Fool me once, fool me twice:

The relationship between statement consistency and veracity across

repeated recalls

Charlotte A. Hudson

Department of Psychology,

University of Portsmouth

September 2020

"We must look for consistency. Where there is a want of it we must suspect deception."

— *Arthur Conan Doyle,*

The Casebook of Sherlock Holmes, Vol. 1

Abstract

The overarching aim of the current thesis is to examine the relationship between veracity and consistency across multiple recalls provided by a single individual in the context of repeated investigative interviews. As a form of communication, deception is a dyadic process which can be examined at two points: *sending* and *receiving*. The content of an interviewee's statement can be examined to determine whether it is consistent over multiple presentations, and a veracity assessor's judgment-making process can be examined to determine the influence of their personal beliefs about the relationship between consistency and deception on their perception of the interviewee's statements.

Chapters II and III examine consistency at the point of sending. In **Experiment 1**, I examined how truth tellers' and liars' consistency differed across two phases of an interview as a result of strategic responses to manipulated interview techniques. In **Experiment 2**, two types of free recall written interviews were compared from the first interview. After a one week delay, truth tellers and liars participated in an oral interview. Across both experiments I found little support that objective measuring of could be used to indicate veracity. Chapters IV and V of this thesis examine consistency from the receiver's perspective. In **Experiment 3**, participants were asked to rate a series of interview transcripts for consistency, based upon objective coding criteria for consistency, and to separately make a holistic consistency judgement. In **Experiment 4**, participants analysed four interview transcripts and reported which factors they found to be indicative of (in)consistencies. Based on this explanation, I created the Layperson's Understanding of Consistency Key (LUCK) coding scheme, which was then compared to a coding scheme built around the experimental definition of consistency. I found that holistic consistency assessments could predict veracity where objective consistency coding could not (Expt. 3 & 4), suggesting that the lay assessment

of consistency relies on features not captured by researcher-generated coding schemes (Expt. 4).

This set of experiments is the first to highlight the importance of differentiating between the examination of consistency as a sending or receiving pattern. Across the programme of research, I found little support for the objective measuring of consistency to infer veracity. However, holistic consistency judgements were found to be relatively accurate at differentiating between truthful and deceptive statements. In the discussion of the results, I consider the applied implications of using consistency to infer veracity, and suggest avenues for future research to decipher the mechanisms that inform the more reliable holistic consistency judgements.

Contents

Part One – Introduction

Part Two – Consistency at the point of sending

Part Three – Consistency at the point of receiving

Part Four – General Discussion

6

Declaration

Whilst registered for the above degree, I have not been registered for any other research

award. The results and conclusions embodied in this thesis are the work of the named

candidate and have not been submitted for any other academic award.

Word count: 42,186

List of Tables & Figures

Figures

Tables

Abbreviations

ANCOVA Analysis of covariance

ANOVA Analysis of variance

CBCA Content-Based Criteria Analysis CI Cognitive

Interview

ICC Intra-class correlation coefficient LUCK Layperson

Understanding of Consistency Key MANOVA Multivariate analysis

of variance OSF Open Science Framework

RM Reality Monitoring

SAI Self-Administered Interview

SFEC Science Faculty Ethics Committee SIP Structured

Interview Protocol VR Virtual reality

WFR Written free recall

Acknowledgements

I will be eternally grateful to my supervisory team for guiding me over the past few years. Aldert, Lucy, Lorraine; you all believed in me even when I didn't believe in myself. Your guidance and encouragement kept me on this rollercoaster, even when it felt like I was about to fall off. You have each shaped me into a better researcher and a better person.

I couldn't have done any of this without the unconditional love and unwavering support of my incredible parents. Mum, Dad; you have always given me your all, and I wouldn't be who I am today without you. You are both my heroes, and I hope you know how important to me you are, because words will never do justice for my love for you.

I am very lucky to have shared the last few years with some awesome people. Liam; forever be my collab-bae so we can talk about science, stats, and our folie à deux. Will; you are me, and I am you, so I think this might make you a Dr now too? Thank you for being the best friend a girl could ask for. Kate and Emma; you've both kept me sane when I thought I might lose my mind. Thank you for the pep talks and tequila. Lauren; you fed me coffee, proofread my work, dealt with my imposter syndrome, and could always make me smile. Thank you for being you.

11

Dissemination

Articles

- Hudson, C. A., Vrij, A., Akehurst, L., & Hope, L. (2019). The devil is in the detail: Deception and consistency over repeated interviews. *Psychology, Crime & Law, 25*(7). 752-770. doi: 10.1080/1068316X.2019.1574790

- Hudson, C. A., Vrij, A., Akehurst, L., & Hope, L. (2020). An examination of the Self Administered Interview as a verbal veracity assessment tool. *Applied Cognitive*

Psychology, 34(5). 1083-1091. doi: 10.1002/acp.3696

- Hudson, C. A., Vrij, A., Akehurst, L., Hope, L., & Satchell, L. P. (2020). Veracity is in the eye of the beholder: A lens model examination of consistency and deception. *Applied Cognitive Psychology, 34*(5). 996-1004. doi: 10.1002/acp.3678

- Hudson, C. A., Vrij, A., Akehurst, L., & Hope, L. (under review). Understanding laypeople's approach to measuring consistency to determine deception. *Applied Cognitive Psychology.*

CH is responsible for the design, material development, data collection, analysis and manuscript preparation of all four experiments. AV, LA, and LH provided guidance and feedback on all four experiment designs and comments on all manuscript drafts. LS provided guidance and assistance on the data analysis and manuscript draft for Experiment 3.

Conference Presentations

- Hudson, C. A., Vrij, A., Akehurst, L., Hope, L., & Satchell (2019, July). *A lens model examination of consistency and deception.* Paper presented at the annual meeting of European Association of Psychology and Law, Santiago de Compostela, Spain.

- Hudson, C. A., Vrij, A., Akehurst, L., & Hope, L. (2019, June). *"He went that way!": Consistent and detailed eyewitness deceptions.* Paper presented at the annual meeting of International Investigative Interviewing Research Group, Stavern, Norway.

▪ Hudson, C. A., Vrij, A., Akehurst, L., & Hope, L. (2018, July). *The SAI and subsequent recall: Examining written and verbal cues to deception.* Paper presented at the annual meeting of Behavioural and Social Sciences in Security, Lancaster, UK.

▪ Hudson, C. A., Vrij, A., Akehurst, L., & Hope, L. (2018, June). *Attention to detail: Deception and consistency in investigative interviews.* Paper presented at the annual meeting of the European Association of Psychology and Law, Turku, Finland.

▪ Hudson, C. A., Vrij, A., Akehurst., L. & Hope, L. (2017, August). *The devil is in the detail: Deception and consistency over multiple interviews*. Poster presented at Decepticon: Second International Conference on Deceptive Behavior, Palo Alto, USA.

▪ Hudson, C. A., Vrij, A., Akehurst., L. & Hope, L. (2016, July). *The effect of veracity and interviewing techniques on statement consistency.* Poster presented at 26th Annual Conference of the European Association of Psychology and Law, Toulouse, France.

Part One – Introduction

Consistency is often influential in credibility assessments in the legal system (Reinhardt & Sporer, 2008) with a majority of investigators, judges, and lawyers believing that inconsistencies are indicative of deception (Fisher, Brewer, & Mitchell, 2009; Vredeveldt, Granhag & van Koppen, 2014). This belief is known as the consistency heuristic (Granhag & Strömwall, 1999). When multiple statements are available, observers rely on the consistency heuristic as an indicator of veracity more so than any other cue (Strömwall, Granhag & Jonsson, 2003). Both laypeople (Krix, Sauerland, Lorei, & Rispens, 2015; Global

Deception Research Team, 2006) and professionals in the legal system (Bogaard & Meier, 2017; Strömwall & Granhag, 2003) report similar beliefs regarding the consistency heuristic. In the interview room, investigators often report relying on (in)consistency to form opinions about an interviewee's veracity and in turn to develop lines of inquiry (Deeb et al., 2018). In the courtroom, lawyers actively try to elicit inconsistencies to discredit witnesses, victims and suspects, and judges direct juries to attend to (in)consistencies when assessing statement credibility (Fisher et al., 2009).

The overarching aim of the current thesis is to examine the relationship between veracity and consistency across multiple experiential recalls provided by a single individual in the context of investigative interviewing. As deception is a form of communication, it is a dyadic process. The receiver (i.e. the veracity assessor) usually makes a judgement about the message (i.e. truth or lie) that is conveyed by the sender (i.e. truth teller or liar). As such, the relationship between consistency and deception can be examined from two points: the point of *sending* and the point of *receiving*. In this thesis I will first examine the content of senders' messages to determine whether they are consistent over multiple presentations. Secondly, I will examine the influence of receivers' personal beliefs about the relationship between

consistency and deception on their judgement-making process. Experiments 1 and 2 of this thesis will examine the relationship between deception and consistency at the point of sending. Experiments 3 and 4 will examine the relationship between deception and consistency at the point of receiving.

To date, research has examined four types of consistency (Vredeveldt et al., 2014), (*i*) within-statement consistency (the consistency of details provided within one statement), (*ii*) between-statement consistency (the consistency between multiple statements provided by the same individual at different times), (*iii*) within-group consistency (the consistency between statements provided by multiple individuals), and (*iv*) statement-evidence consistency (the consistency between an individual's statement and additional evidence). When appraising

different statements from a single suspect, police officers place a similar value on within statement inconsistency and between-statement inconsistency (Deeb et al., 2018). Therefore, the focus of this thesis will be on statements provided by a single individual and will examine both within-statement and between-statement consistency. Within-group consistency and statement-evidence consistency are based on a different approach to consistency which compares consistency between external information and an individual's memory of an event. This thesis examines consistency across multiple recalls provided by a single individual and, as such, I will not discuss within-group or statement-evidence consistency any further.

There are two main opposing explanations for the relationship between consistency and veracity; the consistency heuristic and the repeat versus reconstruct hypothesis (Granhag & Strömwall, 1999). This thesis seeks to reconcile the relationship between consistency and deception at the point of sending and at the point of receiving in relation to these two explanations. But first a definition of consistency. Consistency is explained by the Cambridge Dictionary as the quality of always behaving or happening in a similar way. In this thesis, I examine consistency in the context of investigative interviewing, where

individuals are frequently required to make multiple statements about an event. To objectively establish consistency across multiple recalls provided by a single individual, researchers usually compare details between different parts of one interview or between interviews and classify the details into four categories: repetitions (information that is reported in an initial part/interview and then repeated in a subsequent part/interview), reminiscences (information that is reported in a subsequent part/interview, but not reported in the primary part/interview; sometimes referred to as commissions in deception research), omissions (information that is provided during an initial part/interview, but not reported in a subsequent part/interview), and contradictions (information provided in the subsequent part/interview that directly opposes what was reported in the primary part/interview) (Fisher, Vrij, & Leins, 2013). Consistency is increased with the occurrence of repetitions, whereas the

presence of reminiscences, omissions, and contradictions decreases consistency.

The consistency heuristic

The term 'consistency heuristic' was coined by Granhag and Strömwall in 1999, and describes a commonly held belief about the relationship between deception and consistency. That is, the majority of people (including 82% of police officers, 72% of prosecutors, and 74% of judges) believe that truth tellers' statements will be more consistent than liars' statements (Strömwall & Granhag, 2003). People do not just hold this belief though, they report that they use it to inform their veracity judgements too. In a recent study by Masip et al. (2018) laypeople and police officers were asked to make veracity judgements of several pairs of interviews, and the vast majority (90%) reported basing their judgement on (in)consistencies between the two accounts.

This belief is even encouraged by some; a simple Internet search will return results where investigators and training agencies claim that inconsistency in a suspect's interview is

a reliable cue to deception (e.g. Cosmos Compliance, The Linguistic Statement Analysis Technique). Some published interrogation manuals report the same belief (e.g. The Reid Technique; Buckley, 2000). Even court instructions direct juries to attend to inconsistencies in witnesses statements to inform credibility assessments (Sixth Circuit Criminal Pattern Jury Instructions, 2019), including the suggestion that inconsistencies can make a statement "less believable", and the instruction for jurors to "ask yourself if [the inconsistency] seemed like an innocent mistake, or if it seemed deliberate" (Sixth Circuit Criminal Pattern Jury Instructions, 2019, section 1.07, point F). The instruction is not just about what is said on the stand though; a note regarding the use of inconsistency states "[point F should be used regarding] inconsistent testimony on the stand, or *inconsistent out-of-court statements or conduct, or both*" (emphasis added). Even though the instructions fail to provide any direction about what constitutes inconsistency, they direct jurors to consider the consistency

of statements made by an individual over the course of an entire investigation; where in some situations initial statements may have been taken *years* before the case reaches court.

There is logic to the consistency heuristic. When a statement does not change, and the information is simply repeated, there is little motivation for a credibility assessor to question it. However, if an individual fails to report everything they previously reported, and starts to omit information, this may lead to questions about the quality of their memory, and doubts about the accuracy of the information that *was* remembered. Also, if someone gives a second statement that contradicts what they previously reported, it is reasonable to wonder whether the individual was wrong in their initial statement or in their second statement. Likewise, if an individual includes reminiscence in subsequent statements, an uninformed credibility assessor may wonder why a person's memory has *improved* over time, and doubt the validity of newly included information.

However, the consistency heuristic is actually incongruous with both theoretical and empirical research into memory. Memory research suggests that various forms of inconsistencies, such as reminiscences and omissions, are common features of truthful memory reports (Fisher et al., 2009; Gilbert & Fisher, 2006; Strange, Dysart, & Loftus, 2015). For example, when honest individuals were interviewed twice about witnessing a videotaped mock crime, 98% of truthful participants included reminiscence in their second recall (Gilbert & Fisher, 2006). Further, although memory is known to fade over time, this fade has little detriment on the accuracy of what is recalled (Evans & Fisher, 2011).

Despite the common use of the consistency heuristic to inform veracity judgements, research typically reports that there is no significant difference between truth tellers' and liars' consistency across multiple accounts. In fact, where differences are found, liars tend to be more consistent than truth tellers (Vredeveldt et al., 2014). Accordingly when veracity judgements are made based upon the consistency heuristic, naturally occurring

inconsistencies in a truth teller's account may lead a perceiver to falsely believe they are being deceived.

The repeat vs reconstruct hypothesis

The repeat vs reconstruct hypothesis (Granhag & Strömwall, 1999) is probably the most prominent approach to examining the relationship between consistency and deception. According to this hypothesis, liars do not take their credibility for granted and make a deliberate effort to appear honest (Granhag, Mac Giolla, Strömwall & Rangmar, 2013). As proposed by the consistency heuristic, consistency is an indicator of honesty (Granhag & Strömwall, 1999). To achieve the aim of being consistent within an interview or across multiple interviews, liars often report that they think back to what they said previously and try to *repeat* this information (Granhag et al., 2013). Meanwhile, it is suggested that truth

tellers prioritise consistency to a lesser extent than liars because they are not concerned with appearing truthful. They assume their credibility will be correctly identified. This assumption has been labelled the *illusion of transparency* (the concept that one's inner feelings will externally manifest; Gilovich, Savitsky, & Medvec, 1998). Further, truth tellers hold *just world beliefs* (the belief that people get what they deserve, and deserve what they get; Lerner, 1980). Truth tellers, therefore, often do not apply a particular strategy to appear convincing (Granhag et al., 2013). Instead, the repeat vs reconstruct hypothesis suggests that truth tellers go back to their memory of the event and *reconstruct* the event again. Truth tellers' statements naturally vary over multiple recalls, due to the intrinsic reconstructive nature of memory: Information is forgotten and remembered differently at different times (Tulving, 1984). As a result of these opposing strategies, liars may appear more consistent than truth tellers contrary to what is predicted by the consistency heuristic.

One of the most salient features of memory in relation to repeated recall is it's reconstructive nature (Tulving, 2000). An individual's recall reflects both the contents of the memory as

well as the process of retrieval (Tulving, 1983). Therefore, if the retrieval processes differ across multiple attempts (as is the case in investigative settings where different interview questions are asked at different times, using different phrasing, for different reasons, and by different people), the recall account will likely differ, even when the contents of memory do not change. As such, it is expected that truth tellers who recall an event over several attempts will include new details, forget some details, and change their reporting over time (Odinot, Wolters & van Giez, 2013). Meanwhile, it is proposed that liars adopt an attempted control approach to interviews (Vrij, 1998). The attempted control approach suggests that a liar will plan ahead and control their behaviour so they can appear as normal and credible as possible. As a consequence of this, a liar may have rehearsed their statement (Vrij, 2004). In fact, liars often report preparing a strategy before an interview

(Hartwig et al., 2010). As a successful liar needs a good memory to be able to remember what they reported in previous interviews (Granhag & Strömwall, 2000), adequate planning and rehearsal could be used to facilitate increased consistency.

Research has found support for the repeat vs reconstruct hypothesis across a variety of interviewing techniques, including Assessment Criteria Indicative of Deception (Colwell, Hiscock-Anisman, Memon, Taylor, & Prewett, 2007) and Strategic Use of Evidence (Luke et al., 2013). Leins, Fisher and Vrij (2012) found that 88% of liars reported an attempt to repeat their story when being interviewed twice using different interview modalities. Considering research into truth tellers' and liars' behaviour and self-reported interview strategies, the repeat vs reconstruct hypothesis makes sense. However, it is based on the premise that there are objective differences in the consistency of truth tellers' and liars' accounts – despite research frequently showing no reliable significant differences in the number of repetitions, reminiscences, omissions or contradictions in across the two types of statement (Vredeveldt et al., 2014). Fundamentally, although people often report perceiving differences in the consistency of truth tellers' and liars' statements, this does not mean that objective

differences actually exist.

Perceptions of deception and consistency

People often overestimate the strength of perceived cues to deception, with people expecting far greater differences between truth tellers' and liars' behaviours than actually exist (Hartwig & Bond, 2011). Furthermore, when examining what cues people report using to determine veracity, there is little correlation between what cues people say they use and what cues they actually use (Hartwig & Bond, 2011). Therefore, there seems to be a substantial discrepancy between the perceived strength of a variety of cues to deception (including consistency) and the actual diagnostic value of these cues. Despite the scientific

evidence for the diagnostic value of the consistency heuristic being weak, there is still a strong belief in, and frequent use of, the heuristic (Bogaard & Meijer, 2017; Granhag & Strömwall, 2000, 2001;Vredeveldt et al., 2014). For example, Masip et al. (2018) found that laypeople who made veracity judgements on a series of written statements reported utilising (in)consistency to detect deception in 90% of the cases. In fact, inconsistencies (contradictions) are the most commonly reported verbal cue used to identify deception by both laypeople and practitioners (Bogaard, Meijer, Vrij & Merckelbach, 2016).

People differ in their judgements of whether a given statement is inconsistent. Granhag and Strömwall (2000) asked participants to act as lie detectors and to rate a series of three videos of the same suspect interrogated on three separate occasions. When the subjective cues reported as justification for participants' veracity-judgements were examined, consistency was the most commonly reported cue. Of 125 participants, 78 reported using the consistency heuristic in making their veracity judgement. However, 38 of these participants reported that they considered the three consecutive statements consistent over time (of which 37 made a truth judgement), whereas the other 40 participants reported that the three consecutive statements were inconsistent over time (of which 35 made a lie judgement).

Thus, the same series of consecutive statements were considered consistent by some judges, and inconsistent by others. This inconsistency in the perception of consistency suggests that there are idiosyncratic interpretations of what is or is not consistent, and that a clear understanding of how to define consistency is lacking among laypersons.

It may be possible that the differences in veracity judgements for the same series of consecutive statements are due to individual differences between receivers. Veracity judgements in interviews are typically subjectively made based upon the receiver's judgement of the sender's statement. People often have poor introspection skills for complex behaviours and, when questioned, are more likely to report implicit, a priori causal theories

for their behaviours (Nisbett & Wilson, 1977). It has been suggested that people respond to questions about cues to deception with reasoned, logical explanations based upon common sense expectations of liars' behaviour, whereas their actual veracity judgement-making is guided by intuition and implicit cognitive processes that are outside of the realm of introspection and self-awareness (Hartwig & Bond, 2011). Consequently, people are unlikely to be able to detail the specific cues that influence their veracity judgement making, and they therefore may report using the consistency heuristic due to a lack of insight into their own behaviours.

Consistency and amount of detail provided in accounts

Fundamentally, as consistency is determined by the details that are reported by interviewees across interviews, consistency is inherently influenced by the amount of detail that an individual reports at the outset in the initial interview. The more details that are reported initially, the more opportunities there are for details to be repeated, omitted, or contradicted in a subsequent statement. If fewer details are reported, there are more opportunities to expand the subsequent statement by including reminiscences. Truth tellers are often found to provide more details than liars (DePaulo et al, 2003). In

addition to this, truth tellers have been found to provide more core details (Leal, Vrij, Deeb, & Jupe, 2018), and more verifiable details (Nahari & Vrij, 2014), as well as to provide greater reconstructability and realism in their stories than liars (Mac Giolla, Ask, Granhag, & Karlsson, 2019). It is thought that these differences in performance arise from deviations in the type of demand placed on memory during interviews with liars and truth tellers (Sporer & Schwandt, 2006). Furthermore, different strategies are commonly reported by truth tellers and liars (Vrij & Granhag, 2012). When truth tellers provide honest reports, they rely on existing memory structures, which allows for longer, more elaborate responses. A liar's task

is much more cognitively challenging, as they are required to construct, verbalise, and keep track of their deceptive account (Dando, Bull, Ormerod, & Sandham, 2013), whilst simultaneously supressing the truth, and assessing whether their story is being believed. Liars may lack the creativity or cognitive flexibility to fabricate details to an equivalent level to that provided by truth tellers, and therefore strategically choose to report shorter, less detailed accounts.

It stands to reason that there are strategic benefits for liars in providing shorter deceptive reports. Providing less information means that as an interview progresses there is *(i)* leeway for adaptation and verbal manoeuvre (Dando et al., 2013), *(ii)* fewer opportunities for liars' statements to contradict evidence, *(iii)* fewer leads for investigators to pursue which could reveal deception (Vrij et al., 2015), and *(iv)* easier maintenance of consistency, as there are fewer details to remember to repeat in follow up interviews.

Consistency across truthful and deceptive accounts

Only rarely are reliable differences found between the objective consistency of truth tellers' and liars' accounts. For instance, Granhag, Mac Giolla, Sooniste, Strömwall and Liu Jonsson (2016) examined statements of true and false intent over repeated interviews. Participants were asked to plan a series of tasks and were led to believe they would complete these tasks.

Truth tellers were instructed that they would be required to complete benign tasks including buying a gift for a friend and taking a photo in a photo booth. Liars were instructed that they would be required to complete a number of suspicious tasks including placing a memory stick containing "illegal" material on a particular shelf in a particular shop, and so should plan a cover story masking their criminal intentions. All participants were intercepted before having the chance to execute their plans and they were interviewed three times. In the interviews, truth tellers were honest about their intentions, whereas liars were required to use

their cover story to mask their true intentions. Statements were coded for repetitions, reminiscences and omissions, and results showed "striking" *similarities* between truth tellers' and liars' accounts (Granhag et al., 2016, pg. 1).

When differences in consistency are found, there is no reliable pattern across results. For example, Strömwall, Granhag and Jonsson (2003) sent pairs of truth telling friends for lunch, and asked pairs of lying friends to pretend that they had had lunch together. Each member of the truth telling and lying pairs was interviewed on two occasions; immediately after the supposed lunch had taken place, and again one week later. Strömwall et al. (2003) found that there was no significant difference between truth tellers and liars in the number of repetitions or omissions across accounts, however truth tellers reported more reminiscent detail in their second interview than did liars. Conversely, Granhag and Strömwall (2002) found that liars included significantly more reminiscences than truth tellers when interviewed three times over 11 days.

More recently, Deeb et al. (2017) examined differences in liars' and truth-tellers' statement consistency about two events. Participants were required to view a meeting which featured a non-critical event and a critical event. Across two interviews, truth tellers were honest about both aspects of the meeting. Liars were honest about the non-critical event, however, they were required to lie about the critical event. Overall, no differences were found between the number of omissions, reminiscences or contradictions provided by truth tellers

and liars, however liars' accounts featured fewer repetitions than truth tellers' accounts for both events. It was suggested that requiring liars to lie about the critical event, but remain truthful about the non-critical event, was more cognitively demanding than being honest about both events, and therefore this made it harder for liars to maintain consistency across their statements.

Consistency has also been examined in relation to a variety of interview techniques. Leins, Fisher, and Vrij (2012) asked liars and truth tellers either the same questions or different questions across two interviews. Half of the participants were asked to use the same mode of reporting in both interviews, whereas the other half of participants were required to switch from verbally reporting to drawing a sketch (or vice versa). Truth tellers responses were considered highly consistent regardless of whether they used the same modality of reporting across both interviews or not. Liars were more consistent if the modes of responding were the same than if they differed. When liars reported across different modalities, they included fewer repetitions and more contradictions. There was no significant difference in the number of reminiscences or omissions for truth tellers or liars. It was concluded that liars lacked the cognitive flexibility to adapt their reporting across different reporting modalities.

Whilst there was no change in modality, Shaw et al. (2014) also tested cognitive flexibility by asking participants to report about a mock security meeting in two phases of an interview. In the first phase of the interview, participants were required to recall the event in chronological order, whereas in the second phase of the interview, they were asked to recall the event in reverse order to either the same or a different interviewer. It was found that liars provided significantly fewer repetitions than truth tellers, but only when they spoke to a different interviewer in the second phase of the interview. Meanwhile, when the reverse order instruction was used in interpreted interviews, liars repeated the same number of details and

included significantly fewer reminiscences than truth tellers (Ewens, Vrij, Mann, & Leal, 2015).

Ironically, there seems to be no consistent pattern of results regarding truth tellers' and liars' consistency. One explanation for this is that liars do not spontaneously display diagnostic cues to deception, and as such specific interview techniques are required to elicit

or enhance such cues (Vrij, 2014). Across research studies into consistency and deception, a number of different interview techniques have been tested and these interview techniques may have elicited consistency in different ways potentially leading to this inconsistent pattern of results. Yet contrary to the widely held beliefs of legal professionals and laypeople, the corpus of research examining the consistency of truth tellers' and liars' statements collectively suggest that consistency is not necessarily indicative of truth telling, and inconsistency is not necessarily indicative of deception.

Outline of the thesis

Previous research has not made a distinction between the points of sending and receiving truthful and deceptive accounts. It is therefore built upon the assumption that receivers perceive consistency patterns in the same way that they are sent. This programme of research seeks to distinguish between the point of sending and the point of receiving to better understand the relationship between consistency and deception. As the consistency heuristic is reported to be commonly relied upon to inform veracity judgements, I aimed to establish whether this belief actually reflects the consistency of truth tellers' and liars' statements at the point of sending, by examining the number of repetitions, reminiscences, omissions and contradictions that are provided across repeated interviews. Furthermore, I examined whether the perception of consistency is influenced by the presence of repetitions, reminiscences, omissions and contradictions, with a view to determining the applied value of using objectively measured consistency to inform veracity judgements.

Ethics

All research proposals were evaluated and granted favourable opinion by the University of

Portsmouth Science Faculty Ethics Committee (SFEC; Appendix 3). For each

experiment, a detailed rationale and explanation of the methodology and method of analysis

was provided for peer review. Informed Consent forms were provided to all participants to

read and sign before the experiments commenced. These outlined the procedures, explained

the nature of the research, clearly stated the right to withdraw at any time, how the data

would be stored and used, and how participants could obtain the findings of the studies.

Furthermore, the SFEC favourable opinion reference number and SFEC contact details were

provided. All data will be kept for a period of at least 10 years, in line with APA guidelines

and the University's Data Management guidelines. Participants were all assigned participant

numbers, so their name is not associated with the data, and anonymised copies of the research

data sets are publicly available on the Open Science Framework.

Transparency

Experiments 2, 3, and 4 were pre-registered on the Open Science Framework. Data for

Experiment 1 was collected before the author was aware of the Open Science Framework,

and therefore this study was not pre-registered. To increase the integrity, openness, and

reproducibility of this thesis, all studies have open materials and open data, available from

the Open Science Framework repository. Direct links are provided for each study.

The Experiments

The following provides a brief overview of the main aims, methodologies, and results

for each of the experiments. Each experiment addresses the relationship between deception

and consistency, either from the sender's or from the perceiver's perspective.

Experiment 1

In real-life situations, suspects are frequently interviewed numerous times, with different interviewers using different techniques and they are asked different questions about the same event (Miller & Stiff, 1993). The aim of this study was to establish whether using different interviewers and different interview techniques would elicit differences in consistency between truth tellers and liars. Consistency is of particular importance in interview settings, as it is commonly utilised by interviewers to infer honesty and deception (Deeb et al., 2018). Experiment 1 focuses on examining the amount of detail and consistency of detail provided by truth tellers ($n = 80$) and liars ($n = 80$) over two phases of an interview. At interview, truth tellers discussed a documentary that they had watched. Liars were tasked with stealing a USB stick, containing confidential information, from a laboratory. The liars did not watch the documentary. At interview, the liars were asked to discuss the documentary, as if they had watched it, and to conceal the fact that they had stolen the USB stick.

In the initial phase of the interview, the interviewer asked participants to provide their statement in chronological order. After a short break, the second phase of the interview began. Either the same person who had conducted the first phase of the interview (same interviewer condition, $n = 79$), or another interviewer (different interviewer condition, $n = 81$) entered the room. Participants were asked to provide their statement again. Half of the interviews were conducted using the same questions as in the first phase ($n = 81$), whereas the remainder of the interviews were conducted using the reverse order recall technique ($n = 79$).

As the reverse order recall instruction disrupts the recollection of events from a schema (Gilbert & Fisher, 2006), I hypothesised that reverse order recall prompts would elicit more reminiscent details than chronological order interviews for truth tellers, but would not

affect the amount of reminiscent detail reported by liars due to their concerns about consistency (Hartwig et al., 2007). I also predicted that truth tellers would include more repetitions and fewer omissions when interviewed by a different interviewer in phase two of their interview than when interviewed by the same interviewer across both phases, as it was anticipated that they would not want to bore the same interviewer by telling the same report again (as opposed to a different interviewer who had not yet heard their report). In contrast, I hypothesised that liars would include more repetitions and fewer omissions when interviewed by the same interviewer in phase two of their interview than when interviewed by a different interviewer as they were aware that the same interviewer had heard their previous statement and could more easily cross-compare for consistency.

It was found that truth tellers provided more details overall, and more reminiscent details than liars but there were no differences between veracity groups for the number of omissions made or repetitions reported. There was also no evidence that the reverse order recall or change interviewer manipulations induced inconsistencies in liars. In sum, I found little support that consistency (or a lack thereof) could be used as a diagnostic cue to deception, and that due to the natural occurrence of reminiscences in truth tellers' statements, liars appeared slightly more consistent than truth tellers.

Experiment 2

In Experiment 1, I found truth tellers provided more details than liars. Previous research has suggested that liars lack the imagination to fabricate details to an equivalent level to that provided by truth tellers (Vrij, Hope, & Fisher, 2014), and tend to report as little information as possible to reduce the opportunity for contradicting themselves in subsequent interviews (Vredeveldt, van Koppen & Granhag, 2014). In Experiment 2, I examine how maximising the amount of detail reported by liars ($n = 64$) and truth tellers ($n = 64$) in their

first recall might impact their consistency after a delay of seven days. The primary aim of this

study was to investigate whether differences can be elicited between truth tellers' and liars' statement consistency by encouraging liars to commit to more detail in their initial interviews. Both truth tellers and liars were asked to watch the same first-person perspective video of a crime but were either instructed to honestly report what they saw, or deliberately mislead authorities about the perpetrator of the crime.

Truth tellers and liars were asked to either complete a written free recall (WFR; $n = 64$) or the Self-Administered Interview (SAI; $n = 64$). The SAI is a questionnaire booklet that draws on theory and empirical research into memory to elicit a comprehensive, detailed free recall, using the witness' own words. One week later, participants returned for a verbal interview. As research indicates that *(i)* the SAI elicits more details than a written free recall (Gabbert, Hope, & Fisher, 2009), *(ii)* truth tellers report more details than liars (Amado et al., 2016; DePaulo et al., 2003), and *(iii)* there is little difference in the consistency of truth tellers' and liars' accounts, (Granhag & Strömwall, 2002; Granhag, Strömwall & Jonsson, 2003; Granhag, Mac Giolla, Strömwall, Sooniste & Liu-Jonsson, 2016), I was interested in how veracity and written interview format influence the level of detail and consistency of witnesses' reports.

It was predicted that the truth tellers would include significantly more reminiscences and repetitions in their second interview than liars if they initially completed the SAI, and that all participants completing the SAI would include significantly more details than those completing a WFR. Furthermore, I predicted that statements provided by truth tellers completing the SAI would have significantly higher accuracy rates than statements provided by truth tellers completing the WFR.

In contrast to Experiment 1, I found that truth tellers omitted and repeated more information in their subsequent interview than liars, however there was no significant

difference in the number of reminiscent details reported. Whilst I found that participants who completed the SAI reported more details in their initial account than participants who

completed the WFR, the SAI (cf. WFR) affected truth tellers' and liars' consistency equally. I concluded that whilst the SAI is an effective information-gathering tool, it did not facilitate the detection of deception to any greater level than the WFR.

Experiment 3

Considering that Experiments 1 and 2 suggest that there are no differences between truth tellers' and liars' consistency, Experiment 3 examines whether people truly base their veracity decisions on perceived statement consistency. Whilst there is a large body of research showing that people generally believe that inconsistency is indicative of deception, identifying consistency in a series of consecutive statements is found to be subjective. The aim of this study was to examine whether people utilise the consistency heuristic when making veracity judgements. Our primary research question was; do people's beliefs about consistency as a cue to deception influence their veracity judgements? Participants ($N = 285$) were asked to report their beliefs about a variety of cues to deception, before making veracity judgements about a series of paired statements. They were also asked to rate the paired statements for perceived amount of detail, holistic consistency between the statements, number of repetitions, number of omitted details, number of reminiscent details, and number of contradictions.

Taking into consideration the widely reported use of consistency to inform veracity judgements (Granhag & Strömwall, 2007; Masip et al., 2018), it was predicted that participants would use the consistency heuristic to inform their veracity judgements, such that statements perceived to be higher in consistency would be rated significantly more truthful than statements perceived to be lower in consistency. Truth tellers provided more details than

31

liars in Experiments 1 and 2, so I predicted that statements that were perceived as more detailed would be rated as more truthful than statements that were perceived as lower in detail. Combining these two considerations, it was further predicted that statements perceived

to be both higher in consistency and higher in detail would be considered significantly more truthful than statements that were perceived to be both lower in consistency and lower in detail. Furthermore, it was predicted that participants would prioritise the consistency heuristic over amount of detail in their veracity decision-making, as it is the most commonly reported perceived cue to deception of the two. I thus predicted that statements that were perceived to be higher in consistency and lower in detail would be considered significantly more truthful than statements that were perceived to be lower in consistency and higher in detail. Finally, research suggests that not all participants interpret cues equally (Granhag & Strömwall, 2000). Therefore, it was predicted that significantly more variability in veracity ratings would be accounted for by individual factors of the participant rating the statement than by level of detail and level of consistency within statements.

I found that when participants made a higher rating for holistic consistency, I could predict that they would make an overall truthful judgement for the statement, and also that the statement that they were rating was actually a truthful statement. However, participants' judgements for amount of detail, repetitions, omissions, reminiscences or contradictions could not be used to predict the veracity condition of the statement they were judging. A lens model showed that whilst perceptions of repetitions, omissions, contradictions and amount of detail influenced veracity judgements, these perceptions (and overall veracity judgements) did not represent the veracity conditions, or expert veracity coding of the statements. Therefore, there was a discrepancy between how laypeople identified consistency and how experts coded for consistency. Furthermore, no strong correlations were found between self reported reliance on cues and actual usage of cues, suggesting that participants were not

basing their decisions on what they reported as being important cues to deception. It was concluded that individuals do not understand the cues they personally use to make veracity decisions, and that a holistic perception of consistency could facilitate accurate veracity judgements.

Experiment 4

In Experiment 3, I found that experts, with dedicated time and resources, who code statements for consistency do not refer to the same aspects of the statement as laypersons when they identify 'consistency'. Therefore, it could be that the academic conceptualisation of consistency does not align with laypersons' understanding of consistency. Experiment 4, Part 1 aimed to examine what laypeople mean when they report using consistency to inform their veracity judgements. Participants identified the key aspects of a statement that they use to characterise (in)consistency. Our primary aim was to determine which key aspects of a statement characterise (in)consistency for laypeople.

Participants ($N = 20$) were provided with a set of four different statement pairs about a memorable event and were asked to rate the holistic consistency for each pair of statements (a total of 80 judgements). Once the participants had made their judgements, they were then asked to highlight the key aspects of the statement pair that they used to inform their consistency judgement and to add notes to explain how that aspect of the statement informed their perception of consistency. I found that participants rated truthful statements as significantly more holistically consistent than deceptive statements, but that no significant differences between truthful and deceptive statements emerged for the number of repetitions, omissions, reminiscences, or contradictions. I conducted a thematic analysis of the statements and comments to reveal the themes that defined participants' understanding of consistency and I identified four major themes that contribute to the layperson understanding of

consistency: Core Information, General Information, Logic, and Narrative. From this, a new coding scheme was developed (Layperson Understanding of Consistency Key; LUCK) to test whether laypeople's approach to measuring consistency can be operationalised and used to facilitate discrimination between truthful and deceptive statements to a greater level than current coding schemes (Phase 2).

Before Phase 1, independent participants were asked to provide two reports about a

memorable event e.g. birthday party, wedding, holiday. Participants were randomly assigned to either tell the truth or lie about the memorable event. In Phase 2, six coders, blind to the veracity of the statements, were trained in the LUCK coding scheme and were asked to rate 20 truthful and deceptive statements (a total of 120 observations) using the LUCK criteria. As well as using the LUCK coding scheme, the coders also coded these 20 statements for number or repetitions, reminiscences, omissions, and contradictions (the four features of consistency). I compared the LUCK coding scheme with ratings for the four features of consistency to determine whether it is possible to accurately determine statement veracity by coding for consistency. Experiment 4 was an exploratory study, and as such I did not formulate any hypotheses.

No significant difference in the consistency scores emerged using the LUCK coding scheme nor when counting the four features of consistency. While, in Phase 1, I was able to establish what aspects of a statement inform holistic consistency judgements made by laypeople, the coding scheme derived from this information did not facilitate discrimination between truthful and deceptive statements in the same way that a holistic consistency judgement does. In sum, the thematic analysis suggests that laypeople use context and meaning to inform their consistency judgements, rather than sterile categorisations of details as (in)consistent. The reduction of the complex construct of consistency down to atomistic units (as is done with coding schemes) resulted in a failure to capture meaningful aspects of

complex social behaviours. It was concluded that it may be that holistic consistency judgements are too complex to be reduced to a relatively simplistic coding scheme based upon current understanding of the use of consistency to assess veracity.

General Discussion

In the general discussion, I provide an overview of the key findings from the four experiments, and consider how they contribute to our understanding of the relationship

between consistency and deception. I discuss the theoretical implications of differentiating between the point of sending and receiving when examining consistency, and the practical impact of the interviewer, the effect of interview modality, and the applied value of consistency to inform veracity judgements. I address some of the limitations of the current experiments and wider field, as well as presenting suggestions for future research in light of this programme of study. Finally, I conclude this chapter with recommendations relating to the interpretation and use of consistency to inform veracity decisions in applied settings.

Part Two – Consistency at the point of sending

Chapter II

The devil is in the detail: Deception and consistency over repeated interviews.

Abstract

Research indicates that truthful statements typically contain more details than fabricated

statements, and that truth tellers are no more consistent than liars over multiple interviews. In this experiment, we examine the impact of *(i)* multiple interviewers and *(ii)* reverse order recall on liars' and truth tellers' consistency and amount of reported detail over repeated recall attempts. Participants either took part in a mock crime (lying condition) or an innocent event (truth telling condition) which they were subsequently interviewed about in two separate interview phases. Truth tellers provided more details overall, and more reminiscent details than liars. There were no differences between veracity groups for the number of omissions made or repetitions reported. Despite the popular belief that inconsistency is a cue to deception, we found little support for the notion that consistency (or lack of consistency) offers a diagnostic cue to deception. We found little evidence that switching interviewer or recalling in reverse order induced inconsistencies in liars. In fact, due to the number of reminiscent details in truth tellers' accounts, our findings suggest that accounts provided by liars tend to be slightly more consistent than those provided by truth tellers. Materials for this paper can be found at osf.io/hgvmk/.

Introduction

Research has consistently shown that people are relatively poor at recognising when someone is lying (Hartwig et al., 2011; ten Brinke, Vohs, & Carney, 2016). Meta-analyses reveal that humans' ability to detect deception rarely exceeds chance (Bond & DePaulo, 2006), and that there is little difference in accuracy between laypeople and practitioners who work in fields where attempts to deceive are commonplace (Aamodt & Custer, 2006; Bond & DePaulo, 2006). Even when investigators are trained to look for cues to deception, only a small to medium training effect for accuracy in distinguishing between truth tellers and liars

is found with the largest effect sizes found for training that focuses on verbal content (Hauch, Sporer, Michael, & Meissner, 2016).

Given people's poor ability to differentiate between truth tellers and liars, researchers have been working towards developing interviewing techniques designed to increase the differences in verbal cues between truth tellers and liars (e.g. Strategic Use of Evidence, SUE; Granhag & Hartwig, 2015) and cognitive lie detection (Vrij, Fisher, & Blank, 2017; Vrij, Leal, Mann, Vernham, & Brankeart, 2015). In the current experiment, we examined the effect of two interviewing strategies (requests for reverse order reporting and changing the interviewer midway through an interview) on truth tellers' and liars' reports over two phases of an interview (i.e. two questioning periods with a short break in between), with respect to the amount of detail provided, and the consistency between the two recall attempts.

Veracity: detail and (in)consistency

Truth tellers tend to provide more details in their accounts than liars (Amado, Arce, & Fariña, 2015; Vrij, 2008; Vrij, Mann, Jundi, Hillman, & Hope, 2014). In comparison to liars, truth tellers provide more verifiable details (Nahari & Vrij, 2014), core details (Leal, Vrij, Deeb, & Jupe, in press) and complications (Vrij, Leal, Jupe, & Harvey, 2018), and present greater reconstructability and realism in their stories (Mac Giolla, Ask, Granhag, & Karlsson,

2019). These differences are thought to stem from discrepancies in the nature of demand on memory for truth tellers and liars (Sporer & Schwandt, 2006) and from different strategies used by truth tellers and liars (Vrij & Granhag, 2012). Individuals providing an honest report can draw from existing memory structures, allowing them to provide longer, more elaborate responses. Meanwhile liars face the more cognitively challenging task of constructing, verbalising, and keeping track of their deceptive accounts (Dando, Bull, Ormerod, & Sandham, 2013). There are also strategic benefits associated with shorter deceptive reports. For instance, providing less information means *(i)* there are fewer opportunities for liars to contradict themselves as evidence emerges through an interview (Hartwig, Granhag,

Strömwall, & Kronkvist, 2006), *(ii)* there are fewer leads for investigators to pursue which could reveal deception (Vrij et al., 2015), *(iii)* there is leeway for verbal manoeuvring and adaptation as the interview progresses (Dando et al., 2013), and *(iv)* it is easier to maintain consistency, as there are fewer details to remember to repeat in follow up interviews.

Consistency has been conceptualised in a number of ways relevant to the criminal justice system (for an overview, see Vredeveldt, van Koppen, & Granhag, 2014). For instance, within-group consistency refers to the comparison of statements provided by multiple suspects (Granhag, Strömwall, & Jonsson, 2003), whereas statement-evidence consistency is the evaluation of a suspect's statement in relation to any gathered evidence (Granhag & Hartwig, 2015). Individual statements can be examined to establish the consistency within them (within-statement consistency; Leins, Fisher, Vrij, Leal, & Mann, 2011). However, in investigations, suspects are often interrogated multiple times (Granhag & Strömwall, 1999; Kassin et al., 2007; Miller & Stiff, 1993) which allows for the comparison of multiple statements provided by a single suspect on a number of different occasions (between-statement consistency). When examining different statements from a single suspect,

police have reported that they place a similar value on within-statement inconsistency and between-statement inconsistency (Deeb et al., 2018).

The 'repeat versus reconstruct' hypothesis (Granhag & Strömwall, 1999) is probably the most prominent approach to examining the relationship of statement consistency and veracity in the lie detection literature. This hypothesis contends that, for liars, being consistent is an important aim in interviews. Over multiple interviews, reported details can be compared and classified into four categories in order to quantify consistency; repetition, omissions, reminiscences (sometimes referred to as commissions in deception research) and contradictions[1](Fisher, Vrij, & Leins, 2013).

In repeated interviews, it is hypothesised that liars think back to their earlier

statements and try to repeat the same information to achieve their aim of being consistent. In contrast, truth tellers are thought to be less concerned with consistency between statements (Hartwig, Granhag, & Strömwall, 2007) and more focused on remembering and reporting the encoding event. They are therefore less likely to focus on what was said in a previous interview, but more likely to think back to the originally encoded event and try to recall this event again, hereby generating reminiscent details and potentially omitting previously reported details (Granhag & Strömwall, 1999; Granhag et al., 2003).

The occurrence of repetitions (information that is provided across all interviews or phases of an interview) increases apparent consistency, while the occurrence of omissions (information that is provided during a primary interview, but not mentioned in a subsequent interview) and reminiscences (information that is provided in a subsequent interview, but not provided in the primary interview) decrease consistency. Perceptions of consistency can mediate credibility judgments (Reinhard & Sporer, 2008) and observers often utilise

[1] Contradictions rarely occur in experimental research (e.g. Granhag et al., 2003) and will not be discussed further.

statement consistency as a cue to deception (de Keijser, Malsch, Kranendonk, & de Gruijter, 2012; Roos af Hjelmsäter, Öhman, Granhag, & Vrij, 2014). When individuals are tasked with assessing the veracity of consecutive statements given by one suspect, the most commonly reported cue sought is consistency (Granhag & Strömwall, 2000, 2001), with 82% of police officers believing that truthful consecutive statements are more consistent than deceptive ones (Strömwall & Granhag, 2003). Lay people express similar views about the relationship between consistency and deception (Krix, Sauerland, Lorei, & Rispens, 2015; The Global Deception Research Team, 2006). This tendency has been referred to as the 'consistency heuristic' (Granhag & Strömwall, 2000). However, research demonstrating the malleability and reconstructive nature of memory (Bartlett, 1920; Hemmer & Steyvers, 2009; Michaelian, 2013) is incongruent with the consistency heuristic. With respect to statement consistency,

memory research suggests that some types of inconsistency are a common feature of memory reports (Fisher, Brewer, & Mitchell, 2009; Strange, Dysart, & Loftus, 2015), with both reminiscences and omissions occurring without any detrimental effect on overall accuracy (Gilbert & Fisher, 2006; Odinot, Memon, La Rooy, & Millen, 2013). Consequently, in some situations, the consistency heuristic could be misleading, and cause a receiver to believe that such inconsistencies in an account might indicate that they are being deceived.

Protocols designed to actively elicit and magnify statement consistency as a useful cue to deception are still relatively understudied in lie detection research (Granhag, Mac Giolla, Sooniste, Strömwall, & Liu-Jonsson, 2016; Vredeveldt et al., 2014). However, research examining unanticipated questions (Granhag & Mac Giolla, 2014) and changing interview modality, from verbal descriptions to drawing a sketch of a restaurant layout (Leins et al., 2011) or of a particular room (Leins, Fisher, & Vrij, 2012) indicates that within statement consistency of suspects can be magnified by introducing specific proactive interview protocols, resulting in diagnostic cues to deception.

Practitioners often encourage applied researchers to examine the impact of using multiple techniques together, particularly when the techniques have already been tested individually. To address this, in the current experiment we introduced two different changes in interview format and examined their effects on liars' and truth tellers' accounts: introducing a request for a reverse order recall, similar to Vrij, Leal, Mann, and Fisher (2012) and switching the interviewer for a different interviewer, similar to Shaw et al. (2014). Combining these two techniques, which have both independently shown promise with respect to increasing the amount of detail provided by truth tellers, could magnify the difference in truth tellers' and liars' performance to a level diagnostic of deception.

The reverse order recall strategy was initially developed to support memory retrieval as part of the Cognitive Interview (CI). Use of this technique may trigger recollection of additional details following a traditional chronological recall of an event (Fisher &

Geiselman, 1992). The reverse order recall instruction has been found to provide unsatisfactory results in some eyewitness interviews (Davis, McMahon, & Greenwood, 2005; Memon, Wark, Bull, & Koehnken, 1997), resulting in suggestions by some that it should be removed from the CI (Davis et al., 2005).

Vrij et al. (2008) interviewed 80 mock suspects about stealing £10 from a wallet. Half of the sample were innocent and were instructed to tell the truth about their actions, whereas the other half of the sample were guilty and were instructed to lie about their actions. Participants were asked to give their accounts in either chronological or reverse order. Liars who were asked to recall events in reverse order displayed more signs of cognitive load (e.g. spoke at a slower rate and provided fewer auditory details) than truth tellers, however the same effects were not found for those who recalled in chronological order. This suggests that when the reverse order recall technique is applied for liars, it results in additional cognitive load, as they try to work backwards through what they have previously reported (Vrij et al.,

2008). The change in recall order makes it more difficult for liars (cf. truthtellers) to monitor their consistency, as this unexpected method of questioning requires cognitive flexibility – a challenge for liars, who typically rehearse their deceptive accounts in a unidimensional manner (Leins et al., 2011). Consequently, increase in cognitive load as a result of reverse order recall should have a larger impact on liars than on truth tellers and thus should magnify the differences between them to a level that is diagnostic of deception.

In legal situations individuals may be interviewed a number of times by a range of individuals for example, police officers, social workers and lawyers. Research has suggested that using the same or different interviewers across repeated interviews has no effect on the accuracy of reported details (Odinot et al., 2013). However, previous research examining the effect of switching interviewers during an interview on cues to deceit found that a different interviewer (vs the same interviewer) in a second phase of an interview resulted in truth tellers providing more details, and liars providing fewer details (Shaw et al., 2014). The

participants in Shaw et al.'s (2014) research were interviewed by two interviewers about a mock security meeting that they had attended. In half of the interviews, new interviewers took over the interview half way through (the 'changed interviewer' condition). Otherwise, the same two interviewers conducted the entire interview ('same interviewer' condition). During the first half of the interviews, participants were required to provide their account in chronological order. In the second half of the interviews, participants were asked to provide a reverse order recall. It was found that differences in performance between truth tellers and liars were most prominent in the 'changed interviewer' condition, where truth tellers provided more detail and more repetitions in the second half of the interview than liars. This was not the case for the 'same interviewer' condition.

A new interviewer (vs the original interviewer) may motivate truth tellers to repeat more and omit fewer details because this new interviewer has not heard their story yet. In

contrast, a switch in interviewers may make liars less motivated to be consistent (and therefore less motivated to repeat information) because it is more difficult for a new interviewer (cf. the original interviewer) to confirm consistency between two interviews.

In the current experiment, we asked liars and truth tellers to provide two accounts, across two phases of an interview, with a short break in between. All participants were asked to give a free recall of their actions in the first interview phase. After the break, either the same or a different interviewer asked the participants to report their account again, either as a free recall or using the reverse order recall instruction. As the two accounts were taken within the same interviewing session, we examined the within-statement consistency between the two accounts.

Most lie-detection studies involve assignment of participants to lie-telling conditions (Vrij, 2008). As these participants may lack the motivation and arousal of individuals who are lying of their own volition (Fisher & Perez, 2007), we instead attempted to increase the ecological validity of our experimental paradigm by allowing participants to select whether they would

tell the truth or lie (Wachi et al., 2017). Providing participants with the option to lie or tell the truth has been implemented in a variety of investigations into deceptive behaviour (Gneezy, 2005; Talwar & Lee, 2008; Tyler, Feldman, & Reichert, 2006), however, we recognise problematic features sometimes associated with this methodology. Whilst research has suggested that males and females lie with equal frequency (DePaulo, Kashy, Kirkendol, Wyer, & Epstein, 1996), there is also evidence to suggest a gender difference in predisposition to lie, with research suggesting men are more likely to lie than women (Dreber & Johannesson, 2008), as well as the inverse (Tyler et al., 2006). Furthermore, there is data to suggest younger people lie more than older individuals (DePaulo et al., 1996), with college students telling more lies per day than community members. For this reason, we carefully checked for the effect of gender and age across our truth telling and lie telling groups.

Based on previous literature (Amado et al., 2015; Gilbert & Fisher, 2006; Granhag et al., 2003; Vrij, 2008; Vrij et al., 2014), we hypothesised that truth tellers would provide more details than liars in the first phase of their interviews (Hypothesis 1a) and more reminiscences than liars in the second phase of their interviews (Hypothesis 1b). As liars are more concerned about appearing consistent than truth tellers (Hartwig et al., 2007), it was further predicted that liars would make fewer omissions (Hypothesis 2a) and more repetitions (Hypothesis 2b) in the second stage of their interviews than truth tellers. We also predicted that truth tellers would include more repetitions and fewer omissions when interviewed by a different interviewer in phase two of an interview than when interviewed by the same interviewer (Hypothesis 3a). In contrast, we hypothesised liars would include more repetitions and fewer omissions when interviewed by the same interviewer in phase two of an interview than when interviewed by a different interviewer (Hypothesis 3b). People's recollection is partially influenced by the retrieval cues that they use to search their memory, and the reverse order recall instruction disrupts the recollection of events from a schema (Gilbert & Fisher, 2006). Therefore, we predicted that reverse order recall prompts would

elicit more reminiscent details than chronological order interviews for truth tellers, but would not affect the amount of reminiscent detail reported by liars (Hypothesis 4), due to their concerns about consistency (Hartwig et al., 2007). Finally, we predicted that liars would include fewer repetitions and more omissions in reverse order recall attempts than in chronological order recall attempts, whereas this manipulation would not affect truth tellers' repetitions and omissions (Hypothesis 5).

Method

Participants

A total of 160 participants (49 male, 109 female, and 2 non-binary, with ages ranging from 18 years to 68 years, $M = 24.09$ years, $SD = 9.93$ years) were recruited from a university in the United Kingdom (126 participants were undergraduate students). Opportunity sampling was used for convenience. Participants received either extra course credit as part of a research participation scheme or £5 compensation if they were not eligible for course credit. **Design**

We used a $2 \times 2 \times 2 \times 2$ mixed design, with Veracity (truth teller vs liar), Order of Recall (chronological order recall vs reverse order recall) and Interviewer (same second interviewer vs different second interviewer) as between subject factors, and Interview Phase (i.e. detail reported in Phase 1 vs detail reported in Phase 2) as a within-subject factor. Interviews were transcribed for analysis, and the number of details in each phase, and the repetitions, omissions and reminiscences found in Phase 2 were analysed as the dependent variables.

Materials

Stimulus materials. Truth tellers were asked to watch a BBC nature documentary clip. The documentary concerned caterpillar and butterfly survival, lasted 6 min and 32 s, and was rich in visual and narrative detail. A single page written synopsis of the documentary was also prepared in the style of a review (462 words).

Post interview questionnaires. Two questionnaires were presented to participants after they had finished both phases of their interview. One questionnaire focused on participants' demographic details (age, gender identification, profession), the other questionnaire asked about how well participants remembered what they had done (i.e. the tasks they had completed) and how they perceived the interview process. Nine questions

included 7 point Likert scales (1 = not at all to 7 = completely) regarding how much they remembered of the tasks they completed, their motivation to be believed, and the extent to which they thought they were believed. Two open-ended questions asked participants to report on the strategies they adopted when providing their accounts. The questionnaires can be found at osf.io/ kx82u/.

Procedure

Pre-interview instructions. After giving consent, participants were then presented with two envelopes; one labelled 'liar' and the other labelled 'truth teller'. They were informed that each envelope contained a set of keys to a room on the second floor of the building they were in, and a set of instructions regarding the tasks that they would be asked to complete, and then either later lie about in an interview (the 'liar' condition), or tell the truth about (the 'truth teller' condition). They were told that the instructions asked liars to steal a USB stick from an empty room, whereas truth tellers would be asked to watch a short nature documentary in an empty room. The researcher then left the room to give the participants time to decide which envelope (and, therefore, task) to select. They then placed the other envelope into a metal tin. Participants were informed that the researcher would not know which condition they chose until after they had completed the experiment. The majority of

participants (n = 130) selected their preferred experimental condition in this way.[2]

Following the written instructions in the envelope truth tellers went alone to a research room and accessed a link to the nature documentary on a computer. Once the documentary was finished, truth tellers then logged out of the computer, and returned to the room in which they had originally met the researcher.

---

[2] To correct for imbalances in the sample size between conditions, some participants signed up for the experiment having read an advert about solely the 'liar' condition (n = 30), and thus were not provided with the option to choose the 'truth teller' envelope. All truth tellers chose their condition (n = 87). Analyses to check for the presence of confounds differences between those who chose to lie compared to (i) those who chose to tell the truth and (ii) those who were given no choice but to lie are presented in the results section.

In contrast, but also following written instructions, liars were directed to another room where they were asked to steal a classified USB stick. They were informed that the USB stick contained important computer coding regarding the privacy and safety of the University's intranet firewall, and they were asked to log on using stolen credentials to check the code before returning to the researcher. Liars were also provided with an outline of the nature documentary that they were informed truth tellers would watch. It was suggested that this synopsis might help them to claim as an alibi, during their interviews, that they were watching the documentary. Procedures where liars are provided with information they might use for an alibi (often a task completed by participants in a truth-telling condition) are common in deception research (e.g. Nahari & Vrij, 2014; Granhag, Strömwall & Jonsson, 2003). Once the liars had completed the tasks on the computer and logged off, they were also instructed to return to the room in which they had originally met the researcher.

When participants returned they were then taken through to the interview room.

The interview. Interviews were conducted by two male interviewers blind to the participants' veracity status. All participants were interviewed individually. All interviews were recorded using a digital recorder. A copy of the interview script is available at osf.io/dxfas/.

The first phase of each interview lasted on average 3 min and 33 s ($SD = 1$ min 41 s).

Participants were not offered preparation time, and were told that they needed to convince the interviewer that they could not have stolen the USB device, as they had been watching the nature documentary. Firstly, the interviewer introduced himself, before informing the participants that some classified data had just been stolen. Participants were then asked to recall what they had been doing for the past 10 min. As expected, all participants claimed to have been watching a nature documentary. A second question asked participants to describe the documentary in chronological order, without guessing, and including as much detail as

possible. Interviewers were instructed not to interrupt participants, or verbalise any encouragement, but could nod to indicate they were paying attention.

After participants had finished describing the documentary, they were thanked and asked to wait in the same room with the instruction that 'someone will be with you shortly'. Participants were left unattended for 3 min, before the second phase of each interview began. At this point, either the same person who conducted the first phase of the interview (same interviewer condition, $n = 79$), or another interviewer (different interviewer condition, $n = 81$) entered the room. Participants were informed that there had been a problem with their previous statement, and that they needed to provide their report about the documentary again. Half of the participants received the same free recall instruction as in the first phase (i.e. chronological order, include as much detail as possible, don't guess, $n = 81$), whereas the remainder of the interviewees were asked to recall the documentary in reverse order ($n = 79$). For the 'reverse order' condition, the interviewer explained what recalling the events in reverse order meant, and then participants were given an example of how one would describe making a cup of tea in reverse order. The interviewer asked the participants to confirm that they understood what was required, before instructing them to include as much detail as possible without guessing. The second phase of the interview lasted on average 3 min and 6 s ($SD = 1$ min 36 s).

Post interview instructions. Once the interviews were complete, participants were

asked to fill in the two post interview questionnaires, and to be completely honest in their responses. Following this they were fully debriefed, paid or awarded credits, and thanked for their time.

Coding of interview transcripts

All audio recorded interviews were transcribed for analysis. Each interview transcript was first coded for number of details provided. (e.g. 'the video had a <u>white</u> <u>male</u> <u>narrating</u>'

would contain three details; the descriptors 'white' and 'male', and the action 'narrating'). Then, following Fisher et al. (2013), each detail provided in the second phase of the inter view was coded for consistency. Repetitions were details reported in both phases of the interview, omissions were details reported in the first phase but not in the second phase of the interviews, reminiscences were details reported in the second phase but not in the first phase of the interviews, and contradictions were details reported in the first phase that were reported differently in the second phase. Due to very few contradictions being reported across conditions, this consistency category will not be considered further in the analyses. A subset of 40 interviews (25%) were coded by a second researcher, who was blind to the experimental conditions. The inter-rater reliability between the coders was high for details in phase one of the interviews (intra-class correlation coefficient [ICC]=.83) and details in phase two of the interviews (ICC=.76). Once again, these details were further classified for consistency, with satisfactory reliability being found between the two coders for repetitions (ICC = .73), reminiscences (ICC = .59) and omissions (ICC = .76).

Results

Manipulation checks

Assignment to condition. A one-way MANOVA was conducted with condition allocation method for Liars (free choice vs forced choice) as the independent variable and total unique detail, detail provided at phase one, detail provided at phase two, omissions, reminiscences

and repetitions as dependent variables to examine the effect of free choice or having been assigned a condition on liars' performance. The multivariate effect was not significant, $F(5, 67) = 1.92$, Wilks' $\Lambda = .88$, $p = .10$, $f = .38$. A series of manipulation checks were also conducted to examine whether the participants who chose to be a liar ($n = 43$) were significantly different to those who were assigned to the liar condition ($n = 30$) in motivation,

confidence, age, gender or number of details provided. An independent measures t-test found that there was no significant difference between motivation levels of those who chose to lie compared to those who were assigned to the liar condition, $t(71) = .310$, $p = .76$, $d = .07$. Furthermore, there was no significant difference in the confidence levels, $t(71) = .37$, $p = .72$, $d = .09$, the age of free choice vs forced choice liars, $t(56.54) = 1.78$, $p = .08$, $d = .39$, the amount of detail provided in the first phase of the interview, $t(71) = 1.94$, $p = .15$, $d = .33$, or the second phase of the interview, $t(67.01) = 1.96$, $p = .054$, $d = .44$. A chi-squared test using Fisher's exact test found that there were no significant differences in the gender divide of those who chose to be a liar compared to those who were assigned to the liar condition, $p = .42$. The data for participants who chose to be liars and those who were assigned as liars were therefore collated in the subsequent analyses. An independent measures t-test found no significant difference in the age of those who chose to be truth tellers compared to liars, $t(156) = .66$, $p = .51$, $d = .11$, and a chi-squared test using Fisher's exact test found no significant difference in the gender representation within each experimental cell, $p = .70$.

Motivation. A 2 (Veracity: truth teller vs liar) × 2 (Order of Recall: chronological order recall vs reverse order recall) × 2 (Interviewer: same second interviewer vs different second interviewer) ANOVA revealed no significant main or interaction effects on self reported motivation scores, all F's < 3.08, all p's > .08. The grand means reflected high levels of motivation ($M = 5.87$, $SD = .94$, 95% CI [5.72, 6.02]), with 74.5% of the sample reporting themselves as 'very' (6) or 'completely' (7) motivated.

Confidence in performance. A 2 (Veracity: truth teller vs liar) × 2 (Order of Recall:

chronological order recall vs reverse order recall) × 2 (Interviewer: same second interviewer vs different second interviewer) ANOVA revealed that truth tellers reported higher confidence levels than liars (see Table 1). No other significant main or interaction effects were found for confidence in performance, all F's < 2.38, all p's > .13.

Memory for stimulus. A 2 (Veracity: truth teller vs liar) × 2 (Order of Recall: chronological order recall vs reverse order recall) × 2 (Interviewer: same second interviewer vs different second interviewer) ANOVA revealed that truth tellers reported greater memory strength than liars (for truth tellers: self-reported memory strength for the video; for liars: self-reported memory strength for the synopsis of the documentary, see Table 1). There were no other significant main or interaction effects for memory strength for the stimulus, all F's < 1.45, all p's > .23

Table 1.

The difference in performance of truth tellers and liars.

	Truth tellers			Liars						
	M	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*F*	*p*	*f*	*90% CI f*
Details reported in Phase 2	101.28*	62.19	[80.30, 122.36]	80.54*	46.17	[63.45, 97.63]	5.42	.02*	.19	[.05, .32]
Total Unique Details	185.53*	88.35	[155.58, 215.48]	156.92*	61.91	[154.36, 159.48]	4.91	.03*	.17	[.04, .31]
Omissions	84.59	59.36	[72.23,							

	M	SD	[95% CI]	M	SD	[95% CI]	F	p	[95% CI]
			96.82]	76.66	50.56	[65.62, 88.68]	.07	.79	.00[.00, .14]
Repetitions	54.93	37.69	[47.22, 63.61]	50.14	32.83	[43.02, 57.26]	.05	.82	.00[.00, .12]
Reminiscences	46.01*	33.67	[39.88, 52.83]	30.12*	28.07	[24.44, 37.75]	8.40	.004*	.25[.10, .37]
Confidence	5.75	.96	[5.43, 6.08]	3.51	1.43	[2.98, 4.04]	138.65	.001*	.95[.79, 1.11]
Memory	5.95	.65	[5.72, 6.18]	5.52	1.16	[5.09, 5.95]	8.93	.004*	1.00[.11, .38]

Notes: Statistically significant differences are highlighted with an asterix.

Reported strategies. Overall, 38.4% of the sample reported using a strategy. Truth tellers were less likely to use a strategy, with 23% reporting using a strategy, compared to 56.9% of liars, a difference in proportions of .34, $p < .001$. Thematic examination of the reported strategies found the most common technique to relate to remembering/visualising the documentary (for truth tellers) or the summary of the documentary (for liars) in a lot of detail (reported by 54.1% of those who used a strategy), followed by considerations of nonverbal behaviour (reported by 31.14% of those who used a strategy). Other reported strategies included reporting truthful details (11.48% of those who used a strategy), and not

overthinking (3.28% of those who used a strategy).

Veracity and reporting

A series of analyses were run to examine the difference in performance of truth tellers and liars on a number of factors (see Table 1). An independent measures *t*-test was conducted to compare the amount of detail provided in the first phase of the interview between truth tellers and liars. There was no significant difference in the amount of detail provided by veracity condition, $t(158) = 1.24$, $p = .22$, $d = .20$, and consequently Hypothesis 1a, that truth tellers would provide more details then liars in the first phase of their interviews, was not supported.

A 2 (Veracity: truth teller vs liar) × 2 (Order of Recall: chronological order recall vs reverse order recall) × 2 (Interviewer: same second interviewer vs different second interviewer) ANOVA revealed that truth tellers included more details in the second phase of their interviews than liars (see Table 1).

The number of unique details provided across both interview phases was calculated by combining the number of details reported in Phase 1 with the number of reminiscent details reported in Phase 2. No hypotheses were formed regarding reporting of unique detail, however exploratory analyses were run. A 2 (Veracity: truth teller vs liar) × 2 (Order of

Recall: chronological order recall vs reverse order recall) × 2 (Interviewer: same second interviewer vs different second interviewer) ANOVA revealed that truth tellers reported more unique details in total than liars (see Table 1). No further effects were found, all *F*'s < 2.84, all *p*'s > .09.

Whilst there were no significant differences in the number of omissions or repetitions that truth tellers and liars provided, a 2 (Veracity: truth teller vs liar) × 2 (Order of Recall: chronological order recall vs reverse order recall) × 2 (Interviewer: same second interviewer vs different second interviewer) ANCOVA with the amount of detail provided in Phase 1 of the interview as a co-variate revealed that truth tellers included more reminiscence than liars

(see Table 1). This supports Hypothesis 1b that truth tellers, more so than liars, would provide reminiscences (additional detail) in the second phase of their interviews. Examinations of interactions and main effects for Order of Recall and Interviewer are presented below.

Details analyses

We conducted a 2 (Veracity) × 2 (Order of Recall:) × 2 (Interviewer) × 2 (Interview Phase) mixed ANOVA with Interview Phase (i.e. detail reported in Phase 1 vs detail reported in Phase 2) as a within subject factor. Findings indicated a significant difference between the amount of detail reported in Phase 1 ($M = 133.83$, $SD = 65.52$) and the amount of detail reported in Phase 2 ($M = 91.82$, $SD = 56.26$), $F(1, 152) = 74.38$, $p < .001$, $f = .70$. There was also an interaction effect of Order of Recall and Interview Phase, $F(1, 152) = 18.26$, $p < .001$, $f = .35$. To explore this interaction further, we broke it down to examine the effect on Phase 1 and Phase 2 of the interviews independently. Statistical significance of a simple main effect was accepted at a Bonferroni-adjusted alpha level of .025, in order to not capitalise on probabilities (Cramer et al., 2016). There was a statistically significant main effect of Order of Recall for detail reported in Phase 2, $F(1, 158) = 5.82$, $p = .02$, $f = .15$, but not in Phase 1,

$F(1, 158) = 3.74$, $p = .06$, $f = .19$. Mean detail reported in Phase 2 was higher for recall in chronological order than recall in reverse order, with a mean difference of 21.15, 95% CI [3.84, 38.45], $p = .02$. There was also a three way interaction identified between Interview Phase, Veracity and Interviewer, $F(1, 152) = 4.44$, $p = .04$, $f = .17$. To explore this interaction further, we examine Phase 1 and Phase 2 separately. Statistical significance of a simple two way interaction was accepted at a Bonferroni-adjusted alpha level of .025. There were no statistically significant simple two-way interactions of Interviewer and Veracity for detail reported in Phase 1, $F(1, 156) = 3.86$, $p = .051$, $f = .16$, or for detail reported in Phase 2, $F(1, 156) = .01$, $p = .94$, $f = .00$.

We examined the amount of detail provided in Phase 2 of the interview using a 2

(Veracity: truth teller vs liar) × 2 (Order of Recall: chronological order recall vs reverse order recall) × 2 (Interviewer: same second interviewer vs different second interviewer) ANOVA and found a further main effect for Order of Recall, $F(1, 158) = 6.01$, $p = .02$, $f = .20$. Those who reported their second statement in chronological order included more details overall ($M = 102.26$, $SD = 55.52$, 95% CI [82.75, 121.77]) than those who reported it in reverse order ($M = 81.11$, $SD = 55.31$, 95% CI [61.43, 100.79]). No further main or interaction effects were found, all F's < 1.58, all p's > .21.

Consistency analyses

Consistency is related to the volume of information that is reported in the initial phase of an interview. For example, if more details are reported in phase one of the interview then there are subsequently more opportunities for repetitions or omissions and fewer opportunities for reminiscences in the second phase of the interview. To control for this, we analysed the effect of Veracity, Order of Recall and Interviewer upon repetitions, omissions and reminiscences with the amount of detail provided in phase one of the interview as a covariate. Contradictions did not occur frequently enough for inclusion in the analyses.

54

Omissions. A 2 (Veracity: truth teller vs liar) × 2 (Order of Recall: chronological order recall vs reverse order recall) × 2 (Interviewer: same second interviewer vs different second interviewer) ANCOVA with the amount of detail provided in Phase 1 of the interview as a co-variate revealed a main effect of Order of Recall on the number of omissions, $F(1, 151) = 24.02$, $p < .001$, $f = .40$. Those who reported in reverse order exhibited more omissions ($M = 99.19$, $SD=55.33$, 95% CI [87.79, 112.81]) than those who reported in chronological order ($M = 63.20$, $SD = 49.87$, 95% CI [53.11, 73.98]). No other main effects were found, all F's < 2.74, all p's > .10, and consequently Hypothesis 2a, that liars would prioritise consistency and therefore have fewer omissions than truth tellers, was not supported.

Repetitions. A 2 (Veracity: truth teller vs liar) × 2 (Order of Recall: chronological

order recall vs reverse order recall) × 2 (Interviewer: same second interviewer vs different second interviewer) ANCOVA with the amount of detail provided in Phase 1 of the interview as a co-variate revealed a main effect for Order of Recall on the number of repetitions, $F(1, 151) = 23.44$, $p < .001$, $f = .40$. Individuals who reported in chronological order included more repetitions ($M = 60.60$, $SD = 33.68$, 95% CI [53.28, 68.14]) than those who reported in reverse order ($M = 44.68$, $SD = 35.76$, 95% CI [37.26, 52.75]). No further effects were found, all F's < 2.57, all p's > .11, and consequently Hypothesis 2b, that liars would prioritise consistency and include more repetitions than truth tellers was not supported.

The Veracity × Interviewer interaction was not significant, $F(1,151) = 2.57$, $p = .11$, $f = .00$, and therefore no support was found for Hypothesis 3a that truth tellers would include more repetitions and fewer omissions when interviewed by a different interviewer in phase two compared to the same interviewer, or Hypothesis 3b that liars would include more repetitions and fewer omissions when interviewed by the same interviewer compared to a different interviewer.

Reminiscent details. A 2 (Veracity: truth teller vs liar) × 2 (Order of Recall: chronological order recall vs reverse order recall) × 2 (Interviewer: same second interviewer vs different second interviewer) ANCOVA with the amount of detail provided in Phase 1 of the interview as a co-variate revealed a main effect of Interviewer on number of reminiscences, $F(1, 151) = 4.56$, $p = .03$, $f = .17$, which revealed that those interviewed by the same interviewer included more reminiscences ($M = 45.15$, $SD = 38.20$, 95% CI [36.89, 54.03]) than those interviewed by a different second interviewer ($M = 32.53$, $SD = 23.48$, 95% CI [27.87, 37.58]). No further effects were found, all F's < 2.36, all p's > .13.

Hypothesis 4 predicted that reverse order recall would elicit more reminiscences than chronological order recall in truth tellers. To specifically test this hypothesis we compared the number of reminiscences provided in reverse order and chronological recalls for truth tellers only using a between subjects t-test. No significant difference was found, $t(85) = .65$,

p = .52, *d* = .14 with truth tellers including an average of 43.68 (*SD* = 36.04) reminiscent details in reverse order interviews, and an average of 48.40 (*SD* = 31.30) reminiscent details in chronological order.

Hypothesis 5 predicted that liars would make more omissions during reverse order recall than during chronological order recall. We examined this using a between subjects *t* test and found that liars made significantly more omissions during reverse order recall (*M* = 96.97, *SD* = 55.42, 95% CI [67.76, 126.18]) than during chronological order recall (*M* = 57.95, *SD* = 37.45, 95% CI [38.74, 77.16]), *t*(71) = 3.55, *p* = .001, *d* = .84.

Discussion

Replicating previous findings in deception research, truth tellers provided more details than liars in the second phase of their interviews, as well as over both phases of the interview combined (Amado et al., 2015; Vrij, 2008). However, fewer details were reported

by both truth tellers and liars in the second phase of the interview when compared to the first phase, and there were no differences in the number of details provided by truth tellers or liars in the initial phase of the interviews. Our participants did not know they were going to be re interviewed, and initially liars and truth tellers provided an equivalent amount of detail. When prompted to discuss the matter further, truth tellers were more likely than liars to expand their testimony to include new information, despite both groups providing less detail than in the previous phase of the interview. This expansion on testimony is often found in the 'encouraging interviewees to say more' deception literature (Vrij et al., 2017). It is interesting to note the decrease in detail provided during the second phase of the interview by both truth tellers and liars, despite the increase in reminiscent details for truth tellers compared to liars, and no veracity-based differences for omissions or repetitions. We based our design on that of Leins et al. (2011) and Leins et al. (2012), who also examined within-statement consistency using two accounts taken from the same interviewing session. However, we

instigated a shorter break between phases of the interview, and this could be considered a limitation of the design, resulting in a lack of motivation for participants to repeat the story in full again.

In addition, truth tellers included more reminiscent details during the second phase of their interviews than liars, but there were no significant differences between truth tellers and liars for the number of repetitions or omissions (see also Granhag et al., 2003). This inclusion of reminiscent details suggests that truth tellers, more than liars, prioritised providing new information in Phase 2 of their interviews, instead of simply maintaining their original narrative. Reminiscences are commonly observed in memory research (Gilbert & Fisher, 2006) and the current findings provide further support for the failure of liars to account for natural side effects of rudimentary memory systems (Harvey, Vrij, Leal, Hope, & Mann, 2017). If we define statement consistency as the presence of repetitions and the absence of

omissions and reminiscences, the findings suggest that liars were more consistent than truth tellers due to the relative absence of reminiscences in liars' accounts. This lack of consistency for truth tellers is generally found across the literature (see Vredeveldt et al., 2014 for an overview), although remains at odds with the popular belief that deception involves a lack of consistency.

In the current experiment, we introduced two manipulations during the interview process and examined their effects on liars' and truth tellers' statements: (*i*) introducing a prompt for reverse order recall, similar to Vrij et al. (2012) and (*ii*) switching the interviewer, similar to Shaw et al. (2014). We were particularly interested in whether these manipulations would induce inconsistencies in liars. We found little evidence to support this notion. This may have been due to limitations with the liars' task, whereby liars were required to read a review-style synopsis of a Sir David Attenborough documentary, and to pretend that they had watched it. Due to the immense popularity of Sir David Attenborough's work in the United Kingdom (Smith, 2018), it is possible that liars were familiar with the style of Sir David

Attenborough's work, and therefore able to perform at an overall similar level to the truth tellers.

When interviewed for the second phase of the interview, all individuals reporting in chronological order included more details and more repetitions than those reporting in reverse order. The number of omissions also increased for all individuals asked to recall in reverse order rather than chronological order, but the type of recall had no discernible effect on the number of reminiscences. The increase in omissions is most likely due to the increased cognitive load that results from reporting events in reverse order. The reverse order technique was designed as a memory enhancing task as part of the Cognitive Interview (Fisher & Geiselman, 1992) designed to break script reliance and help trigger items without script rehearsal. Theoretically, this could have led to a greater number of reminiscences in truthful

interviews, as honest interviewees were required to examine their memory utilising different retrieval processes to trigger further recollections (Fisher et al., 2009), though in this study we found no effect of Order of Recall on reminiscences. There was a significant effect of Order of Recall for liars, who exhibited significantly more omissions when reporting in reverse order. However, there was no significant difference in the number of omissions between truth tellers and liars when providing either chronological order or reverse order recalls, which suggests both groups found the tasks equally difficult. We speculate that this may be due to the brevity of the break between interview phases, whereby the information that participants had provided in Phase 1 of the interview was still fresh in their mind.

The number of reminiscent details reported increased for individuals interviewed by the same second interviewer, compared to those who were interviewed by a different interviewer in the second phase of their interview. One possible explanation for this finding is that the interviewees thought that the same interviewer had returned because s/he was not satisfied with the amount of information s/he had initially gathered. This suggestion is speculative though, as there could be a variety of explanations for this increase in

reminiscence, such as the context of talking to the same person again, or the familiarity of being interviewed by the same person. However, there were no significant differences in the number of repetitions provided, or omissions made by veracity condition as a result of changing interviewer, which may be due to the relatively short nature of the interviews. Establishing the effect of changing interviewers on statement consistency is important, considering how many people interview individuals being processed through the criminal justice system (e.g. investigators, lawyers, psychologists), and who operate with the belief that consistency is indicative of honesty (Bogaard & Meijer, 2017). A better examination of the effect of this would be to compare changing or maintaining interviewers when

interviewees are providing longer statements for personally experienced episodic events, as would likely be found in the criminal justice system.

The design of this study was somewhat similar to that of Shaw et al. (2014), although we were unable to replicate the effects they found when switching interviewer. Shaw et al. (2014) found that when the interviewers were changed partway through the interviews, truth tellers provided more detail in the second phase of the interviews than when presented with the same interviewer, and that their answers across both phases of the interviews included more repetitions than the answers provided by liars. In the current study, however, we only noted a difference in the number of reminiscences provided by interviewees due to changing interviewer. There are a number of methodological differences between the studies that might account for this discrepancy. First, Shaw et al. (2014) asked participants to fabricate information about an experience rather than providing them with information on which to base their alibi. Providing our participants with this alibi information prevented interviewees from crafting their own deception, which may be where the source of variance in performance arises. In addition, Shaw et al. (2014) used a coding system whereby participant statements were scored with respect to 18 key points and then subjectively rated for

consistency. In comparison, our statements were objectively coded for every detail that the participant mentioned, which provided more opportunities for consistencies or inconsistencies to arise.

Most participants were provided with the opportunity to select their veracity condition, however there was a small subsection who were recruited in the knowledge they would be asked to lie. We propose that the latter group of participants were also afforded the option of participating as a liar, but this choice was presented at an earlier stage when they volunteered to participate knowing they would be required to lie. Research has shown there to be no difference in performance between those exhibiting forced and voluntary dishonesty

(Geven, Selle, Ben-Shakhar, Kindt, & Verschuere, 2018; Nahari, Breska, Elber, Klein Selle, & Ben-Shakhar, 2017). Whilst the mixed nature of having free choice liars and forced liars may be considered a limitation, as there were no differences in the demographics, motivation or performance of participants in the current study, it was reasonable to combine these two groups for analysis. Furthermore, providing participants with the choice to be a liar or truth teller could be considered a limitation, as it may have introduced a systematic bias whereby the individuals who volunteered as liars were more practiced, and therefore possibly more skilled at deception. However, we argue that this choice is reflective of real life, where individuals make a decision as to whether they wish to lie about something. Therefore, those who may be uncomfortable with lying would be less likely to lie. When examining who chose to be a liar compared to those who chose to be a truth teller, we found no differences in background characteristics such as gender (consistent with DePaulo et al., [1996]) or age, nor did we find any difference in motivation levels across experimental cells, suggesting that those who chose to tell the truth, lie or were asked to lie were all equally motivated to succeed. We encourage further research into the impact of choice upon deceptive behaviour, and suggest future studies could consider examining the performance of both free choice and forced truth tellers and liars.

Considering the number of individuals who work within the criminal justice system, and operate under the belief that consistency is indicative of honesty (Bogaard & Meijer, 2017; Strömwall & Granhag, 2003), examining the statement consistency for truthful and deceptive suspects is important to help practitioners make informed veracity judgements. In the current study, we investigated the influence of recall order and a change in interviewer on the within-statement consistency of both truth tellers and liars. This has important implications for those involved in the criminal justice system, where suspects may be interviewed multiple times by different people, and be subject to a variety of interviewing

techniques throughout their interviews. No effects were found for changing the interviewer and changing the recall order on repetitions and omissions. It seems therefore that the combination of these interviewing techniques is not necessarily beneficial for enhanced deception detection.

Despite the popular belief that inconsistency is a cue to deception, we also found little support for the notion that consistency (or lack of consistency) offers diagnostic cues to deception. If anything, due to the natural occurrence of reminiscences in our truth tellers' statements, as is commonly found in memory research (Gilbert & Fisher, 2006), our findings suggest that liars are more consistent than truth tellers, and that the consistency heuristic is an inconsistent and impractical cue to deception.

CH is responsible for the design, material development, data collection, analysis and manuscript preparation of Experiment 1. AV, LA, and LH provided guidance and feedback on experimental design and comments on the manuscript draft.

Chapter III

An examination of the Self-Administered Interview (SAI) as a verbal veracity assessment tool.

Abstract

The self-administered interview (SAI) is a written eyewitness recall tool that elicits more information from cooperative witnesses than written free recall (WFR) formats. To date, SAI research has examined the accounts of cooperative people providing honest reports. In the current experiment, truthful *and* fabricating participants ($N = 128$) either completed a WFR or a SAI after witnessing a crime (initial account). After a 1-week delay, participants were interviewed verbally (subsequent interview). Truth tellers reported significantly more detail than liars in both the initial account and subsequent interview, and participants who completed the SAI reported more detail than those completing the WFR. Truth tellers repeated and omitted more information in the subsequent interview than liars; however, there was no significant difference in the number of reminiscent details reported. Although the SAI is effective in eliciting information as an initial eyewitness reporting tool, no benefits for the detection of deception were demonstrated.

Introduction

Eyewitness evidence is often crucial for police investigations. In the case of serious events, the number of eyewitnesses can outstrip the level of police resources available at the scene. Ideally, eyewitness testimony should be gathered as soon as possible after an event to reduce memory decay and exposure to post-event misinformation, which may compromise the quality and quantity of later recall (Hope, Gabbert, & Fisher, 2011). Although investigators try to prioritise managing witnesses at the scene, there are often factors, such as a large volume of witnesses to deal with, that restrict comprehensive eyewitness interviews from taking place for days, or even weeks, after the event (Hope, Gabbert, Fisher, & Jamieson, 2014). Post-event misinformation in high-stress environments can spread especially quickly through modern social networks such as Twitter (Huang, Starbird, Orand, Stanek, & Pedersen, 2015), a site heavily used during high-profile incidents (Saleem, Xu, & Ruths, 2014), which increases the need for authorities to gather comprehensive witness reports as soon as possible.

The majority of research into eyewitness testimony has been conducted in the context that, on the whole, witnesses try to provide genuine reports, and that any inaccuracies that may arise in their statements are the result of honest mistakes. Yet, in some circumstances, hostile or uncooperative individuals may deliberately provide misleading reports to derail investigations. Law enforcement professionals often try to discern the accuracy of witness' statements to help direct the investigation (Desmarais & Yarmey, 2004; Masip & Herrero, 2015). While many witnesses are cooperative and can be trusted to provide credible and reliable accounts, in certain contexts or circumstances, witnesses who are hostile to the aims of the police or investigation may intentionally mislead investigators. They may do this for a variety of reasons including attempts to protect the perpetrator, fear of retribution from the perpetrator, or even distrust of the authorities (Parliament & Yarmey, 2002; Yarmey, 2004).

It has been shown across a variety of modalities that truth tellers often report more information than liars (Amado, Arce, Fariña, & Vilarino, 2016; DePaulo et al., 2003). Liars may lack the creativity or imagination to fabricate details to an equivalent level to that provided by truth tellers (Vrij, Hope, & Fisher, 2014), or may intentionally minimise the amount of detail reported out of fear that additional detail may provide leads for investigators (Vrij, Granhag, Mann & Leal, 2011). Liars may also decide to report as little information as possible, to reduce the opportunity to contradict themselves in subsequent interviews (Vredeveldt, van Koppen & Granhag, 2014). Thus, liars' accounts are often shorter and less detailed than those typically provided by truth tellers.

In repeated interviews, details can be compared and classified into four categories to discern consistency; repetition, reminiscences (sometimes referred to as commissions in deception research), omissions and contradictions (Fisher, Vrij, & Leins, 2013). The occurrence of repetitions (information that is reported during both interviews) increases consistency, and the occurrence of reminiscences (information that is reported in a subsequent interview, but not reported in the primary interview), omissions (information that is provided during an initial interview, but not reported in a subsequent interview) and contradictions (information provided in the subsequent interview that directly opposes what was reported in the primary interview) decrease consistency.

Perceptions of consistency can influence credibility judgements (Reinhard & Sporer, 2008), with 8 out of 10 police officers reporting that they believe consecutive statements given by the same individual will be more consistent if the individual is telling the truth as opposed to telling a lie (Strömwall & Granhag, 2003). In fact, when more than one statement is available, judges rely on the perceptions of consistency as a cue to veracity more than any other cue (Strömwall, Granhag & Jonsson, 2003), despite research frequently finding there is very little difference between truth teller and liar consistency (Granhag & Strömwall, 2002;

Granhag, Strömwall, & Jonsson, 2003; Granhag, Mac Giolla, Sooniste, Strömwall, & Liu-

Jonsson, 2016). For instance, Granhag and Strömwall (2002) found that over three interrogations truth tellers repeated more details and omitted more information that liars did, and that there was no difference in the amount of reminiscent detail reported by truth tellers and liars. As such, truth tellers' and liars' statements were about equally consistent over time, which was corroborated by the consistency ratings that the truth tellers' and liars' provided about their own statements. Conversely, Granhag et al. (2003) found that there was no difference between the number of repetitions or omissions given by truth tellers and liars over two interrogations, and that truth tellers provided more reminiscent detail in the second interrogation than liars did. Furthermore, when the statements were subjectively rated for consistency, liars' and truth tellers' statements were perceived equally consistent. Granhag et al. (2016) also found that the consistency of truth tellers and liars was similar when examining repetitions, reminiscences and omissions. Yet when Masip et al. (2018) asked uninformed laypeople to make veracity judgements on a series of written statements, they found that 90% of the laypeople reported using consistency/inconsistency to assist in making their judgement. This is incongruous with research examining truth tellers' and liars' consistency, as well as memory research, which suggests that some types of inconsistencies, such as omissions and reminiscence, are commonly found in memory accounts (Fisher, Brewer & Mitchell, 2009; Gilbert & Fisher, 2006; Strange, Dysart, & Loftus, 2014).

When individuals were interviewed about witnessing a videotaped mock crime, 98% of truthful participants included reminiscence in their second recall (Gilbert & Fisher, 2006). Liars, however, may be less likely to include new details in subsequent statements, in order to maintain a greater level of consistency, or due to an unfamiliarity with natural memory phenomena (Harvey, Vrij, Hope, Leal, & Mann, 2017). The repeat versus reconstruct hypothesis (Granhag & Strömwall, 1999) suggests that liars have the aim of being consistent

across interviews to present themselves as being honest, and therefore avoid introducing reminiscent detail. However, truth tellers are less concerned with appearing consistent, and

therefore reconstruct the event from memory, thereby introducing reminiscent detail. **The Self-Administered Interview ©**

The Self-Administered Interview (SAI©; Gabbert, Hope & Fisher, 2009) was initially developed as a tool to enable investigators to gather an extensive initial report from cooperative eyewitnesses, either at the scene of an event, or shortly after, without placing any additional strain on available resources. The SAI is a reporting tool that draws on memory theory and empirical research to promote a comprehensive free recall, from witnesses, in their own words. The SAI has been shown to facilitate the reporting of more correct details, compared to a written Free Recall (WFR) statement collection method (Gabbert et al., 2009), and greater consistency (e.g., participants completing the SAI included proportionally more repetitions and fewer reminiscences during a second recall after a 1-week delay than those completing a WFR; Hope et al., 2014), which is beneficial in legal settings where consistency is valued (Fisher et al., 2009). It has been found that completing a SAI shortly after witnessing an event leads witnesses to recall more correct information in a delayed recall test, to report less misleading post-event information, and to be more resistant to misleading questions (Gabbert, Hope, Fisher, & Jamieson, 2012). Research suggests that the SAI should be administered as soon as possible, as recall accuracy decreases and post-event misinformation susceptibility increases when the SAI is administered more than 24 hr after an event (Paterson, Eijkemans, & Kemp, 2015).

Due to the greater amount of detail prompted with the SAI (Gabbert et al., 2009), we predicted that in the initial account all participants completing the SAI would include significantly more details than those completing a WFR (Hypothesis 1), and that statements provided by truth tellers completing the SAI would have significantly higher accuracy rates

than statements provided by truth tellers completing the WFR (Hypothesis 2). With research indicating that truth tellers report more details than liars (Amado et al., 2016; DePaulo et al., 2003), it was predicted that truth tellers would include significantly more repetitions in their

subsequent interview than liars, particularly if they initially completed the SAI, as they would have initially reported more details which they could repeat (Hypothesis 3). It was also predicted that as reminiscences are commonly found in repeated honest recalls (Fisher et al., 2009; Gilbert & Fisher, 2018; Strange et al., 2014), truth tellers will include significantly more reminiscences in their subsequent interview than liars, who may avoid including new information, as they wish to be perceived as consistent (Granhag & Strömwall, 1999), particularly if they initially completed the SAI (Hypothesis 4). Liars may feel that by completing the SAI, they have already provided an adequately detailed account and be reluctant to expand their statement further, in an effort to keep their story simple (Granhag & Strömwall, 2002; Hartwig, Granhag & Strömwall, 2007).

Method

Participants

A total of 128 participants (45 male, 83 female, with ages ranging from 18 to 74 years, $M = 26.92$ years, $SD = 9.84$ years) were recruited from a British university using opportunity sampling. Participants received a £10 honorarium on completion of the research session. A total of 46 participants were undergraduate students, with the remaining 82 being local community members.

Design

This study used a 2 (Veracity: truth teller vs. liar) × 2 (Initial account type: SAI vs. WFR) × 2 (Time of interview: initial account vs. subsequent interview) mixed design. Veracity and Initial account type were between-subjects measures, and time of interview was

a within-subjects measure as all participants provided both an initial account and subsequent interview. Participants were randomly allocated to either the truth teller or liar condition before being asked to watch the stimulus video, and were then randomly allocated to either the SAI or WFR condition to give their initial account of what they had witnessed. Following

a delay of 1 week, all participants returned to be interviewed. This study was preregistered on the Open Science Framework (OSF), and all of the materials and datasets can be found at osf.io/fjuzm/. The preregistration can be found at osf.io/y4hfw.

Materials

Stimulus materials. A film was shot in first-person perspective and featured the view of an individual who walks down a footpath in a park, checks a wristwatch for the time, accidentally bumps into another person, receives a phone call, and then witnesses a car hitting a pedestrian in a car park. The viewer then sees the driver exit the car, stab the pedestrian who had been hit by the car, and then shout, wave a knife, and chase other witnesses. The perpetrator comes toward the viewer and speaks directly to them before continuing past. A man dressed in military uniform then approaches the viewer and informs the viewer that this was a terrorist attack, and that they need to complete a written report about what they had witnessed. The event lasted 1 minute 55 seconds.

Initial accounts. Participants in the SAI condition were provided with a copy of the SAI to complete. The SAI contains several sections (as described in Gabbert et al., 2009), and asks for descriptions of; what happened, the scene, people present at the scene, the perpetrator, any vehicles involved, how well participants could see the incident, and any other information about the event. The first section promotes the importance of following the instructions, highlighting the requirement to complete the sections in sequential order. In the second section, witnesses are requested to complete a free recall, drawing in the Context Reinstatement and Report Everything components of the Cognitive Interview. The next

section focuses on details regarding the perpetrators while in the following section witnesses are instructed to provide a sketch the scene. In the final section, witnesses are asked to report aspects they may have not previously considered, such as viewing conditions at the scene (e.g., distance from incident, weather, etc.). Participants completed the SAI in an average of 25 minutes 18 seconds (SD = 9 minutes 24 seconds).

Participants in the WFR condition were provided with instructions for completing the WFR and were given a blank sheet of paper for their statement. Participants were instructed that they should report all of the details about the incident and the people involved that they could remember, and to avoid making guesses about things they were unable to remember. They were advised that they could report the event in whatever manner they preferred (e.g., paragraphs, bullet points, etc.) and that the information that they provided should be as detailed and accurate as possible. Participants completed the WFR in an average of 13 minutes 24 seconds (SD = 7 minutes 44 seconds).

Post-report questionnaires. A 17-item post-report questionnaire was administered after participants had provided their initial account to gather information about their motivation, comprehension of the task, perceived interview performance and memory of the video (rated on 10-point scales). Five questions concerned comprehension of the task (e.g., "I found the instructions easy to understand"), five questions concerned perceived interview performance (e.g., "I was very detailed in my account"), and five questions concerned memory for the video (e.g., "My memory of the video is very clear"). The questionnaire also included an open-ended question about strategies used by truth tellers and liars. **Procedure**

Participants were randomly allocated to act as a truth teller or a liar in the study, and briefed accordingly. Truth tellers were informed that they were going to watch a video of a short fictional attack being conducted by enemy forces, and that they were to treat it as an

event that they were witnessing live. They were told to be completely truthful in their written report to help their team catch the perpetrator. Liars were also informed that they were going to watch a video of a short fictional attack, and to treat it as an event that they were witnessing live. However, they were told that they were an undercover colleague of the perpetrator, and that their role was to protect this colleague by misleading the enemy forces who would investigate the incident. Therefore, they were informed that they should lie about the perpetrator when writing their report, so both they (the participant) and their colleague

(the perpetrator) could evade detection by the enemy forces.

Participants were then invited to put on the virtual reality headset and a set of headphones to watch the video. Virtual reality presentation was used to increase immersion in the event (Bowman & McMahan, 2007). After watching the stimulus film, all participants removed the headset and headphones, and were asked to complete either the SAI or WFR initial account regarding what they had witnessed. Regardless of interview condition, all participants were instructed to complete their account in their own time, using their own words, whilst bearing in mind their objective as either a truth teller or liar. They were shown to a quiet room where they completed the report independently. Upon completing their report, they were asked to complete the demographic and post-report questionnaires. Descriptive statistics and group comparisons for these questionnaires are provided in Appendix 1.

Participants returned to the laboratory 7 days later for their subsequent interview. On arrival they were asked if they remembered their veracity condition (all participants did), and were reminded of their objective. Truth tellers were informed that the person who was going to interview them was on the same team as they were, and consequently they were instructed to provide a completely honest report about what they saw. Liars were informed that the person who was going to interview them was working for enemy forces, and that they should

deceive the interviewer to protect the colleague who conducted the attack. Participants were then introduced to the interviewer and informed that the interview would be audio recorded. All participants were interviewed individually. The interviewer, who had not read the participants' previous accounts, was instructed to elicit as much information as possible using a modified Structured Interview Protocol (SIP; Gabbert et al., 2016). The SIP is a flexible interview protocol derived from best practice policy (e.g., PEACE). It opens with rapport building behaviours (the engage and explain phase of PEACE interviewing) and starts with an open-ended request for interviewees to provide a detailed free narrative. Interviewers are

instructed to prioritise the use of open questions throughout the interview to maximise information gathering. In each interview conducted for the current research, the interviewer asked for an open-ended free recall, and two open-ended prompt questions regarding the critical incident. Interviews took on average 7 minutes 39 seconds (SD = 1 minute 1 second). After the interview, participants were asked to complete the same post-report questionnaire that they had completed after their initial account. They were then fully debriefed, provided with another opportunity to ask any questions, and paid a £10 honorarium. **Coding and analysis**

The transcripts were first coded for detail (specified as "person", "object", "action" and "location", following the coding protocol used by Gabbert, Hope & Fisher, 2009). For example, "a man got out of the car, he was a skinny man" would contain two person details ("man" and "skinny"), one action detail (got out), and one object detail (car). Details were only counted the first time they were mentioned in each account. A total detail score for each account was computed by adding the number of person, object, action, and location details provided. The details found in truth tellers' statements were further coded for accuracy (specified as "correct detail", "incorrect detail", and "confabulation", as used in Vredeveldt,

Baddeley & Hitch, 2013). With respect to liars' statements[3], as liars were required to embed their lies in an account of the same incident (i.e. acknowledge they were there but misrepresent what happened and who was involved), we coded for "truthful" details (i.e. details that were accurate), as well as "fabricated" details (i.e. details that were completely fabricated and of which there was no evidence in the film, such as mention of additional witnesses), and "distorted" details (i.e. amending details in the video, such as altering the description of the perpetrators' clothes, so they were a distorted version of actual details).

For both truth tellers and liars, the details provided in their subsequent interviews were compared to those provided in their initial accounts, and categorised as the four elements of

consistency (specified as "repetition", "omission", "reminiscent", and "contradiction[4]", as described by Fisher, Vrij & Leins, 2013). A subset of 26 interviews (20%) were coded by a second researcher, who was blind to the experimental conditions. Any disagreements between the two raters were discussed and resolved. The inter-rater reliability between the coders was high for detail in the initial accounts (intra-class correlation coefficient [ICC] = .94) and detail in the subsequent interviews (ICC = .91). Satisfactory reliability was also found across the two coders for repetitions (ICC = .86), reminiscences (ICC = .73), contradictions (ICC = .77) and acceptable for omissions (ICC = .51). Truth tellers' accounts were also coded for accuracy, and inter-rater reliability was found to be good for the initial accounts (ICC = .91), and the subsequent interviews (ICC = .86). Liars' initial accounts coded for truthful (ICC = .96), fabricated (ICC = .97) and distorted details (ICC = .95) were also found to have high inter-rater reliability, as were their

[3] This coding scheme is exploratory and was developed after collecting the data to identify the types of details liars used to build their narrative. Given it's exploratory nature, this coding and subsequent analyses were not included in the pre-registration.

[4] Contradictory details were relatively infrequent within the subsequent interviews (as has previously been found by Granhag and Strömwall, 2000, 2001) and were not suitable for meaningful statistical analysis. When contradictions did occur, there was only one contradictory detail in 18.9% of statements, and two contradictory details in 11.8% of statements, with the data being skewed towards absence.

subsequent interviews (truthful ICC = .88, fabricated ICC = .92, and distorted ICC = .94). For each of the analyses, parametric assumptions were checked and met, all p's > .05.

Results

Analyses relating to the hypotheses[5]

Initial account. A 2 (Veracity: truth teller vs liar) x 2 (Initial account type: SAI vs WFR) ANOVA revealed a main effect for Veracity, $F(1, 124) = 12.99$, $p < .001$, $f = .30$, $BF_{(10)} = 13.43$, with truth tellers reporting more details than liars (Table 2). As was predicted in Hypothesis 1, there was also a significant effect of the Initial account type on the number of details provided, $F(1, 124) = 46.04$, $p < .001$, $f = .61$, $BF_{(10)} = 5.33 \times 10^6$, with those

completing the SAI providing significantly more details than those who completed a WFR (see Table 2). No significant interaction effect was found, $F(1, 124) = .35$, $p = .55$, $BF_{(01)} = 2.28$.

Table 2.

Number of details reported in the initial account and subsequent interview (SD) Initial account

	Truth teller	Liar	Total
SAI 48.75 (14.14) 95% CI [43.65, 53.84]	WFR 48.84 (14.54) 95% CI [44.29, 53.40]	33.93 (13.98) * 95% CI [30.84, 37.03]	
WFR 35.06 (12.87) 95% CI [30.41, 39.70]	Total 49.61 (14.08) ** 95% CI [46.39, 52.82]	43.88 (12.24) 95% CI [39.32, 48.43]	
Total 41.91 (15.08) * 95% CI [38.81, 45.00]	Notes: * $p < .001$, ** $p = .001$	38.97 (11.21) 95% CI [34.42, 43.52]	
Subsequent interview	42.09 (12.20) 95% CI [37.69, 46.49]	41.42 (11.90) ** 95% CI [38.20, 44.64]	
SAI 50.38 (13.79) 95% CI [45.82, 54.93]	25.78 (10.55) 95% CI [21.98, 29.58]	45.42 (13.52) *	

95% CI [42.33, 48.52]   47.13 *(13.34)*   95% CI [40.69, 47.13]

30.42 *(12.57)* *   95% CI [43.91, 50.34]

95% CI [27.33, 33.52]   43.91 *(13.81)*

An independent samples *t*-test found no significant difference in the accuracy of the

initial accounts given by truth tellers completing the SAI or a WFR, $t(62) = 1.46$, $p = .55$,

$BF_{(01)} = 3.37$. Therefore, no support was found for Hypothesis 2. Furthermore, there was no

significant difference in the accuracy of the subsequent interviews given by truth tellers

completing the SAI compared to truth tellers completing a WFR , $t(62) = .65$, $p = .52$, $BF_{(01)}$

$= 3.28$.

Repetitions. We ran a 2 (Veracity: truth teller vs liar) x 2 (Initial account type: SAI vs WFR)

ANOVA, to examine the number of repetitions provided by truth tellers and liars. There was

a main effect of Veracity on the number of details that were repeated, $F(1, 124) = 10.64$, $p =$

$.001$, $f = .29$ $BF_{(10)} = 11.23$, with truth tellers providing significantly more

75

repetitions ($M = 30.27$, $SD = 10.87$, 95% CI [27.61, 32.93] than liars ($M = 24.22$, $SD = 10.81$,

95% CI [21.57, 26.87]. We also found a main effect of Initial account type, $F(1, 124) = 9.89$,

$p = .002$, $f = .28$, $BF_{(10)} = 15.60$, with those who initially provided a SAI including more

repetitions in their second account ($M = 30.16$, $SD = 11.77$, 95% CI [27.28, 33.04]) than

those who used the WFR ($M = 24.33$, $SD = 9.89$, 95% CI [21.01, 26.75]). There was no

significant interaction between Veracity and Initial account, $F(1, 124) = .81$, $p = .37$, $BF_{(10)} =$

$.36$, and, consequentially, no support was found for Hypothesis 3.

Reminiscences. We ran a 2 (Veracity: truth teller vs liar) x 2 (Initial account: SAI vs

WFR) ANOVA to examine the number of reminiscences provided by truth tellers and liars.

All participants provided at least one reminiscent detail in their subsequent interview, though

there were no significant main or interaction effects revealed by this analysis, all F's < 3.07, all p's > .08, BF(10) > .76. Therefore, no support was found for Hypothesis 4. **Additional analyses.**

Details provided in subsequent interview. A 2 (Veracity: truth teller vs liar) x 2 (Initial account: SAI vs WFR) ANOVA revealed a main effect for Veracity, $F(1, 124) =$ 12.67, $p = .001$, $f = .32$, with truth tellers reporting more details in their subsequent interview than liars (Table 2). There were no significant Initial account type main effects or interaction effects, all F's < 1.96, all p's > .16.

Comparison of detail across accounts. We conducted a 2 (Veracity: truth teller vs liar) x 2 (Initial account: SAI vs WFR) x 2 (Time: initial account vs subsequent interview) repeated measures ANOVA, with Time as a within subjects factor, on the amount of detail provided in the initial account and subsequent interview. Findings indicated a main effect of Time, with significantly more detail reported in the subsequent interview ($M = 45.52$, $SD =$ 13.62, 95% CI [43.24, 47.79]) than the initial account ($M = 37.92$, $SD = 15.03$, 95% CI [35.73, 40.11]), $F(1, 124) = 101.75$, $p < .001$, $f = .91$.

There was also an interaction effect of Time and Initial account type, $F(1, 124) =$ 61.23, $p < .001$, $f = .70$, such that there was a significant main effect of Time (initial account vs subsequent interview) for those in the WFR condition, $F(1, 124) = 160.42$, $p < .001$, $f =$ 1.14, but not those in the SAI condition, $F(1, 124) = 2.56$, $p = .11$. To further explore the effect of Time on those in the WFR condition, we broke the interaction down into the amount of detail reported in each account. Statistical significance of a simple main effect was accepted at a Bonferroni-adjusted alpha level of .025, in order to not capitalise on probabilities (Cramer et al., 2016). For those in the WFR condition, there were significantly fewer details reported in the initial account ($M = 30.42$, $SD = 8.84$, 95% CI [27.33, 33.52]) than in the subsequent interview ($M = 43.91$, $SD = 9.20$, 95% CI [40.69, 47.13]). There was no significant difference between the amount of details reported in the initial account ($M =$

45.42, *SD* = 13.52, 95% CI [42.33, 48.52]) and the subsequent interview (*M* = 47.13, *SD* = 13.34, 95% CI [43.91, 50.34]) for those in the SAI condition. The interaction effect of Time and Veracity was not significant, $F(1, 124) = .02$, $p = .89$, nor was the interaction of Time, Veracity and Initial account, $F(1, 124) = .06$, $p = .80$.

Omissions. To explore the amount of detail that liars and truth tellers omitted from their initial accounts, we conducted a 2 (Veracity: truth teller vs liar) x 2 (Initial account: SAI vs WFR) ANOVA on the number of omitted details. There was a main effect of Initial account, $F(1, 123) = 142.77$, $p < .001$, $f = 1.06$, with those completing the SAI in their initial account omitting more details in their subsequent interview (*M* = 15.05, *SD* = 5.30, 95% CI [13.76, 16.36]) than those who initially completed a WFR (*M* = 5.33, *SD* = 3.83, 95% CI [4.39, 6.27]). There was also a main effect for Veracity, $F(1, 123) = 4.24$, $p = .042$, $f = .18$, with truth tellers omitting more details (*M* = 11.02, *SD* = 6.88, 95% CI [9.33, 12.71]) than liars (*M* = 9.27, *SD* = 6.46, 95% CI [7.69, 10.85]). No significant interaction was found, $F(1, 124) = .16$, $p = .69$.

Details provided in deceptive accounts. To compare the types of details that liars provided in their accounts, we conducted a series of independent samples *t*-tests. Statistical significance of the six *t*-tests was accepted at a Bonferroni-adjusted alpha level of .008, in order to not capitalise on probabilities. In the initial account, liars who completed the SAI provided more truthful details (*M* = 23.00, *SD* = 10.19, 95% CI [19.26, 26.74]) than liars who completed the WFR (*M* = 15.94, *SD* = 9.20, 95% CI [12.62, 19.26]), $t(61) = 2.89$, $p = .005$, $d = .73$. Liars completing the SAI provided more distorted detail in the initial account (*M* = 11.68, *SD* = 5.86, 95% CI [9.53, 13.83]) than liars completing the WFR (*M* = 5.22, *SD* = 3.80, 95% CI [3.85, 6.59]), $t(61) = 5.20$, $p < .001$, $d = 1.31$. There was no significant difference in the initial account between the number of fabrications reported by liars who completed the WFR (*M* = 4.69, *SD* = 4.22, 95% CI [3.23, 6.15]) and liars who completed the SAI (*M* = 7.26, *SD* = 8.00, 95% CI [4.49, 10.03]), $t(61) = 1.60$, $p = .11$. In the subsequent

interview, we found no significant difference between the amount of distorted detail reported by liars who had previously completed the SAI ($M = 9.84$, $SD = 5.09$, 95% CI [7.97, 11.71]) and liars who had previously completed the WFR ($M = 7.38$, $SD = 4.62$, 95% CI [5.71, 9.05]), $t(61) = 2.01$, $p = .048$. We also found no significant difference between liars who completed the WFR ($M = 7.69$, $SD = 8.40$, 95% CI [4.78, 10.60]) or SAI ($M = 7.90$, $SD = 8.35$, 95% CI [5.01, 10.79]) for the amount of fabricated details provided in the subsequent interview, $t(61) = 0.10$, $p = .92$. There was also no significant difference for the amount of truthful information provided by liars who completed the SAI ($M = 26.48$, $SD = 8.67$, 95% CI [23.47, 29.48]) and the WFR ($M = 24.53$, $SD = 9.22$, 95% CI [21.34, 27.72]) in the subsequent interview, $t(61) = .87$, $p = .39$.

Discussion

Replicating previous findings in deception literature (DePaulo et al., 2003), truth tellers provided significantly more details than liars in their initial accounts and subsequent interviews. Truth tellers also repeated more and omitted more details in the subsequent interview than liars did. As truth tellers reported a greater amount of detail in the initial account than liars, this provided a greater opportunity for more details to be repeated and omitted.

Contrary to the repeat vs reconstruct hypothesis (Granhag & Strömwall, 1999), no difference emerged between truth tellers and liars in the number of reminiscent details introduced in the subsequent interview. This may be due to the extensive nature of the subsequent interview, which may have exhausted the memory of all interviewees. The overall pattern, truth tellers repeated and omitted more details than liars did but no difference in reminiscences, replicates Granhag and Strömwall (2002).

Consistent with previous research (Gabbert et al., 2009; Hope et al., 2014), participants who completed the SAI reported more details in their initial account than participants who completed the WFR (supporting Hypothesis 1). However, unlike Hope et al. (2014), no difference was found in the amount of detail reported in the subsequent interview by participants who had previously completed the SAI compared to those who had completed the WFR. In the subsequent interview, Hope et al. (2014) used the Cognitive Interview (CI) technique (which the SAI is based upon) resulting in overlap between the two interviews. This overlap may have allowed SAI participants to use similar processing across both retrieval tasks. As we did not use CI in the current experiment, it is possible that the questions in the subsequent interview did not facilitate SAI participants' retrieval in the same way.

Consistent with Gilbert and Fisher (2006), all participants provided some reminiscent details during their subsequent interviews, and in both the SAI and WFR conditions,

significantly more detail was provided in the subsequent interviews than in the initial accounts. These findings may be due to the change in modality from a written to a verbal account, as people tend to report more information when they speak than when they write (Elntib, Wagstaff, & Wheatcroft, 2015; Sauerland & Sporer, 2011).

It was found that participants who initially completed a SAI repeated more details and omitted more details in their subsequent interviews than those completing the WFR. This is likely due to the greater amount of detail reported in the initial accounts prompted by the SAI, which gave participants more opportunity to repeat and omit details in the subsequent interview.

We did not find any difference in accuracy rates during the initial account or subsequent interviews for truth tellers completing the SAI compared to those completing a WFR, and therefore no support was found for Hypothesis 2. This could be due to participants being informed that they needed to provide a truthful report about the event before witnessing it. They may have thus paid close attention to the event.

When examining the types of detail constituting the initial deceptive accounts, it was found that liars completing the SAI provided more truthful details and more distorted details than liars completing the WFR. The current study required liars to embed their deception within truthful peripheral detail, as oppose to fabricate an entire scenario. This reflects real life better, where deceptive individuals are likely to embed their fabrications within truthful details (Leins, Fisher & Ross, 2013; Vrij, 2008).

In real life interviews, interviewers would have access to the initial account to assist in developing an interview plan. Since SAI has been developed for, and is recommended for use in, incidents involving multiple witnesses, statements can be compared with other witnesses' accounts (or physical evidence if available). Such comparisons would give investigators an idea which details reported in the SAI are truthful and which are