditions and this finding was supported by a Bayes Factor analysis. This result is unsurprising given that a truth-bias is

found across deception detection studies (Bond & DePaulo, 2008) even though instructions may vary.

I also explored the impact of question phrasing on observers' confidence. As expected, observers were more confident in their judgments of truth-tellers than lie-tellers (e.g., Masip et al., 2006). I also conducted an exploratory analysis of the effect of phrasing on confidence, as this had not been examined in previous literature, but I did not find an effect of question phrasing. My Bayes Factor analysis indicated that this test provided strong evidence that phrasing did not affect confidence. Like observers' discrimination, their confidence was likely not affected by phrasing as they did not attend to the change in wording. In general, laypeople and professionals (e.g., law enforcement personnel) tend to be highly confident in their judgments, with increases in confidence occurring in response to training or repeated exposure (e.g., Meissner & Kassin, 2002). However, my observers were not trained, and were only asked to make a typical veracity judgment as they were not alerted to the phrasing manipulation. It is, thus, unsurprising that observers' confidence was unimpacted across phrasing conditions.

The null results may be explained using the literature on survey formulation and conversational maxims. Recall that comprehension and wording are central to effective question creation (e.g., Loftus & Palmer, 1974; Payne, 1951). Also, placing excessive emphasis on the semantic or literal meaning of words within a question is problematic (Sudman et al., 2010). When people understand language, it is not their comprehension of individual words that matters, but their interpretation of the message's overall meaning (Clark & Schober, 1992). Thus, observers may have not attended to the experimentally relevant focus of the question and simply treated all variations as a request to render a

veracity judgment. A similar explanation of the non-significant results lies in Grice's (1989) maxims of conversation, specifically the maxim of manner. People assume the most obvious meaning of a question. It is, therefore, possible that observers did not carefully attend to the focus of the question (i.e., the details, event, or person) semantically, but rather, thought that the meaning was implied – "Is this a truth or lie?"

4.4.1 Limitations and Future Directions

There were limitations that arose from the chosen methodology. In particular, participants were not alerted to the importance of the wording of the veracity question. In studies where wording is paramount to the research question (e.g., Pickel et al., 2013; Wilson et al., 2013), researchers have added specific instructions for participants to pay close attention to the way that the question is phrased. Whether such an instruction would alter my results could be examined further. I chose, however, not to alert observers to question phrasing as that approach is more representative of typical deception detection studies. Thus, I was able to investigate whether phrasing affects observers without artificially inflating the effect.

Another limitation of this study was that it was unclear whether observers attended to the questions' phrasing. My non-significant findings may have been caused by a failed manipulation. In a review of methods to improve surveys, Collins (2003) suggested that cognitive testing is important in the development of a survey. Thus, future researchers could test whether participants noticed the question phrasing. Participants could be asked to talk through the deception detection task as they completed it, for example (Ericsson & Simon, 1980). In this study, however, whether participants explicitly attended to the phrasing was not central to the research question. My intent was to replicate the minute

changes in question phrasing present in typical deception detection studies to determine if these subtle variations led to differences in performance. They did not.

Finally, this research was conducted at the beginning of the COVID-19 pandemic and the results may have been impacted by the use of a modified experimental paradigm. I had originally proposed in-person data collection to mimic the conditions of typical deception detection studies; instead, the study had to be conducted online. Thus, it is possible that null effects were, in part, due to this change. In a recent presentation about the impact of COVID-19 on research, Maeder (2022) stated that there were stark differences in results when the same study was conducted in different modalities. In her work, null effects were only found in the online study, similar to what I observed here. In a chapter about the drawbacks and benefits of online research, Finley and Penningroth (2015) list a multitude of potential problems reported in online (vs. laboratory) experimentation, most important of which was the potential for poor data quality. Taken together, this implies that participants may have been less likely to attend to phrasing. Consequently, future researchers may consider replicating this study in an in-person context.

4.4.2 Implications

This work makes an important contribution to the deception detection literature. To date, the deception detection task has varied across studies and between laboratories (e.g., Elliott & Leach, 2016; Vrij et al., 2008). The goal of this study was to determine whether this inconsistency might account for the variability in findings across the literature. With strong support from additional statistical analyses (i.e., Bayes factor analyses), my findings indicate that altering the phrasing in terms of three foci did not make a

difference. That suggests that phrasing is not a significant source of variability in accuracy and bias across studies. On the whole, there is a broader implication that people seem to be insensitive to phrasing in the deception detection (vs. other) literature. Perhaps this is due to the nature of deception. When people are asked to discern lie-tellers from truth-tellers they may not dwell on the phrasing of the question because they are familiar with the task, as it is part of their everyday life (e.g., Camden et al., 1984; DePaulo & Kashy, 1998).

Nonetheless, psychological research faces a replication crisis (Wiggins & Christopherson, 2019). Although I did not find support for the hypothesis that wording changes in the deception detection task impact discrimination, continued ambiguity around methodology is problematic. Deception research cannot be replicated, reexamined, and extended if the methodology is not prescriptive. If researchers are not aware of methodological differences, then they can make errors in interpretation. For example, if two laboratories examine the effect of details on deception detection but do not disclose the way in which they asked participants to report that metric (e.g., open-ended question vs. Likert scale), then they may (mis)attribute the results to the variables of interest rather than underlying methodological variability. More broadly, insufficient transparency from each individual laboratory might be responsible for the lack of standardized methodology in deception detection research overall. Specifically, deception research varies in terms of methodological choices, such as paradigms (e.g., Evans et al., 2017), types of questioning (e.g., Vrij et al., 2008, 2010), response options (e.g., Albrechtsen et al., 2009; Leal et al., 2015), or measure of deception (e.g., Evans et al.,

2013; Palena et al., 2021). Thus, the field would benefit from greater emphasis on the reporting of methodology and reproducibility.

4.5 Conclusion

In this study, I found that deception detection was not affected by the phrasing of the deception detection task question. It was not that the stimuli were not representative of those used in the field, as I replicated previously established findings (e.g., higher accuracy when judging truth-tellers than lie-tellers; Bond & DePaulo, 2006; Masip et al., 2006). Rather, a slight change in the target of deception detection did not affect performance. It is important to note that just because something is assumed in the field does not mean that it should not require empirical validation. My findings suggest that question phrasing – as it was manipulated here – is unlikely to be a source of variability between deception detection studies. Yet, on the whole, there is a need to be explicit about the wording of the deception detection task in the interests of replication and open science.

Chapter 5: General Discussion

There have been decades of research on deception, yet its detection continues to be regarded as a difficult task with chance-level accuracy (Bond & DePaulo, 2006). I postulated that performance rates might be impacted by previously unexplored components of the deception detection task. In Study 1, I used a multi-method approach to explore whether a false dichotomy was created by labelling participants either 'lie-tellers' or 'truth-tellers' (i.e., whether their accounts were exclusively honest or deceptive). Next, I examined the notion that observers' decisions and biases change over time (i.e., throughout an account) using a novel dynamic approach to measuring deception detection. Finally, in Study 3, I varied the focal element of the deception detection task to determine whether question phrasing impacted veracity decisions.

There is an implicit assumption within the deception detection literature that lie-tellers and truth-tellers are inherently different from one another (e.g., Nahari et al., 2019). In turn, researchers might have underestimated the intrinsic overlap between the groups. After all, individuals who enter the laboratory to participate in a study are not liars (i.e., not everything they utter is fabricated) but, rather, people who are simply instructed to lie about something specific (e.g., an opinion, a behavior, an event). I examined the proportions of accurate, inaccurate, and omitted details in the accounts of witnesses who were assigned to a veracity group. Overall, witnesses' self-reports revealed that both veracity groups incorporated truthful and deceptive elements into their accounts. Next, I tested whether witnesses' perceptions of their utterances were reflected in their accounts. I found that accurate information was provided by both lie-tellers and truth-tellers, inaccuracies were equally used by both groups, and lie-tellers tended to

employ omissions more than truth-tellers. Although researchers had previously focused on the quantity of details (e.g., Palena et al., 2021), I shifted the focus to the verified quality of the details (e.g., presence, accuracy). My results suggest that the types of details (i.e., accurate or omitted) used may reveal the veracity of an account.

If witnesses' accounts are fraught with both honest and deceptive details, then observers' decisions might evolve through the course of an account. For this reason, I investigated whether observers' decisions change over time in my second study. I used a novel approach to aid observers in assessing deception: a dynamic sliding scale. When observers were given the opportunity to record their decisions, they changed their minds up to 200 times during a single account. Whereas, in previous studies, observers made holistic decisions by indicating their judgment after viewing an account (e.g., ten Brinke et al., 2014; Vrij et al., 2010), participants in the dynamic condition in my study were able to record their decisions moment-to-moment. This more naturalistic approach was not related to an increased ability to discriminate between lie- and truth-tellers, however. Perhaps because observers were suspicious, the freedom to render a decision throughout an account may have compelled them to change their minds often. I also found that the well-established truth-bias (Bond & DePaulo, 2006) disappeared when observers were told that they could change their minds. These findings call into question whether the standard deception detection task itself (i.e., post hoc decision-making) is responsible for discrimination rates and truth-biases. Consequently, previous research might be overestimating observers' performances and underlying biases.

In the last study, I examined the effects of another methodological choice: the focus of the deception detection task. Observers were assigned to one of three phrasing

conditions, in which they assessed the deception of either the witness (i.e., lie- or truth-teller), details, or the event. I found that phrasing did not have an effect on observers' discrimination nor bias. This suggests that phrasing might not be a significant source of variability among observers. By comparing research on instructions and question formulation in other fields (e.g., Gabbert et al., 2012; Sudman et al., 2010), it became apparent that the most central aspect of deception detection research – the veracity question – was rarely explicitly defined. Thus, the field as a whole may need to improve its reporting standards.

5.1 Limitations

There are relevant overarching limitations to this dissertation beyond those individually discussed in each study. The first limitation is related to the paradigm used in these studies. Typically, deception detection studies are comprised of two phases: stimuli acquisition and deception detection. The researcher is tasked with deciding which deception-eliciting paradigm they will use. For this study, I created stimuli with the suspicious event paradigm because it had successfully been used in other studies (i.e., there were significant differences between veracity groups; Elliott & Leach, 2016). The paradigm might have afforded certain types of responses in both witnesses and observers (e.g., contained an unnatural amount of variability). In a study where deception detection paradigms were used as covariates to conduct deception-relevant analyses, however, paradigm type did not have an impact on the findings (Evans et al., 2017). Thus, it is unlikely that the results in this dissertation are exclusively generalizable to studies conducted using the suspicious event paradigm.

Another limitation is that deception was manufactured in these studies. That is, instead of recording and observing witnesses who deceived by their own volition, participants either lied or told the truth according to their randomly-assigned condition. An argument can be made that laboratory lies are not naturalistic. Although witnesses were asked to deceive the experimenter about what they had seen, I did not limit the deception process by prescribing a strategy (e.g., denial of guilt) or artificially encouraging differences between veracity groups using known methods (e.g., directing participants to divulge details). In fact, Hartwig and Bond's (2011) meta-analysis revealed that deception is equally detectable across a multitude of situations, including high-stakes (e.g., deposition for a crime) and laboratory lies.

5.2 Implications

I explored several "known facts" within the literature and found that they were not infallible. Observers' truth-biases and slightly above average accuracy (i.e., 54%) have been central to deception detection research (e.g., Bond & DePaulo, 2006). However, my research revealed that the methodological components underlying deception detection studies may have been responsible for these results.

In the first study, there was an overlap in the deceitful and honest information provided in the accounts of lie-tellers and truth-tellers. Deception detection studies are often conducted by assigning participants to veracity conditions with rarely any direction beyond, "tell the truth" or "lie" (e.g., Elliott & Leach, 2016). Although this practice is meant to mimic realistic situations in which people deceive, there is typically no information provided about the extent to which participants should lie, or the strategy they should use. My research suggests that, when behaving naturalistically, people cannot

uniformly be called 'liars'[9] or 'truth-tellers' simply because of random assignment to condition. For example, lie-tellers may provide inaccurate information or omit certain details, but it is unlikely that their entire accounts will be fabricated. In fact, I found that a compelling proportion of lie-tellers' accounts contained accurate details. Thus, Study 1 highlights the potential disconnect between labelling a participant with a veracity condition and their account. I found that there is a lot of variability within veracity groups in terms of the quality of information. Perhaps researchers are not finding many behavioral differences between lie- and truth-tellers because the amount of false information provided is the same across veracity groups. Furthermore, observers may be poor detectors of deception because they are looking for someone who fabricated their entire account – my results imply that this rarely occurs. Similarly, professional deception detectors are likely not met with individuals who fabricate every detail. Law enforcement officers, for example, may need to recalibrate their expectations regarding lie-tellers. My findings do suggest that lie-tellers' use of omissions and accurate details is different from truth-tellers'. Thus, future researchers may study ways for officers to strategically exploit these differences to improve deception detection.

Modifying the deception detection task can affect performance. Study 2 was created to more closely mimic the experiences of decision-makers who made veracity judgments on a daily basis (e.g., police officers, border patrol agents, judges). Making moment-to-moment decisions – as opposed to holistic verdicts – about a person's veracity lowered discrimination rates and eliminated response biases. This finding suggests that the

[9] Recall that it is common practice within deception detection literature to call individuals who had been assigned to tell a lie in a study: liar(s).

experiences of professional deception detectors – and even laypersons – might not be represented in the typical deception detection literature. These results have strong implications for the field of deception detection. First, the methodology that observers use to render a decision may need to be re-evaluated. Second, a more sophisticated method of assessing dynamic decision-making should be created to assess the nature of variability. Finally, and most consequentially, previous results may need to be called into question.

Varying the phrasing of the deception detection task did not impact discrimination or bias. Yet, the implications from the third study are three-fold. First, question phrasing does not appear to be responsible for the variability between study findings. This revelation indicates that phrasing should not be a significant concern for researchers, and thus, comparisons can be made across studies. Second, the impact of varying phrasing within the criminal justice system remains unknown. I examined laypersons' decision-making in an online deception detection task, but there are a multitude of situations in which individuals are asked similar questions in significant contexts, such as officers during an interrogation, jury members, or border patrol agents. My findings suggest that the phrasing of a deception task does not impact outcomes; however, this remains an empirical question in practice. Third, the documentation and reporting of the methodologies used in deception detection studies needs to be improved. The deception detection literature, in particular, lacks transparency regarding methodological choices, such as detailed paradigm descriptions. Even in seminal meta-analyses in deception detection (Bond & DePaulo, 2006; DePaulo et al., 2003; Hartwig & Bond, 2011; Palena et al., 2021), there have been no analyses regarding question phrasing. That is likely

because this critical methodological information is not often provided by researchers. For replication to be possible, and for the study of deception detection to remain scientific, it is fundamental for researchers to know the exact methodology used to study a phenomenon (e.g., details regarding the veracity question).

Chapter 6: Conclusion

Using novel methodological approaches, I explored components underlying decision-making during deception detection. On the whole, the findings indicate that deception and its detection have not been fully captured in the literature. At least some of what we "know" about deception may, in fact, be wrong. These findings have significant implications for the study of deception detection and practice, and highlight the need for the reform of data gathering methodology and increased transparency in research overall.

References

Aamodt, M. G., & Custer, H. (2006). Who can best catch a liar? A meta-analysis of individual differences in detecting deception. *The Forensic Examiner*, *15*(1), 6–11.

Abeler, J., Nosenzo, D., & Raymond, C. (2019). Preferences for truth-telling. *Econometrica*, *87*(4), 1115–1153. https://doi.org/10.3982/ECTA14673

Albrechtsen, J. S., Meissner, C. A., & Susa, K. J. (2009). Can intuition improve deception detection performance? *Journal of Experimental Social Psychology*, *45*(4), 1052–1055. https://doi.org/10.1016/j.jesp.2009.05.017

Aldridge, J., & Cameron, S. (1999). Aldridge and Cameron 1999.pdf. *Applied Developmental Science*, *3*(2), 136–147. https://doi.org/10.1207/s1532480xads0302_7

Alpert, G. P., Dunham, R. G., Stroshine, M., Bennett, K., & MacDonald, J. (2006). *Police officers' decision making and discretion: Forming suspicion and making a stop* (No. 213004; pp. 1–151). U.S. Department of Justice.

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. https://doi.org/10.1016/j.ejpal.2014.11.002

Amado, B. G., Arce, R., Fariña, F., & Vilariño, 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. https://doi/org/10.1016/j.ijchp.2016.01.002

Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin*, *111*(2), 256–274. https://doi.org/10.1037/0033-2909.111.2.256

Aryadoust, V. (2020). Dynamics of item reading and answer changing in two hearings in a computerized while-listening performance test: An eye-tracking study. *Computer Assisted Language Learning*, *33*(5–6), 510–537. https://doi.org/10.1080/09588221.2019.1574267

Ask, K., Granhag, P. A., Juhlin, F., & Vrij, A. (2013). Intending or pretending? Automatic evaluations of goal cues discriminate true and false intentions. *Applied Cognitive Psychology*, *27*, 173–177. https://doi.org/10.1002/acp.2893

Baron, J. (2004). Normative models of judgment and decision making. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 19–36). Blackwell Publishing. https://doi.org/10.1002/9780470752937.ch2

Battigalli, P., Charness, G., & Dufwenberg, M. (2013). Deception: The role of guilt. *Journal of Economic Behavior & Organization*, *93*(1), 227–232. https://doi.org/10.1016/j.jebo.2013.03.033

Beckmann, J. F. (2010). Taming a beast of burden – On some issues with the conceptualisation and operationalisation of cognitive load. *Learning and Instruction*, *20*(3), 250–264. https://doi.org/10.1016/j.learninstruc.2009.02.024

Benjamin, L. T., Cavell, T. A., & Shallenberger, W. R. (1984). Staying with initial answers on objective tests: Is it a myth? *Teaching of Psychology*, *11*(3), 133–141. https://doi.org/10.1177/009862838401100303

Black, D., & Reiss, A. (1967). *Patterns of behavior in police and citizen transactions* (Studies in Crime and Law Enforcement in Major Metropolitan Areas). U.S. Department of Justice.

Bodenhausen, G. V., & Wyer, R. S. (1987). Social cognition and social reality: Information acquisition and use in the laboratory and the real world. In H. J. Hippler, N. Schwarz, & S. Sudman (Eds.), *Social information processing and survey methodology* (pp. 6–41). Springer Verlag. https://doi.org/10.1007/978-1-4612-4798-2_2

Bond, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review*, *10*(3), Article 3. https://doi.org/10.1207/s15327957pspr1003_2

Bond, C. F., & DePaulo, B. M. (2008). Individual differences in judging deception: Accuracy and bias. *Psychological Bulletin*, *134*(4), 477–492. https://doi.org/10.1037/0033-2909.134.4.477

Bonner, H. S. (2015). Police officer decision-making in dispute encounters: Digging deeper into the 'black box.' *American Journal of Criminal Justice*, *40*(3), 493–522. https://doi/org/10.1007/s12103-014-9274-2

Brandt, D. R., Miller, G. R., & Hocking, J. E. (1980). Effects of self-monitoring and familiarity on deception detection. *Communication Quarterly*, *28*(3), 3–10. https://doi/org/10.1080/01463378009369370

Brown, S. G., & Daus, C. S. (2015). The influence of police officers' decision-making style and anger control on responses to work scenarios. *Journal of Applied*

Research in Memory and Cognition, *4*(3), 294–302. https://doi.org/10.1016/j.jarmac.2015.04.001

Buchanan, L., & O'Connell, A. (2006). A brief history of decision making. *Harvard Business Review*, *84*(1), 32–41.

Camden, C., Motley, M. T., & Wilson, A. (1984). White lies in interpersonal communication: A taxonomy and preliminary investigation of social motivations. *Western Journal of Speech Communication*, *48*(4), 309–325. https://doi.org/10.1080/10570318409374167

Carson, T. L. (2006). The Definition of Lying*. *NOÛS*, *40*(2), 284–306. https://doi.org/10.1111/j.0029-4624.2006.00610.x

Chisholm, R. M., & Feehan, T. D. (1977). The Intent to Deceive. *The Journal of Philosophy*, *74*(3), 143–159. https://doi.org/10.2307/2025605

Choshen-Hillel, S., Shaw, A., & Caruso, E. M. (2020). Lying to appear honest. *Journal of Experimental Psychology: General*, *149*(9), 1719–1735. https://doi.org/10.1037/xge0000737

Clark, H. H., & Schober, M. F. (1992). Questions about questions: Inquiries into the cognitive bases of surveys. In J. M. Tanur (Ed.), *Asking questions and influencing answers* (pp. 15–48). Russell-Sage.

Collins, D. (2003). Pretesting survey instruments: An overview of cognitive methods. *Quality of Life Research*, *12*, 229–238. https://doi.org/10.1023/A:1023254226592

Da Silva, C. S., & Leach, A.-M. (2013). Detecting deception in second-language speakers. *Legal and Criminological Psychology*, *18*(1), 115–127. https://doi.org/10.1111/j.2044-8333.2011.02030.x

Deeb, H., Vrij, A., Hope, L., Mann, S., Granhag, P. A., & Strömwall, L. A. (2018). Police Officers' Perceptions of Statement Inconsistency: *Criminal Justice and Behavior*. https://doi.org/10.1177/0093854818758808

Deeb, H., Vrij, A., Hope, L., Mann, S., Granhag, P.-A., & Lancaster, G. L. J. (2017). Suspects' consistency in statements concerning two events when different question formats are used: Suspects' consistency in statements. *Journal of Investigative Psychology and Offender Profiling*, *14*, 74–87. https://doi.org/10.1002/jip.1464

DePaulo, B. M. (1992). Nonverbal behavior and self-presentation. *Psychological Bulletin*, *111*(2), 203–243. https://doi.org/10.1037/0033-2909.111.2.203

DePaulo, B. M., Charlton, K., Cooper, H., Lindsay, J. J., & Muhlenbruck, L. (1997). The accuracy-confidence correlation in the detecting of deception. *Personality and Social Psychology Review*, *1*(4), 346–357. https://doi.org/10.1207/s15327957pspr0104_5

DePaulo, B. M., & Kashy, D. A. (1998). Everyday lies in close and casual relationships. *Journal of Personality and Social Psychology*, *74*(1), 63–79. https://doi.org/10.1037/0022-3514.74.1.63

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin*, *129*(1), Article 1. https://doi.org/10.1037/0033-2909.129.1.74

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, *5*(781), 1–17. https://doi/org/10.3389/fpsyg.2014.00781

Dillon, S. M. (1998). *Descriptive decision making: Comparing theory with practice*. 1–10.

Ekman, P. (1985). *Telling lies: Clues to deceit in the marketplace, marriage, and politics.* W. W. Norton.

Ekman, P., & Friesen, W. V. (1974). Detecting deception from the body or face. *Journal of Personality and Social Psychology*, *29*(3), 288–298. https://doi.org/10.1037/h0036006

Elaad, E. (2022). Tunnel vision and confirmation bias among police investigators and laypeople in hypothetical criminal contexts. *SAGE Open*, *12*(2), 1–10. https://doi.org/10.1177/21582440221095022

Elliott, E., & Leach, A.-M. (2016). You must be lying because I don't understand you: Language proficiency and lie detection. *Journal of Experimental Psychology: Applied*, *22*(4), Article 4. https://doi.org/10.1037/xap0000102

Ericsson, A. K., & Simon, H. A. (1980). Ericsson and Simon 1980.pdf. *Psychological Review*, *87*(3), 215–251.

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*, 33–41. https://doi.org/10.1016/j.jarmac.2013.02.002

Evans, J. R., Pimentel, P. S., Pena, M. M., & Michael, S. W. (2017). The ability to detect false statements as a function of the type of statement and the language proficiency of the statement provider. *Psychology, Public Policy, and Law*, *23*(3), 290–300. https://doi.org/10.1037/law0000127

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), Article 2. https://doi.org/10.3758/BF03193146

Finley, A. J., & Penningroth, S. L. (2015). Online versus in-lab: Pros and cons of an online prospective memory experiment. In A. M. Columbus (Ed.), *Advances in Psychology Research* (Vol. 113, pp. 135–162). Nova Science Publishers, Inc.

Frank, M. G., & Feeley, T. H. (2003). To catch a liar: Challenges for research in lie detection training. *Journal of Applied Communication Research*, *21*(1), 58–75. https://doi.org/10.1080/00909880305377

Gabbert, F., Hope, L., Fisher, R. P., & Jamieson, K. (2012). Protecting against misleading post-event information with a self-administered interview: Protecting against misinformation. *Applied Cognitive Psychology*, *26*(4), 568–575. https://doi.org/10.1002/acp.2828

Gawrylowicz, J., Fairlamb, S., Tantot, E., Qureshi, Z., Redha, A., & Ridley, A. M. (2016). Does practice make the perfect liar? The effect of rehearsal and increased cognitive load on cues to deception. *Applied Cognitive Psychology*, *30*(2), 250–259. https://doi.org/10.1002/acp.3199

Gneezy, U. (2005). Deception: The role of consequences. *The American Economic Review*, *95*(1), 384–394. https://doi.org/https:/doi/org/10.1257/0002828053828662

Granhag, P. A., & Hartwig, M. (2008). A new theoretical perspective on deception detection: On the psychology of instrumental mind-reading. *Psychology, Crime & Law*, *14*(3), 189–200. https://doi.org/https:/doi.org/10.1080/10683160701645181

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. Wiley & Sons, Inc.

Grice, P. (1989). *Studies in the way of words*. Harvard Univeristy Press.

Hanway, P., Akehurst, L., Vernham, Z., & Hope, L. (2021). The effects of cognitive load during an investigative interviewing task on mock interviewers' recall of information. *Legal and Criminological Psychology*, *26*(1), 25–41. https://doi.org/10.1111/lcrp.12182

Hart, C. L., Fillmore, D. G., & Griffith, J. D. (2009). Indirect detection of deception: Looking for change. *Current Research In Social Psychology*, *14*(9), 135–142.

Hartwig, M., & Bond, C. F. (2011). Why do lie-catchers fail? A lens model meta-analysis of human lie judgments. *Psychological Bulletin*, *137*(4), Article 4. https://doi.org/10.1037/a0023589

Hartwig, M., Granhag, P. A., & Strömwall, L. A. (2007). Guilty and innocent suspects' strategies during police interrogations. *Psychology, Crime & Law*, *13*(2), Article 2. https://doi.org/10.1080/10683160600750264

Hartwig, M., Granhag, P. A., Stromwall, L. A., & Doering, N. (2010). Impression and information management: On the strategic self- regulation of innocent and guilty Suspects. *The Open Criminology Journal*, *3*(2), 10–16. https://doi.org/10.2174/1874917801003020010

Hartwig, M., Granhag, P. A., Strömwall, L. A., & Vrij, A. (2005). Detecting deception via strategic disclosure of evidence. *Law and Human Behavior*, *29*(4), Article 4. https://doi.org/10.1007/s10979-005-5521-x

Inbau, F. E., Reid, J. E., Buckley, J. P., & Jayne, B. C. (2013). *Criminal interrogations and confessions* (5th ed.). Jones & Bartlett Learning.

Jarvis, B. G. (2018). *MediaLab* (2018.1.117). Empirisoft Corporation.

Johnson, M. K., & Raye, C. L. (1981). Reality Monitoring. *Psychological Review*, *88*(1), Article 1.

Kassin, S. M., Drizin, S. A., Grisso, T., Gudjonsson, G. H., Leo, R. A., & Redlich, A. D. (2010). Police-induced confessions: Risk factors and recommendations. *Law and Human Behavior*, *34*(1), 3–38. https://doi.org/10.1007/s10979-009-9188-6

Kassin, S. M., Leo, R. A., Meissner, C. A., Richman, K. D., Colwell, L. H., Leach, A.-M., & La Fon, D. (2007). Police interviewing and interrogation: A self-report survey of police practices and beliefs. *Law and Human Behavior*, *31*(4), 381–400. https://doi.org/10.1007/s10979-006-9073-5

Krix, A. C., Sauerland, M., Gabbert, F., & Hope, L. (2014). Providing eyewitnesses with initial retrieval support: What works at immediate and subsequent recall? *Psychology, Crime & Law*, *20*(10), 1005–1027. https://doi.org/10.1080/1068316X.2014.902456

Lancaster, G. L. J., Vrij, A., Hope, L., & Waller, B. (2013). Sorting the liars from the truth Tellers: The benefits of asking unanticipated questions on lie detection. *Applied Cognitive Psychology*, *27*(1), 107–114. https://doi.org/10.1002/acp.2879

Leach, A.-M., & Da Silva, C. S. (2013). Language proficiency and police officers' lie detection performance. *Journal of Police and Criminal Psychology*, *28*(1), 48–53. https://doi.org/10.1007/s11896-012-9109-3

Leal, S., Vrij, A., Mann, S., & Fisher, R. P. (2010). Detecting true and false opinions: The Devil's Advocate approach as a lie detection aid. *Acta Psychologica*, *134*(3), 323–329. https://doi.org/10.1016/j.actpsy.2010.03.005

Leal, S., Vrij, A., Warmelink, L., Vernham, Z., & Fisher, R. P. (2015). You cannot hide your telephone lies: Providing a model statement as an aid to detect deception in insurance telephone calls. *Legal and Criminological Psychology*, *20*, 129–146. https://doi.org/10.1111/lcrp.12017

Leins, D. A., Fisher, R. P., & Ross, S. J. (2013). Exploring liars' strategies for creating deceptive reports. *Legal and Criminological Psychology*, *18*(1), 141–151. https://doi.org/10.1111/j.2044-8333.2011.02041.x

Leins, D. A., Zimmerman, L. A., & Polander, E. N. (2017). Observers' real-time sensitivity to deception in naturalistic interviews. *Journal of Police and Criminal Psychology*, *32*(4), 319–330. https://doi.org/10.1007/s11896-017-9224-2

Levine, T. R. (2014a). Truth-Default Theory (TDT): A theory of human deception and deception detection. *Journal of Language and Social Psychology*, *33*(4), 378–392. https://doi.org/10.1177/0261927X14535916

Levine, T. R. (2014b). Active deception detection. *Policy Insights from the Behavioral and Brain Sciences*, *1*(1), 122–128. https://doi.org/10.1177/2372732214548863

Levine, T. R. (2019). Five reasons why I am skeptical that indirect or unconscious lie detection is superior to direct deception detection. *Frontiers in Psychology*, *10*, 1354. https://doi.org/10.3389/fpsyg.2019.01354

Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior*, *13*(5), 585–589. https://doi.org/10.1016/S0022-5371(74)80011-3

Loftus, E. F., & Zanni, G. (1975). Eyewitness testimony: The influence of the wording of a question. *Bulletin of the Psychonomic Society*, *5*(1), 86–88. https://doi.org/10.3758/BF03336715

Luke, T. J. (2019). Lessons from Pinocchio: Cues to deception may be highly exaggerated. *Perspectives on Psychological Science*, *14*(4), 646–671. https://doi.org/10.1177/1745691619838258

Maeder, E. (2022, March). *Methodological Considerations in Online Jury Research: Presentation Modality and Data Quality*. 2022 American Psychology - Law Society annual conference, Denver, CO.

Mahon, J. E. (2016). The definition of lying and deception. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition). https://plato.stanford.edu/archives/win2016/entries/lying-definition/

Markowitz, D. M. (2020). The deception faucet: A metaphor to conceptualize deception and its detection. *New Ideas in Psychology*, *59*, 100816. https://doi.org/https:/doi.org/10.1016/j.newideapsych.2020.100816

Marsili, N. (2014). Lying as a scalar phenomenon: Insincerity along the certainty-uncertainty continuum. In S. Cantarini, W. Abraham, & E. Leiss (Eds.), *Certainty-uncertainty—And the attitudinal space between* (pp. 153–173). John Benjamins Publishing Company. https://benjamins.com/catalog/slcs.165.09mar

Masip, J., Garrido, E., & Herrero, C. (2003). When did you conclude she was lying? The impact of the moment the decision about the sender's veracity is made and the sender's facial appearance on police officers' credibility judgments. *The Journal of Credibility Assessment and Witness Psychology*, *4*(1), 36.

Masip, J., Garrido, E., & Herrero, C. (2006). Observers' decision moment in deception detection experiments: Its impact on judgment, accuracy, and confidence. *International Journal of Psychology*, *41*(4), 304–319. https://doi.org/10.1080/00207590500343612

Masip, J., Garrido, E., & Herrero, C. (2009). Heuristic versus systematic processing of information in detecting deception: Questioning the truth bias. *Psychological Reports*, *105*(1), 11–36. https://doi.org/10.2466/PR0.105.1.11-36

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), Article 1. https://doi.org/10.1080/10683160410001726356

Mathews, C. O. (1929). The effect of the order of printed response words on an interest questionnaire. *Journal of Educational Psychology*, *20*(2), 128–134. https://doi.org/10.1037/h0074399

McCornack, S. A. (1992). Information manipulation theory. *Communication Monographs*, *59*(1), 1–16. https://doi.org/10.1080/03637759209376245

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. https://doi.org/10.1023/A:1020278620751

Meissner, C. A., Kelly, C. E., & Woestehoff, S. A. (2015). Improving the effectiveness of suspect interrogations. *Annual Review of Law and Social Science*, *11*(1), Article 1. https://doi.org/10.1146/annurev-lawsocsci-120814-121657

Minson, J. A., VanEpps, E. M., Yip, J. A., & Schweitzer, M. E. (2018). Eliciting the truth, the whole truth, and nothing but the truth: The effect of question phrasing on deception. *Organizational Behavior and Human Decision Processes*, *147*, 76–93. https://doi.org/10.1016/j.obhdp.2018.05.006

Nahari, G. (2018). Reality monitoring in the forensic context: Digging deeper into the speech of liars. *Journal of Applied Research in Memory and Cognition*, *7*(3), 432–440. https://doi.org/10.1016/j.jarmac.2018.04.003

Nahari, G., Ashkenazi, T., Fisher, R. P., Granhag, P.-A., Hershkowitz, I., Masip, J., Meijer, E. H., Nisin, Z., Sarid, N., Taylor, P. J., Verschuere, B., & Vrij, A. (2019). 'Language of lies': Urgent issues and prospects in verbal lie detection research. *Legal and Criminological Psychology*, *24*, 1–23. https://doi.org/10.1111/lcrp.12148

Nahari, G., Vrij, A., & Fisher, R. P. (2014). The verifiability approach: Countermeasures facilitate its ability to discriminate between truths and lies. *Applied Cognitive Psychology*, *28*, 122–128. https://doi.org/10.1002/acp.2974

O'Sullivan, M., & Ekman, P. (2004). The wizards of deception detection. In P. A. Granhag & L. A. Stromwall (Eds.), *The detection of deception in forensic contexts* (pp. 269–286). Cambridge University Press. https://doi.org/10.1017/CBO9780511490071.012

Palena, N., Caso, L., Vrij, A., & Nahari, G. (2021). The verifiability approach: A meta-analysis. *Journal of Applied Research in Memory and Cognition*, *10*(1), 155–166. https://doi.org/10.1016/j.jarmac.2020.09.001

Park, H. S., Levine, T. R., McCornack, S. A., Morrison, K., & Ferrara, M. (2002). Park, Levine, McCornack, Morrison, and Ferrara 2002.pdf. *Communication Monographs*, *69*(2), 144–157. https://doi.org/10.1080/714041710

Payne, S. L. B. (1951). *The art of asking questions: Studies in Public opinion*. Princeton University Press.

Pickel, K. L., Kulig, T. C., & Bauer, H. M. (2013). Judging veracity impairs eyewitnesses' memory of a perpetrator. *Memory*, *21*(7), 843–856. https://doi.org/10.1080/09658211.2013.765486

Porter, S., & Yuille, J. C. (1996). The language of deceit: An investigation of the verbal clues to deception in the interrogation context. *Law and Human Behavior*, *20*(4), 443–458. https://doi.org/10.1007/BF01498980

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59–108.

Reed, R. (1980). Jury simulation: The impact of judge's instructions and attorney tactics on decision making. *The Journal of Criminal Law and Criminology*, *71*(1), 68–72. https://doi.org/10.2307/1142800

Rogers, T., Zeckhauser, R., Gino, F., Norton, M. I., & Schweitzer, M. E. (2017). Artful Paltering: The risks and rewards of using truthful statement to mislead others. *Journal of Personality and Social Psychology*, *112*(2), 456–473. http://dx.doi.org/10.1037/pspi0000081

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*(2), 225–237. https://doi.org/10.3758/PBR.16.2.225

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science*, *300*(5626), 1755–1758. https://doi.org/10.1126/science.1082976

Schwarz, N. (2008). The Psychology of Survey Response. In W. Donsbach & M. Traugott (Eds.), *The SAGE Handbook of Public Opinion Research* (pp. 374–387). SAGE Publications Ltd. https://doi.org/10.4135/9781848607910.n35

Shanks, D. R. (2017). Regressive research: The pitfalls of post hoc data selection in the study of unconscious mental processes. *Psychonomic Bulletin & Review*, *24*(3), 752–775. https://doi.org/10.3758/s13423-016-1170-y

Sharman, S. J., & Powell, M. B. (2012). A comparison of adult witnesses' suggestibility across various types of leading questions: Question structure and misleading information. *Applied Cognitive Psychology*, *26*(1), 48–53. https://doi.org/10.1002/acp.1793

Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics*, *69*(1), 99–118. https://doi.org/10.2307/1884852

Sona Systems. (n.d.). *Sona Systems: Cloud-based participant management software*. Sona Systems, Ltd. https://www.sona-systems.com

Sporer, S. L. (2016). Deception and cognitive load: Expanding our horizon with a working memory model. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00420

Sporer, S. L., & Ulatowska, J. (2021). Indirect and unconscious deception detection: Too soon to give up? *Frontiers in Psychology*, *12*, 601852. https://doi.org/10.3389/fpsyg.2021.601852

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*(1), 137–149. https://doi.org/10.3758/BF03207704

Street, C. N. H., & Richardson, D. C. (2015a). Lies, damn Lies, and expectations: How base rates inform lie-truth judgments. *Applied Cognitive Psychology*, *29*(1), 149–155. https://doi.org/10.1002/acp.3085

Street, C. N. H., & Richardson, D. C. (2015b). Descartes versus Spinoza: Truth, uncertainty, and bias. *Social Cognition*, *33*(3), 227–239. https://doi.org/10.1521/soco.2015.33.2.2

Street, C. N. H., & Richardson, D. C. (2015c). The focal account: Indirect lie detection need not access unconscious, implicit knowledge. *Journal of Experimental Psychology: Applied*, *21*(4), 342–355. https://doi.org/10.1037/xap0000058

Sudman, S., Bradburn, N. M., & Schwarz, N. (Eds.). (2010). *Thinking about answers: The application of cognitive processes to survey methodology*. Jossey-Bass.

Sutter, M. (2009). Deception through telling the truth?! Experimental evidence from individuals and teams. *The Economic Journal*, *119*(534), 47–60. https://doi.org/10.1111/j.1468-0297.2008.02205.x

ten Brinke, L., Stimson, D., & Carney, D. R. (2014). Some evidence for unconscious lie detection. *Psychological Science*, *25*(5), 1098–1105. https://doi.org/10.1177/0956797614524421

Tourangeau, R., Rips, L. J., & Rasinski, K. (Eds.). (2000). *The psychology of survey response*. Cambridge University Press. https://doi.org/10.1017/CBO9780511819322

Tourangeau, R., & Smith, T. W. (1996). Asking sensitive questions: The impact of data collection mode, question format, and question context. *Public Opinion Quarterly*, *60*(2), 275–304.

Verigin, B. L., Meijer, E. H., Bogaard, G., & Vrij, A. (2019). Lie prevalence, lie characteristics and strategies of self-reported good liars. *PLOS ONE*, *14*(12), e0225566. https://doi.org/10.1371/journal.pone.0225566

Verigin, B. L., Meijer, E. H., & Vrij, A. (2020). Embedding lies into truthful stories does not affect their quality. *Applied Cognitive Psychology*, *34*(2), Article 2. https://doi.org/10.1002/acp.3642

Vrij, A. (1993). Credibility judgments of detectives: The impact of nonverbal behavior, social skills, and physical characteristics on impression formation. *The Journal of Social Psychology*, *133*(5), 601–610. https://doi.org/10.1080/00224545.1993.9713915

Vrij, A. (2000). *Detecting lies and deceit: The psychology of lying and the implications for professional practice*. John Wiley & Sons.

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

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. https://doi.org/10.1111/lcrp.12088

Vrij, A., Granhag, P. A., & Porter, S. (2010). Pitfalls and opportunities in nonverbal and verbal lie detection. *Psychological Science in the Public Interest*, *11*(3), Article 3. https://doi.org/10.1177/1529100610390861

Vrij, A., Leal, S., Deeb, H., Chan, S., Khader, M., Chai, W., & Chin, J. (2020). Lying about flying: The efficacy of the information protocol and model statement for detecting deceit. *Applied Cognitive Psychology*, *34*(1), Article 1. https://doi-org./10.1002/acp.3614

Vrij, A., Leal, S., Mann, S., Warmelink, L., Granhag, P. A., & Fisher, R. P. (2010). Drawings as an innovative and successful lie detection tool. *Applied Cognitive Psychology*, *24*, 587–594. https://doi.org/10.1002/acp.1627

Vrij, A., Mann, S. A., Fisher, R. P., Leal, S., Milne, R., & Bull, R. (2008). Increasing cognitive load to facilitate lie detection: The benefit of recalling an event in reverse order. *Law and Human Behavior*, *32*(3), Article 3. https://doi.org/10.1007/s10979-007-9103-y

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. https://doi.org/10.1080/10683160902740633

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. https://doi.org/10.1016/j.newideapsych.2014.03.001

Walczyk, J. J., Igou, F. P., Dixon, A. 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*. https://doi.org/10.3389/fpsyg.2013.00014

Walczyk, J. J., Mahoney, K. T., Doverspike, D., & Griffith-Ross, D. A. (2009). Cognitive lie detection: Response time and consistency of answers as cues to deception. *Journal of Business and Psychology*, *24*(1), 33–49. https://doi.org/10.1007/s10869-009-9090-8

Wiggins, B. J., & Christopherson, C. D. (2019). The replication crisis in psychology: An overview for theoretical and philosophical psychology. *Journal of Theoretical and Philosophical Psychology*, *39*(4), 202–217. https://doi.org/10.1037/teo0000137

Wilson, J. P., Hugenberg, K., & Bernstein, M. J. (2013). The cross-race effect and eyewitness identification: How to improve recognition and reduce decision errors in eyewitness situations. *Social Issues and Policy Review*, *7*(1), 83–113. https://doi.org/10.1111/j.1751-2409.2012.01044.x

Wixted, J. T. (2004). The psychology and neuroscience of forgetting. *Annual Review of Psychology*, *55*(1), 235–269. https://doi.org/10.1146/annurev.psych.55.090902.141555

Woolridge, L. R., Leach, A.-M., & Elliott, E. (2020, March). *Perceptions of interviewees' accents during deception detection*. American Psychology-Law Society Annual Conference, New Orleans, LA.

Wright, D. B., Loftus, E. F., & Hall, M. (2001). Now you see it; now you don't: Inhibiting recall and recognition of scenes. *Applied Cognitive Psychology*, *15*(5), 471–482. https://doi.org/10.1002/acp.719

Wright, P. (1974). The harassed decision maker: Time pressures, distractions, and the use of evidence. *Journal of Applied Psychology*, *59*(5), 555–561. https://doi.org/10.1037/h0037186

Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. In *Advances in experimental social psychology* (Vol. 14, pp. 1–58). Academic Press Inc.

Appendices

Appendix A: Study 1

A1. Deceptive consent

You have volunteered to take part in our research project: Details. We are examining the way individuals tell their accounts. You will be asked to watch a video talk about what you saw. You will also be asked demographic questions (i.e., age, gender, race) about yourself. This study should take about 30 minutes.

This study has no risks, other than potential discomfort associated with being interviewed. You will be given course credit for your help here today and entered into a draw for $50. You may still withdraw from this study (i.e., tell the experimenter that you do not want to take part); however, you must do so before you leave here today. If you decide not to take part in the study, all of your data (e.g., video, forms) will be deleted/shredded and you will still be given course credit.

Data will be stored on password-protected computers, on external hard drives, and in cabinets in a locked room. Only the primary researchers (Dr. Amy Leach and Elizabeth Elliott), their assistants, and other researchers interested in social psychology will have access to the data (e.g., for meta-analyses). Your individual performance will remain confidential and will not be released to professors, employers, or in publications; only group results will be reported (e.g. conference presentations, journal articles).

During this study, a video camera will be used to record your interview with the experimenter. We use recordings of individuals in different studies. These videos will be used in other studies and may be presented at conferences and in classes or shared with other psychology researchers.

This study has been approved by the UOIT Research Ethics Board (REB #14220) on March 9th, 2017. If you have any questions about this project, please contact the principal investigator, Dr. Amy Leach (amy.leach@uoit.ca). Any questions regarding your rights as a participant, complaints or adverse events may be addressed to the Research Ethics Board through the Ethics and Compliance Officer - reseachethics@uoit.ca or 905.721.8668 x. 3693.

I have read and understand the information listed above. I understand that this is my last chance to withdraw my data and video, which I can do now by stating that request; otherwise by signing this consent form, I give my informed consent to researchers and allow them to use the data that I have provided. I understand that I have not given up my legal rights.

Name (please print): _____ Signature: _____
Date: _____

A2. Video release form and true consent

You have volunteered to take part in our research project: Details. We are examining truthful and deceptive accounts. You were asked to watch a video and then lie or tell the truth about what you saw. Also, you were asked demographic questions (i.e., age, gender, race) about yourself. This study took about 30 minutes.

The goal of this study is to examine truthful and deceptive accounts. We are trying to find out whether the types of details (i.e., truthful vs. deceptive) that people include in their accounts might impact judgments of deception. For example, if someone includes more lies than truths in their deceptive account, are people more likely to accurately judge them as lie-tellers? Currently, the impact of truthful and deceptive details on an account are unknown.

This study has no risks, other than potential discomfort associated with lying or being interviewed. You will be given course credit for your help here today and entered into a draw for $50. You may still withdraw from this study (i.e., tell the experimenter that you do not want to take part); however, you must do so before you leave here today. If you decide not to take part in the study, all of your data (e.g., video, forms) will be deleted/shredded and you will still be given course credit.

Data will be stored on password-protected computers, on external hard drives, and in cabinets in a locked room. Only the primary researchers (Dr. Amy Leach and Elizabeth Elliott), their assistants, and other researchers interested in social psychology will have access to the data (e.g., for meta-analyses). Your individual performance will remain confidential and will not be released to professors, employers, or in publications; only group results will be reported (e.g. conference presentations, journal articles).

During this study, a video camera was used to record your interview with the experimenter. We use recordings of individuals in different studies; the videos are shown to people who are asked to say who is lying or telling the truth. We could also use these videos to look for cues to lying, training, or a number of other purposes. These videos will be used in other studies and may be presented at conferences and in classes or shared with other psychology researchers.

This study has been approved by the UOIT Research Ethics Board (REB #14220) on March 9th, 2017. If you have any questions about this project, please contact the principal investigator, Dr. Amy Leach (amy.leach@uoit.ca). Any questions regarding your rights as a participant, complaints or adverse events may be addressed to the Research Ethics Board through the Ethics and Compliance Officer - reseachethics@uoit.ca or 905.721.8668 x. 3693.

I have read and understand the information listed above. I understand that this is my last chance to withdraw my data and video, which I can do now by stating that request; otherwise by signing this consent form, I give my informed consent to researchers and allow them to use the data and the video that I was produced. I understand that I have not

given up my legal rights.
Name (please print): _____ Signature: _____
Date: _____

A3. Experimental script

Lie-Tellers

Experimenter: Hi, my name is _____ and I am a research assistant. This study will take approximately 30 minutes to complete. Please keep in mind that participation is voluntary, confidential, and you are able to withdraw from the study at any time without penalty or loss of compensation. The purpose of this study will be explained in the following video. Please pay close attention to the instructions. Once you are done watching the video, I will be back to ask you several questions.

A computer program will randomly select whether a participant is placed in the lie or truth condition - participants in the lie condition will view the Suspicious Video.

Video prompts:

The video that you have just watched showed an office that contained suspicious items. For the next part of the experiment, you will be interviewed by the experimenter regarding the video that you saw.

The experimenter does not know what you saw in the video - your job is to convince the experimenter that you saw **a video of an average office** (i.e., **lie** about what you saw). If you are successful at convincing him/her, then you might be able to earn an extra $50.

You have two minutes to prepare your description of the office before the experimenter will return to interview you. **REMEMBER, YOU ARE SUPPOSED TO LIE ABOUT WHAT YOU SAW.**

Truth-Tellers

Experimenter: Hi, my name is _____ and I am a research assistant. This study will take approximately 30 minutes to complete. Please keep in mind that participation is voluntary, confidential, and you are able to withdraw from the study at any time without penalty or loss of compensation. The purpose of this study will be explained in the following video. Please pay close attention to the instructions. Once you are done watching the video, I will be back to ask you several questions.

A computer program will randomly select whether a participant is placed in the lie or truth condition - participants in the lie condition will view the Innocent Video.

Video prompts:

The video that you have just watched showed an office that contained various items. For the next part of the experiment, you will be interviewed by the experimenter regarding the video that you saw.

The experimenter does not know what you saw in the video - your job is to convince the experimenter that you saw **a video of an average office** (i.e., **tell the truth** about what you saw). If you are successful at convincing him/her, then you might be able to earn an extra $50.

You have two minutes to prepare your description of the office before the experimenter will return to interview you. **REMEMBER, YOU ARE SUPPOSED TO TELL THE TRUTH ABOUT WHAT YOU SAW.**

A4. Interview questions

Experimenter: Please tell me in as much detail as possible, what you saw in the office.

Experimenter: I know that you saw what was in the office, so I am going to ask you questions about that.

1. What did you see on the wall?
2. What items were on the desk?

I have interviewed another person, so I have to verify the information:

3. There were papers on the desk. What was written on them?
4. What was marked on the calendar?
5. Where was the gun?
6. What books were on the table?
7. What locations were marked on the map?
8. What is the significance of the locations marked on the map?
9. Can you describe who was in the pictures?
10. There was a package on the desk. Describe what was in the box.
11. Who do you think was working at that desk?
12. What do think that person is planning?
13. Is there anything that we have to be worried about?
14. Are you lying to me?

Experimenter: This is a serious matter and I have to get to the bottom of this.

15. I know there was a bomb in the room. Where was the bomb?
16. How big was the bomb?
17. If he/she wasn't making a bomb, then why were there tools on the table?
18. A date was marked on the calendar, when is the bomb planned to go off?
19. There were places marked on the map, where is he going to place the bomb?
20. I heard that he/she plans to plant a bomb on a bridge. Did you see anything that supports that?
21. Are you lying to me?

A5. Experimental manipulation questionnaire

Please select one of the following:

- ☐ I was **LYING**
- ☐ I was **TELLING THE TRUTH**

Place a check mark next to the items that you remember seeing in the video:

- ☐ Letters
- ☐ Blueprint
- ☐ Candy
- ☐ Blueprint (i.e., HVAC system)
- ☐ Blueprint (i.e., bridge with scribbled writing)
- ☐ Coffee mug with pens, pencils, highlighters
- ☐ Coffee mug with pens, scissors, screwdriver
- ☐ Tape dispenser
- ☐ Books
- ☐ Plant
- ☐ Instructional bomb-making manual and books
- ☐ Empty package (i.e., rectangular cardboard box)
- ☐ Package with imitation pipe bomb
- ☐ Laptop
- ☐ Photo of Obama
- ☐ Photo of Obama with eyes scratched out
- ☐ Map with pictures of tourists at famous bridges (e.g., Golden Gate Bridge)
- ☐ Map with pictures of circled famous bridges (e.g., Golden Gate Bridge)

- ☐ Printer
- ☐ Calendar with circled date
- ☐ Newspaper clippings of events
- ☐ Newspaper clippings of previous bridge malfunctions, demolitions, and explosions

How much of what you told me was a lie? (Scaled response: 1 to 100%)

How much of what you told me was the truth? (Scaled response: 1 to 100%)

Were you motivated to convince the interviewer that you were telling the truth? YES/ NO

Do you think the interviewer believed you? YES / NO

A6. Demographics questionnaire

What is your current age (type age)? ____

What is your gender (select one)?

- ☐ Male
- ☐ Female
- ☐ Other

Which of the following races/ethnicities best describes you (select one)?

- ☐ Aboriginal (i.e., Inuit, Metis, North American Indian)
- ☐ Arab/West Asian (e.g., Armenian, Egyptian, Iranian, Lebanese, Moroccan)
- ☐ Black (e.g., African, Haitian, Jamaican, Somali)
- ☐ Chinese
- ☐ Filipino
- ☐ Japanese
- ☐ Korean
- ☐ Latin American
- ☐ South Asian
- ☐ South East Asian
- ☐ White (i.e., Caucasian)
- ☐ Other

A7. Debriefing

First of all, you did a great job! I don't know whether you were supposed to be lying or not, but I believed you. You did exactly what we wanted you to do. Because of your help here today, I am going to enter you into a draw to win an extra $50.

The goal of this study is to examine truthful and deceptive accounts. We are trying to find out whether the types of details (i.e., truthful vs. deceptive) that people include in their accounts might impact judgments of deception. For example, if someone includes more

lies than truths in their deceptive account, are people more likely to accurately judge them as lie-tellers? Currently, the impact of truthful and deceptive details on an account are unknown.

Deception was use to ensure the ecological validity of the study. In real life situations, witnesses typically do not know in advance that they are about to witness a suspicious event and that they will have to lie about it later. Alerting you to the type of video that you will see might have led you to behave unnaturally in the subsequent interview. The same might have happened if you knew that we will be analyzing the contents of your account and how it will be perceived by observers.

Thank you so much for taking part in our study! I want to let you know that it is *very* important that you do not talk to anyone else about this study who hasn't done it. If people know what we're studying before they arrive, they might change their behavior, and then we wouldn't be able to successfully run the study. That's why we could not tell you exactly what you would be doing at the beginning of the study. So, it is very important that you do not discuss this study with others. Would you promise that you won't tell anyone about the study?

[Wait for promise.]

Thank you again for participating!

Appendix B: Study 2

B1. Consent form

You have been invited to participate in the following research project: Decisions. You will be asked to watch video clips of adults answering questions about a video that they watched, and you will be asked questions about each person's truthfulness. You will also be asked demographic information about yourself (e.g., age, gender, race). The purpose of this study is to learn how people make decisions about deception. The entire session should last approximately 1.5 hours.

There are no known physical, psychological, economic, or social risks associated with this study. You are not obliged to answer any questions that you find objectionable or which make you uncomfortable. You will be given 1.5 credits for your participation in this study whether you complete the study or not. Your participation in this study is completely voluntary and you may withdraw from this study at any time (by informing the experimenter) without any consequences or penalties. If you withdraw, we will delete your data. Please note that you cannot withdraw from the study after you leave here today because there is no way for us to retrieve your specific lie detection decisions.

All information will be coded and stored in a secure area. Only the primary researchers (Dr. Amy Leach and Elizabeth Elliott), their assistants, and other researchers interested in social psychology will have access to the data (e.g., for meta-analyses). Individual performance will remain confidential and will not be released to professors, employers, or in publications. Only group results will be reported (e.g., conference presentations, journal articles). If you have any questions concerning the research study or experience any discomfort related to the study, please contact Dr. Amy Leach at (905) 721-8668 ext. 3706 or amy.leach@uoit.ca.

This study has been approved by the UOIT Research Ethics Board (REB # 14220) on March 9th, 2017. Any questions regarding your rights as a participant, complaints, or adverse events may be addressed to Research Ethics Board through the Research Ethics Coordinator – researchethics@uoit.ca or 905.721.8668 x. 3693.

I have read and understood the statements above. I have had my questions answered to my satisfaction and I understand that I may ask additional questions at any time. My signature, below, indicates my free and informed consent to participate in this research. I understand that I have not waived my legal rights.

Name (please print): _____
Signature:_____ Date: _____

B2. Lie detection task

The following is a script of the instructions that will appear on the screen for participants. The first (general) instructions will appear once, however, the question (i.e.,

'was the participant in the video lying/telling the truth') will appear depending on condition.

Instructions – Please read carefully

You will watch video clips of different adults speaking with an experimenter. In each clip, you will see an adult answering a series of questions. Some adults will be lying, whereas others will be telling the truth. The clips will be randomly presented, so that each adult has a 50-50 likelihood of telling the truth or lying. The segments are also independent. This means that if the person in Clip 1 is telling the truth, there is still a 50-50 change that the person in Clip 2 is telling the truth or lying.

What you are asked to do:

You will be asked to indicate whether you think each adult is lying or telling the truth about what they saw. For each of the following videos, please indicate whether the individual in the video is lying or telling the truth by clicking the mouse and sliding it along the scale.

Reminder:

If you have any questions, please ask the experimenter.

[Participant clicks 'continue']

New page/screen

[Video plays with sliding scale underneath]

[Upon the completion of the video, a 'continue' button appears. Participant must click 'continue']

New page/screen

Was the participant in the video:
□ Lying
□ Telling the truth

[Participant clicks 'continue']

B3. Sliding scale

Below is an imagine (static) of the sliding scale that was used in our lie detection task:

B4. Demographics questionnaire

What is your current age (type age)? ____

What is your gender (select one)?

- Male
- Female
- Other

Which of the following races/ethnicities best describes you (select one)?

- Aboriginal (i.e., Inuit, Metis, North American Indian)
- Arab/West Asian (e.g., Armenian, Egyptian, Iranian, Lebanese, Moroccan)
- Black (e.g., African, Haitian, Jamaican, Somali)
- Chinese
- Filipino
- Japanese
- Korean
- Latin American
- South Asian
- South East Asian
- White (i.e., Caucasian)
- Other

B5. Debriefing

As you read in the consent form, the goal of this study is to learn how people make decisions about deception. More specifically, we are testing when and if a decision changes throughout an account.

We are interested in this issue because when people make decisions about guilt or innocence, they are often asked to make a dichotomous decision (i.e., lying or telling the truth). We believe that this type of decision is dynamic and changes throughout an account. That is, an observer might change their mind about the truthfulness of an account as the individual speak. We suspect that this might be related to the specific content of the truth or lie that is told or exhibited at that particular moment. It is currently unknown whether a lie detection decision changes throughout an account or is static from beginning to end.

Thank you so much for taking part in our study! It is *very* important that you do not talk to anyone else about this study who hasn't done it. If people know what we're studying

before they arrive, they might change their behavior, and then we wouldn't be able to successfully run the study. That's why we could not tell you exactly what you would be doing at the beginning of the study. So, it is very important that you do not discuss this study with others.

Thank you again for participating!

This study has been approved by the UOIT Research Ethics Board (REB # 14220) on March 9th, 2017.

Appendix C: Study 3

C1. Consent form

Title of Research Study: Choices

Name of Principal Investigator (PI): Dr. Amy Leach

PI's contact number/email: (905) 721-8668 ext. 3706 or amy.leach@uoit.ca

Names of Co-Investigator/ Student Lead, and contact email: Elizabeth Elliott, email: elizabeth.elliott@uoit.ca

Departmental and institutional affiliation: Social Science and Humanities

External Funder/Sponsor: none

Introduction

You are invited to participate in a research study entitled Choice. You are being asked to take part in a research study. Please read the information about the study presented in this form. The form includes details on the study's procedures, risks, and benefits that you should know before you decide if you would like to take part. You should take as much time as you need to make your decision. You should ask the Principal Investigator (PI) or study team to explain anything that you do not understand and make sure that all of your questions have been answered before signing this consent form. Before you make your decision, feel free to talk about this study with anyone you wish including your friends and family. Participation in this study is voluntary.

This study has been reviewed by the University of Ontario Institute of Technology (Ontario Tech University) Research Ethics Board [REB # 15794] on March 18th, 2020.

Purpose and Procedure:

The purpose of this study is to learn how people make decisions about deception. You have been invited to participate in this study because you are part of the psychology participant pool.

You will be asked to watch video clips of adults answering questions about a video that they watched, and you will be asked questions about deceit. You will also be asked demographic information about yourself (e.g., age, gender, race). The entire session should last approximately 30 minutes. Overall, 135 participants will take part in this study. The data collected will be used as part of the student lead's dissertation and may also be adapted for conference presentations or academic journals.

Potential Benefits:

The findings may have beneficial implications for the research community. You will not directly benefit from participation in this study; however, participants typically report that these studies are interesting and rewarding.

Potential Risk or Discomforts:

There are no known or anticipated risks to you from participating in this study.

Use and Storage of Data:

All information will be coded and stored in a secure area on password-protected computers and using online cloud storage solutions indefinitely. The only personal information that will be collected will be demographics (e.g., gender, age, race). This information will be used to explain the data and help to align it in the context of previous or future research. Although demographics will be collected, participants will remain anonymous because data will be coded with variables that only the researchers can understand (e.g., 0, 1). Furthermore, each participant will be randomly assigned an ID number by the computer in order for information to never be traced back to specific participants. Only the primary researchers (Dr. Amy Leach and Elizabeth Elliott), their assistants, and other researchers interested in social psychology will have access to the data (e.g., for meta-analyses). Individual performance will remain confidential and will not be released to professors, employers, or in publications. Only group results will be reported (e.g., conference presentations, journal articles).

All information collected during this study, including your demographics, will be kept confidential and only aggregate data may be shared. You will not be named in any reports, publications, or presentations that may come from this study.

Confidentiality:

Your privacy shall be respected. No information about your identity will be shared or published without your permission, unless required by law. Confidentiality will be provided to the fullest extent possible by law, professional practice, and ethical codes of conduct. Please note that confidentiality cannot be guaranteed while data is in transit over the Internet.

This research study includes the collection of demographic data which will be aggregated (not individually presented) in an effort to protect your anonymity. Despite best efforts it is possible that your identity can be determined even when data is aggregated.

Voluntary Participation:

Your participation in this study is voluntary and you may partake in only those aspects of the study in which you feel comfortable. You may also decide not to be in this study, or to be in the study now, and then change your mind before the debriefing. You may leave the study at any time (by closing the browser) without affecting your research credit.

Right to Withdraw:

If you withdraw from the research project at any time before the debriefing, any data that you have contributed will be deleted from the study and you do not need to offer any reason for making this request. Please note that you cannot withdraw from the study after

you see the debriefing page because data is anonymous and there is no way for us to retrieve your specific information.

Conflict of Interest:

Researchers have an interest in completing this study. Their interests should not influence your decision to participate in this study.

Compensation, Reimbursement, Incentives:

You will be given 0.5 research credits for your participation in this study whether you complete the study or not.

Debriefing and Dissemination of Results:

If you are interested in being informed of the results of the study, please contact the PI or student lead after the completion of the study. However, you will not receive feedback about your individual participation because only aggregate results will be published.

Participant Rights and Concerns:

Please read this consent form carefully and feel free to ask the researcher any questions that you might have about the study. If you have any questions about your rights as a participant in this study, complaints, or adverse events, please contact the Research Ethics Office at (905) 721-8668 ext. 3693 or at researchethics@uoit.ca. If you have any questions concerning the research study or experience any discomfort related to the study, please contact Dr. Amy Leach at (905) 721-8668 ext. 3706 or amy.leach@uoit.ca.

By signing this form, you do not give up any of your legal rights against the investigators, sponsor or involved institutions for compensation, nor does this form relieve the investigators, sponsor or involved institutions of their legal and professional responsibilities.

Consent to Participate:

I have read the consent form and understand the study being described.

I freely consent to participate in the research study, understanding that I may discontinue participation at any time. A copy of this Consent Form has been made available to me.

☐ I agree

C2. Lie detection task

The following is a script of the instructions that will appear on the screen for participants. The first (general) instructions will appear once, however, the questions specific to condition (i.e., 'was the participant in the video lying/telling the truth' and 'how confident are you in your decision') will appear after each video.

Instructions – Please read carefully

You will watch video clips of different adults speaking with an experimenter. In each clip, you will see an adult answering a series of questions. Some adults will be lying, whereas others will be telling the truth. The clips will be randomly presented, so that each adult has a 50-50 likelihood of telling the truth or lying. The segments are also independent. This means that if the person in Clip 1 is telling the truth, there is still a 50-50 change that the person in Clip 2 is telling the truth or lying.

(Individual condition)

What you are asked to do:

Please pay close attention to the videos because you will asked questions after each video.

[Participant clicks 'continue']

New screen

[Video plays]

[Upon the completion of the video, a 'continue' button appears. Participant must click 'continue']

New screen

The person was telling:

- a lie
- the truth

[Participant clicks 'continue']

How confident are you in your decision (from 0%-100%)? _____

(Details condition)

Please pay close attention to the videos because you will asked questions after each video.

[Participant clicks 'continue']

New screen

[Video plays]

[Upon the completion of the video, a ‘continue’ button appears. Participant must click ‘continue’]

New screen

The details provided were:

- a lie
- the truth

[Participant clicks ‘continue’]

How confident are you in your decision (from 0%-100%)? _____

(Event condition)

Please pay close attention to the videos because you will asked questions after each video.

[Participant clicks ‘continue’]

New screen

[Video plays]

[Upon the completion of the video, a ‘continue’ button appears. Participant must click ‘continue’]

New screen

The event described was:

- a lie
- the truth

[Participant clicks ‘continue’]

How confident are you in your decision (from 0%-100%)? _____

C3. Demographics

What is your current age (type age)? ____

What is your gender (select one)?

- Male
- Female
- Other

Which of the following races/ethnicities best describes you (select one)?

- Aboriginal (i.e., Inuit, Metis, North American Indian)
- Arab/West Asian (e.g., Armenian, Egyptian, Iranian, Lebanese, Moroccan)
- Black (e.g., African, Haitian, Jamaican, Somali)
- Chinese
- Filipino
- Japanese
- Korean
- Latin American
- South Asian
- South East Asian
- White (i.e., Caucasian)
- Other

C4. Debriefing

As you read in the consent form, the goal of this study is to learn how people make decisions about deception. We are trying to find out whether the types instructions that people receive when making lie detection decisions might impact judgments of deception. For example, if someone is told to focus on the details provided rather than the individual, are they more likely to accurately judge them as lie-tellers? Currently, the impact of instructions on lie detection is unknown.

Thank you so much for taking part in our study! I want to let you know that it is *very* important that you do not talk to anyone else about this study who hasn't done it. If people know what we're studying before they arrive, they might change their behavior, and then we wouldn't be able to successfully run the study. So, it is very important that you do not discuss this study with others. Would you promise that you won't tell anyone about the study?

Thank you again for participating!