Amplifying deceivers' flawed metacognition:

Encouraging disclosures after delays with a Model Statement.

Adam Charles Harvey*

Department of Psychology, University of Portsmouth

Aldert Vrij

Department of Psychology, University of Portsmouth

Sharon Leal

Department of Psychology, University of Portsmouth

Lorraine Hope

Department of Psychology, University of Portsmouth

Samantha Mann

Department of Psychology, University of Portsmouth

*Correspondence concerning this article should be addressed to Adam C. Harvey, University of Portsmouth, Department of Psychology, King Henry Building, King Henry Street, Portsmouth PO1 2DY, Portsmouth, United Kingdom. Email: adam.harvey@port.ac.uk

This work is funded by the High-Value Detainee Interrogation Group, DJF-15-1200- V-0010196 awarded to the University of Portsmouth (UK). Any opinions, findings, conclusions, or recommendations expressed in this article are those of the authors and do not necessarily reflect the views of the U.S. Government.

Abstract (136 words)

Truth tellers provide less detail in delayed than in immediate interviews (likely due to forgetting), whereas liars provide similar amounts of detail in immediate and delayed interviews (displaying a metacognitive stability bias effect). We examined whether liar's flawed metacognition after delays could be exploited by encouraging interviewees to provide more detail via a Model Statement. Truthful and deceptive participants were interviewed immediately (n = 78) or after a three-week delay (n =78). Half the participants in each condition listened to a Model Statement before questioning. In the Immediate condition, truth tellers provided more details than liars. This pattern was unaffected by the Model Statement. In the Delayed condition, truth tellers and liars provided a similar amount of detail in the Model Statement-absent condition, whereas in the Model Statement-present condition, liars provided more details than truth tellers.

Key Words (6): Deception, investigative interviewing, suspects, metacognition, forgetting, delay

Word counts: **1083** (introduction); **1959** (discussion); **[7610]** (article total, excluding references and supplementary materials).

General Audience Summary/ public significance statement (137 words)**:**

Truth tellers provide less detail about an event or incident when interviewed after a delay compared to when interviewed immediately. This reduction in the reporting of detail is likely due to forgetting. In contrast, liars provide similar amounts of detail in immediate and delayed interviews, and so display flawed *metacognition* in that they fail to consider patterns of forgetting. We sought to investigate if liar's failure to simulate forgetting could be further amplified by encouraging interviewees to provide more detail via a Model Statement. In the Immediate interview condition, truth-tellers provided more details than liars, and this pattern was unaffected by the Model Statement. In the Delayed interview condition, truth tellers and liars reported a similar amount of detail in the Model Statement-absent condition, whereas in the Model Statement-present condition liars reported more details than truth tellers.

Amplifying deceivers' flawed metacognition:

Encouraging disclosures after delays with a Model Statement.

Whilst good memory is acknowledged as fundamental to successful deception (Sporer & Schwandt, 2006; Vrij & Granhag, 2014), it is less commonly observed that deceivers require astute metacognitive insights to lie effectively (Vrij et al., 2009). For example, when lying about events in the past, liars must accurately estimate the appropriate extent of forgetting, and this process requires metacognitive insight generally (i.e. cognition regarding mental processes; Flavell, 1979) and, more specifically, insight concerning metamemory (i.e. beliefs about memory performance; Kornell, Rhodes, Castel & Tauber, 2011).

Truth tellers display forgetting and report less information in delayed compared to immediate interviewing conditions (e.g. Ebbinghaus, 1885; Murre & Dros, 2015; Wixted & Carpenter, 2007). Liars show no equivalent forgetting after delays (e.g. Harvey, Vrij, Hope, Leal & Mann, 2017a; Harvey, Vrij, Leal, Hope & Mann, 2017b) but rather display a metacognitive stability bias effect: a failure to accurately take account of well-established patterns of forgetting over delay (see Kornell & Bjork, 2009; Koriat, Bjork, Sheffer & Bar, 2004). Whereas in immediate interviewing paradigms truth tellers typically provide more detailed statements than liars (Amado, Arce, Fariña, & Vilarino, 2016; Oberlader et al. 2016), detail is less diagnostic in delayed interviewing paradigms (Harvey et al., 2017a, b). The current study examines whether liar's flawed metacognition (underestimating forgetting) can be exploited by encouraging interviewees to disclose more detail after delays. Specifically, we investigate the hypothesis that liars will erroneously report *more* detail than truth tellers after a delay when provided with a Model Statement.

Eliciting cues in delayed interviewing conditions

As liars overestimate typical memory performance after delays, it may be possible to exploit this metacognitive error in delayed interviews by encouraging interviewees to disclose more information. Thus, techniques designed to promote the disclosure of information (e.g. the Model Statement, Leal et al., 2015; Vrij, Leal & Fisher, 2018) may be beneficial for lie detection in delayed interviewing settings.

The Model Statement is a technique developed to elicit more detailed statements from interviewees (Leal et al., 2015). The technique involves providing interviewees with a detailed statement, on an unrelated topic, that indicates the expected level of detail the interviewee should provide (Harvey, Vrij, Leal, Lafferty & Nahari, 2017c). Rather than providing retrieval enhancement, the Model Statement directs attention to the social norm 'be sufficiently detailed' (see Eriksson, Strimling & Coultas, 2015; Grice, 1975). In an interview setting where objective information about how much detail is required is not available, interviewees can use the Model Statement as a point of reference, and if the Model Statement is detailed, interviewees should attempt to be provide a similar level of detail.

When recall immediately follows an event, the use of a Model Statement results in the elicitation of more detailed statements from both adult truth tellers and liars, compared to a Model Statement-absent condition (See Vrij, Leal, & Fisher, 2018 for an overview of Model Statement deception research). Research indicates Model Statements affect honest and deceptive interviewees similarly in immediate interviewing contexts, and do not increase the diagnostic utility of 'richness of detail' as a dependent measure (Vrij et al., 2018).

Truth tellers' statements are constrained by their ability to successfully retrieve of information. Thus, truth tellers can *only* disclose additional information if they can successfully retrieve detail from memory. Liars' statements, however, are constrained by the extent to which they experience cognitive load as a result of lying (Vrij, Mann, Fisher, Leal,

Milne & Bull, 2008; Vrij, Granhag, Mann & Leal, 2011) and the limits of their imagination (e.g. Vrij, 2015; see Johnson & Raye, 1981; Johnson, Foley, Suengas, & Raye, 1988). In delayed interviewing scenarios, memory decay may constrain truth tellers' typical forthcoming verbal disclosure strategies (Harvey et al., 2017b). Truth tellers unable to retrieve detailed information may be less able to disclose additional detail in response to the Model Statement. In contrast, liars' reporting is less constrained by the extent of successful genuine retrieval, as their statements contain fabricated information. If encouraged to disclose more information following delayed interviewing, liars may do so for two reasons. First, liars may be motivated to appear cooperative by providing a detailed statement in response to the Model Statement (see Nahari, Vrij et al., 2012; Masip & Herrero, 2013). This strategy makes sense as more detailed statements are perceived as being more credible statements (Bell & Loftus, 1989; Johnson, Foley, Suengas, & Raye, 1988). Second, liars may overestimate genuine recollection following a delay. Such an error is consistent with both the stability bias effect (e.g. Kornell & Bjork, 2009; Koriat, Bjork, Sheffer & Bar, 2004) and the finding individuals typically hold false beliefs regarding the permanence of information in memory (Legaut & Laurence, 2007; Ost, et al., 2015; Simons & Chabris, 2011). Thus, in delayed interviewing, a liar's metacognitive error (failing to take into account the effects of forgetting) may be further magnified after hearing a Model Statement, encouraging them to provide more detail.

When rememberers cannot simultaneously be accurate *and* informative, accuracy is sacrificed to preserve informativeness (Ackerman & Goldsmith, 2008). According to metacognitive regulation models, without enhancing memory retrieval, providing more informative and detailed statements can *only* be achieved by increasing the likelihood of reporting incorrect information (Koriat & Goldsmith, 1996). For example, a witness

statement honestly describing a suspect as 'between 160 and 180 cm' is likely to be accurate, but at the cost of informativeness (the response is broad). In contrast, an estimate of '168cm' is more informative (the response is precise) but less likely to be accurate. Truth tellers can be expected to monitor their memories and only volunteer details they are confident are correct (Ackerman et al., 2008; Koriat et al., 1996). Plausibly, truth tellers may maintain a stable accuracy rate across time (e.g. Sauer & Hope, 2016). In contrast, liars can disclose fabricated (i.e. inaccurate) details as they require, irrespective of if they are interviewed immediately or after a delay.

Based on these theoretical considerations, we predict truth tellers in the Immediate interviewing condition will report more detailed statements than truth tellers in the Delayed interviewing condition, whereas there will be no difference in the amount of information reported by liars irrespective of whether the interview takes place immediately or after a delay (Hypothesis 1). Interviewees in the Model Statement-present condition will report statements containing more overall detail than interviewees in the Model Statement-absent condition who are not given a Model Statement before providing their statement (Hypothesis 2). Finally, in the delayed interviewing condition, liars will report more details that truth-tellers in the Model Statement-present condition, but liars and truth-tellers will report a similar number of details in the Model Statement-absent condition (Hypothesis 3).

Method

Design

A 2 (Veracity: Truth-teller vs. Liar) x 2 (Interview Time: Immediate vs. Delayed) x 2 (Reporting Instruction: Model Statement-present vs Model Statement-absent) between-subject experimental design was used.

The study used two target events: An immersive social interaction event and a recorded

'intelligence' stimulus. The social interaction event was rich in spatial, temporal and perceptual information and therefore suitable to code for experiential detail. The intelligence video was suitable for coding accuracy rates for reported detail. We used Reality Monitoring (RM) criteria for our detail coding. This measure has been used in previous research (e.g. Harvey et al., 2017a) and provides a well-defined and reliable measure of statement detail (see Vrij, 2008; 2015 for specific RM criteria). The dependent variables for the social interaction event were i) total number of details reported (spatial, temporal and perceptual RM details combined) in the initial statement and ii) the number of new details reported (unique RM details reported only in the final statement). The dependent variables for the intelligence video event were iii) the total accuracy rate (the proportion of the total number of reported RM details that were correct) and iv) the accuracy rate for new details (the proportion of new RM details reported in the final statement that were correct).

Participants

A power analysis using G*Power (Faul, Erdfelder, Lang & Buchner, 2007), assuming a medium effect size of f=0.25 ($\alpha = 0.05$) for six groups, indicated a sample size of 128 would be sufficient for an acceptable power of 0.80 (Cohen 1992). A total of 156 volunteers, comprising of 110 females and 46 males, aged between 18 and 69 years ($M = 24.49$ years, $SD = 9.11$, 95% CI [23.16, 26.06]), from the University's undergraduate (n=129) postgraduate (n=16) and staff communities (n=11) participated in the study in exchange for a small *honorarium* (£10 in the Immediate condition, £15 for the Delayed condition) and a chance to win in a raffle one of three additional cash prizes (£50, £100 or £150).

Procedure

The procedure was adapted from previous research (Harvey et al., 2017b). Individuals were recruited via adverts on the University's online participant recruitment platform. The

advert solicited individuals to participate in a study on deception in an intelligence setting.

All participants were randomly allocated to either the truth teller (n = 78) or liar (n = 78) Veracity conditions. Half of the participants per group where then randomly allocated to either the Delayed (n = 78) or the Immediate (n = 78) condition. Participants in the Delayed condition (both liars and truth tellers) were informed at the outset of the experiment that they would need to return in three-weeks to complete the study. In contrast, participants in the Immediate condition (both liars and truth tellers) were informed at the outset of the experiment that they would not need to return in three-weeks to complete the study. Half of the participants in each of the resulting four groups were then randomly allocated into either the Model Statement-present (n = 80) or Model Statement-absent (n = 76) reporting conditions.

All participants were told that the experiment involved assuming the role of an intelligence operative with access to a 'classified recording' of an intelligence briefing about a group of operatives who are planning to plant a surveillance device. (see Appendix 3 for participant instructions). This recorded film has been used in previous research (Shaw et.al, 2013; Ewens et al., 2015; Harvey et al., 2017a). All participants were told they should try and remember as many details about the intelligence recording as possible. Additionally, it was explained i) that note taking was prohibited and ii) that the intelligence recording could only be viewed once. All participants were told they would be interviewed later about the recording.

Truth tellers (n = 78) were told that for the experiment they are in the 'Blue' team and will be interviewed by a member of their own team. As such, they should be totally truthful to the interviewer and provide them with as much information as they can recall. Additionally, truth tellers were informed that there are also participants on the 'Red' team

taking part in the experiment at the same time, and thus they should be mindful as to what members of other teams may be doing. They were instructed that if they encountered anyone using the code words 'Rocket Science' this meant that those people were also on their 'Blue' team. They were also informed that if they did not hear those words then they could assume that the other participants were members of the opposing Red team and the truth tellers should pay attention to anything they do. It was explained that such information would be useful to the Blue team later in the experiment.

Liars (n = 78) were told that for the experiment they were on the 'Red' team and would be interviewed by a member of the opposing 'Blue' team and as such their task was to mislead the interviewer about certain details of the intelligence recording, including (i) what the surveillance device looked like, (ii) its functions and, also (iii) the location that was chosen to plant the device. Liars were told that the interviewer knew that the device would be placed somewhere, but did not know where. They were instructed that they should not reveal the location that was selected to hide the surveillance device and their objective was to mislead the investigator by using the third location mentioned in the intelligence recording as the location that was selected to plant the device. Liars were informed the interviewer knew something about the device but did not have all the details, and that it was not clear exactly what the interviewer knew. Because of this, liars were told to provide some truthful and some false information about the surveillance device, as this would help them appear cooperative without having to tell the interviewer everything. Finally, liars were informed they should be mindful as to what other Red team members may be doing in the experiment. Liars were told that if they encountered anyone during the study who used the code word 'Thermodynamics', those individuals were also on the Red team. Critically, the interviewer of the opposing 'Blue' team would probably be aware that members of the 'Red' team were taking part so

they should not deny seeing them. However, participants were also instructed that they should protect these individuals' identities by not telling the truth about what Red team members looked like and what they said, if asked by the interviewer.

All participants were told that if the interviewer judged their report as credible, they would be entered into a prize draw to win up to £150 in prize money in addition to receiving their *honorarium* (£10 in the Immediate condition, £15 in the Delayed condition). However, if participants did not appear credible (i.e. if the interviewer judged their reported as not truthful), they would not be included in the prize draw and instead be required to write a time-consuming statement about what happened during the study. Participants were instructed not to discuss the study with others and then taken to the waiting room and told to wait until the experimenter collected them to view the intelligence video.

Target Events

Our study used two target events: A recorded 'intelligence' event and an immersive social interaction event. All participants experienced, and were interviewed about, both events.

While waiting to watch the intelligence recording, all participants witnessed the staged social interaction event. This event consisted of a conversation, followed by a document exchange, between two confederates (unknown to the participants). Upon entering the waiting room, the participant was instructed to take a seat (the seat location was identical for all participants) and wait to be collected by the experimenter. One confederate (A) was already seated in the waiting room (again, this confederate location was identical for all participants). After 30 seconds, a second confederate (B) entered the waiting room and walked past the participant to sit next to the first confederate. Both confederates then engaged in the scripted exchange (which included the code word 'thermodynamics'), before a third

confederate entered the waiting room with an inquiry before leaving. The exchange then continued between confederates A and B before the experimenter returns and collects the participant (for a full description of the exchange, see Appendix 1). To check the standardisation of the scripted protocol, the duration of the staged social interaction was recorded for each participant, (M = 104.65 seconds; range: 80- 130; SD = 12.67, 95% CI [102.60, 106.70]). To examine the relationship between duration and the experimental conditions, a 2 (Veracity: Truth-teller vs. Liar) x 2 (Reporting Instruction: Model Statement-present vs. Model Statement-absent) x 2 (Interview Time: Immediate vs. Delayed) Analysis of Variance (ANOVA) was conducted. This analysis revealed no significant main effects and no significant interaction effects for duration, all F's < 2.32, all p's >.130.

Immediately after completion of the social interaction, the experimenter entered the room to collect the participant and escort him/her to watch the intelligence recording in the laboratory. Once seated, the participants viewed the intelligence recording, which comprised a filmed event lasting 6 minutes and 29 seconds. After completion, participants in the Delayed condition where told that they had completed the first phase of the study. On their return after three weeks, participants in the delay condition progressed to the second phase of the study. Participants in the Immediate condition progressed immediately into the second phase of the study.

The experimenter began the second phase of the experiment by informing participants that they would be questioned about both the intelligence recording and the social interaction in the waiting room. Experimental instructions (from Phase 1) were repeated and summarised for all participants. Participants were offered as much time as they required prior to the interview to prepare themselves. After indicating they were ready for the interview, all participants completed the pre-interview questionnaire. The participants were asked for their

demographic information (age, gender, occupation) and to rate their preparation for the interview (on 7-point Likert scales, ranging from 1 (very poor) to 7 (very good); 1 (pointless) to 7 (useful); 1 (insufficient) to 7 (sufficient); and 1 (incomplete) to 7 (thorough). Upon completion, participants were taken to be interviewed.

The interview

The same interviewer was used throughout the experiment and all interviews were audio and video recorded. The interviewer was blind to the experimental condition of the interviewees, and had not viewed the intelligence recording.

All interviews (both the Model Statement-present and Model Statement-absent conditions) began identically. Our interview protocol consisted of eight questions: the first set (questions 1-5) concerned the waiting room interaction and the second set (questions 6-8) concerned the intelligence video (see Appendix 2 for the full questioning protocol). After the participant answered the last question, the interviewer instructed the participant to wait in the interview room and exited for two minutes.

Upon the interviewers return after two minutes, all participants (n = 156) were told that the interviewer needed to clarify some information and so would ask the questions one final time. Participants in the Model Statement-absent condition where then questioned for second (final) time. Participants in the Model Statement-present condition (n=78) were given an additional instruction. The experimented informed them that it can be difficult for interviewees to gauge the appropriate level of detail to report in interview settings. Therefore, they should listen carefully to an audio recorded detailed short statement providing an indication of the appropriate level of detail to provide (i.e. the 'Model Statement'). We used the same Model Statement as previous research (Leal et al., 2015). The Model Statement used was 346 words in length and no specific detail category was emphasised (c.f. Harvey et

al., 2017c). That is, the Model Statement contained a mix of spatial, temporal and perceptual detail. The audio Model Statement lasted 1 minute and 34 seconds. A brief extract from of the Model Statement is included below. For clarity, we have indicated the spatial details (superscript ^{S}), the temporal details (superscript ^{T}) and the perceptual details (superscript ^{P}) but this information was not provided to the interviewees.

'...At this point^{T} I^{P} got out^{S} my^{P} phone^{P} to do a bit of filming^{P} as it's not every day^{T} that you get to be on^{S} the grid^{P} at a formula 2 race^{P}. I was trying to capture the feel of it in the video^{P}. Just then^{T}, I saw^{P} Tom^{P} pull up^{P} in his^{P} formula 2 car^{P}, he was in full^{P} risk gear^{P} with his helmet^{P} on^{S}...'

After the Model Statement was administered, participants in the Model Statement condition were questioned for a second (final) time. The second set of questions were identical in content and sequence for both participants in the Model Statement-present condition and Model Statement-absent condition, and used the same eight questions previously asked (see Appendix 1).

After the final interview, all participants returned to the laboratory and were given a post interview questionnaire. This asked participants to report their motivation for performing well during the interview (on a 7-point Likert Scale, ranging from 1 extremely unmotivated to 7 extremely motivated), to estimate the likelihood (on an 11-point Likert scale, ranging from 0% to 100% likely) of (i) receiving the monetary reward and (ii) having to write the statement, and to report percentage of truthful information they disclosed in the interview (also on an 11-point Likert Scale, ranging from 0% to 100%)). Upon completion participants where thanked, debriefed and compensated for their time.

Coding

All audiotapes were transcribed and the verbal coding was conducted using these

transcripts. All statements were rated by one coder (blind to the experimental conditions) who scored the occurrence of perceptual detail (information about what was seen, heard, felt and smelt during the described activities, e.g. 'There was man[P] in a jacket[P] there'), spatial detail (information about locations or the arrangement of persons and/or objects, e.g. 'the sofa in the far[S] left[S] corner[S] of the room under[S] the window') and temporal detail (information about when the event happened and explicit descriptions of the sequence of various events, 'After[T] no one replied, she then[T] left'). Spatial, temporal and perceptual detail that did not relate to the coding script was considered irrelevant. Such irrelevant detail did not occur for the video description and rarely occurred for the social interaction. Note that the three sub-categories of detail were introduced to facilitate (inter-rater) reliability coding. As no hypothesis was formulated regarding any individual detail category, we only report total detail scores (summed total of spatial, temporal and perceptual detail).

All statements were then coded for new information (i.e. new details reported in the second but *not* the first statement) as specified by Fisher et al., (2013) and Vredeveldt et al., (2014). Finally, the statements pertaining to the intelligence recording (i.e. responses to Questions 6-8) were subject to an additional level of coding. Each detail was further classified as either accurate or inaccurate. To code for accuracy of the information provided, answers where compared to the intelligence video. Reported detail that matched the content of the recording was scored as accurate, whereas detail that did not match the script was scored as inaccurate. Accuracy rates were calculated as the number of accurate details divided by the number of total details reported.

Inter-rater reliability.

A second coder (also blind to the veracity of the statements) coded a random selection of 32 statements (20%) for all the dependent measures. Inter-rater reliabilities between the

two coders for the occurrence frequency of perceptual, spatial and temporal detail, frequency of commissions, as well as for detail accuracy, were measured via intra-class correlation coefficients (ICC). The ICC was high and therefore satisfactory for total details [ICC = .91], new details (commissions) [ICC] = .96, and for correct [ICC] = .94 and incorrect details [ICC] = .92.

Results

Veracity manipulation check.

Two 2 (Veracity: Truth-teller vs. Liar) x 2 (Interview Time: Immediate vs. Delayed) x (Reporting Instruction: Model Statement-present vs. Model Statement-absent) ANOVAs were conducted with the estimated likelihood of (i) receiving the monetary reward and (ii) having to write a statement as the dependent variables. These analyses revealed significant main effects for Veracity regarding both the monetary incentive, $F(1, 148) = 24.57$, $p < 0.001$, $\eta_p^2 = 0.14$, and writing a statement, $F(1, 148) = 55.02$, $p < 0.001$, $\eta_p^2 = 0.19$. Truth tellers ($M = 5.59$, $SD = 1.50$, *95% CI* [5.25, 5.91]) thought it more likely they would receive the reward than liars ($M = 4.47$, $SD = 1.25$, *95% CI* [4.19, 4.74]), Cohen's $d = .81$, 95% CI [.47, 1.13], whereas liars ($M = 3.58$, $SD = 1.55$, 95% CI [3.23, 3.90]) thought it more likely they would have to write a statement than truth tellers ($M = 2.38$, $SD = .83$, *95% CI* [2.19, 2.57]), $d = .97$, 95% CI [.62, 1.28]. No other significant effects emerged for either analysis, all F's < .70, all p's > .41.

A 2 (Veracity: Truth-teller vs. Liar) X 2 (Interview Time: Immediate vs. Delayed) X 2 (Reporting Instruction: Model Statement-present vs. Model Statement-absent) ANOVA was conducted with the self-reported percentage of truthful information disclosed as the dependent variable. The analysis showed a main effect for Veracity, $F(1, 148) = 1109.80$, $p < 0.001$, $\eta_p^2 = .88$. Liars ($M = 31.04\%$, $SD = 11.65$, *95% CI* [28.50, 33.68] reported providing

significantly less truthful information during the interview than truth tellers (M = 94.62%, SD = 12.03, *95% CI* [91.83, 96.94]), d= 5.37, 95% CI [4.62, 5.95]**)**. No other significant effects emerged, all $F's < 2.49$, all $p's > .12$.

Hypothesis testing

Compared to the Null Hypothesis Significance Testing (NHST) approach (e.g. Cohen, 1992; Cummings, 2014), Bayesian analysis confers advantages to researchers (Etz, Gronau, Dablander, Edelsbrunner & Baribault, 2017; Etz & Vandekerckhove, 2017), including a means of quantifying the extent to which the data support hypotheses (Wagenmakers et al., 2016; 2017a). In order to assess the strength of evidence for our experimental hypothesis (following Otgaar, Romeo, Ramakers & Howe, 2017; Otgaar, Howe, Smeets & Wang, 2016), and in addition to significance testing, we calculated a Bayes Factor (BF) score (e.g. Wagenmakers et al., 2016) using a default Bayesian *t* test and open-source JASP software (https://jasp-stats.org; see Wagenmakers et al., 2017b). Note that the default Bayesian *t* test used here is based on the assumptions for when researchers lack knowledge about the expected effect size under investigation (e.g. Wagenmakers et al., 2011). Table 1 shows the interpretation of Bayes Factor derived from Jeffreys (1961) and adapted by Wagenmakers et al. (2011).

[Please add Table 1 about here]

Social Interaction: Quality of reported detail

Since the Reporting Instruction (Model Statement-present or Model Statement-absent) was introduced after the initial statement, this factor is not introduced in the analyses regarding the initial statement. We have, however, included the number of details reported in

the initial statement as a covariate in the analyses of new details (because new details reported in the second statement relate necessarily to the amount of details reported in the initial statement). That is, the more details reported in the initial statement, the fewer new details are likely to be available for subsequent reporting.

A 2 (Veracity: Truth-teller vs. Liar) x 2 (Interview Time: Immediate versus Delayed) ANOVA with the total number of details reported during the Initial Statement as the dependent variable revealed a significant main effect of Veracity, $F(1,152) = 13.14$, $p < .001$, Cohen's $d = .56$, *95% CI* [.23, .87], such that more details were provided by truth tellers than by liars. The main effect for Interview Time was not significant, $F(1,152)= 1.31$, $p = .254$, $d = .17$, 95% CI [-.15, .48]. However, a significant Veracity x Interview Time interaction effect emerged, $F(1,152)= 14.81$, $p < .001$, $\eta_p^2 = .09$.

[Table 2 about here]

Regarding the Veracity x Interview Time interaction, Table 2 shows, liars in the Immediate interviewing condition and liars in the Delayed condition reported similar amounts of detail. To examine how strongly our data supported the hypothesis, we also calculated a Bayes Factor (BF) score (e.g. Wagenmakers et al., 2016). Bayesian analysis showed our data were more in support of the null hypothesis ($BF_{10} = 0.90$). In contrast, truth tellers in the Immediate condition reported significantly more details than truth tellers in the Delayed condition. Bayesian analysis showed our data were more in support of the alternative hypothesis ($BF_{10} = 109.00$). These results support Hypothesis 1.

A 2 (Veracity: Truth-teller vs. Liar) x 2 (Interview Time: Immediate versus Delayed) x 2 (Reporting Instruction: Model Statement-present versus Model Statement-absent) ANCOVA with the total amount of new detail reported in the final statement as the

dependent variable, and the total amount of detail reported during the initial statement as covariate revealed a significant main effect emerged for Veracity, $F(1,147) = 7.43$, $p = .007$, $d = .22$, 95% CI [-.10, .53], such that more new details were provided by liars than by truth tellers. Also a significant main effect for Reporting Instruction emerged, $F(1,147) = 64.21$, $p < .001$, $d = 1.19$, 95% CI [.83, 1.51], such that more new details were provided in the Model Statement-present condition than in the Model Statement-absent condition. The Veracity x Reporting Instruction interaction was significant, $F(1,147) = 5.86$, $p = .018$, $\eta_p^2 = .04$. Truth tellers and liars (see Table 3 for *M* and *SD*) in the Model Statement-absent condition reported similar amounts of new details, $t(74)=1.63$, $p=.108$, $d= .37$, 95% CI [-.09, .82]. Bayesian analysis showed our data were more in support of the null hypothesis ($BF_{10} = .74$). In contrast, liars in the Model Statement-present condition reported significantly more new details than truth tellers in the Model Statement-present condition, $t(58.01)=2.47$, $p=.017$, $d= .55$, 95% CI [.10, .99]. Bayesian analysis showed our data were more in support of the alternative hypothesis ($BF_{10} = 3.09$). No other effects were significant, all F's < 2.65, all p's > .106.

[Table 3 about here]

The Veracity x Interview Time x Reporting Instruction three-way interaction statistics ($p = .325$) refers to any type of interaction. However, in Hypothesis 3 we predicted a specific type of interaction: that in the Delayed interviewing condition, liars will report more details that truth-tellers in the Model Statement-present condition, but liars and truth-tellers will report a similar number of details in the Model Statement-absent condition (Hypothesis 3). Therefore, a more informative test of Hypothesis 3 is to statistically test for significant differences between liars and truth tellers in the Delayed interviewing condition, for the

Model Statement-present and Model Statement-absent condition separately. This approach, introduced by Nahari and Ben-Shakhar (2011), has now been adapted by others (Nahari, 2017). Thus, we test Hypothesis 3 directly.

In the Delayed interviewing condition, truth tellers in the Model Statement-absent condition (*M*=34.42, *SD*=16.00, 95% CI [27.94, 42.05]) and liars in the Model Statement-absent condition (*M*=34.05, *SD*= 21.95, 95% CI [25.09, 44.79]), reported similar amounts of new detail, $t(36)=.06$, $p=.953$, $d= .02$, 95% CI [-.62, .65]. Bayesian analysis showed our data were more in support of the null hypothesis ($BF_{10} = 0.32$).

In contrast, in the Delayed interviewing condition, liars in the Model Statement-present condition (*M*= 79.75, *SD*= 51.37, 95% CI [56.74, 102.00]) reported significantly more new details compared to truth-tellers in the Model Statement-present condition (*M*= 55.75, *SD*= 20.43, 95% CI [46.65, 64.47]), $t(24.87)=1.94$, $p=.030$, $d= .61$, 95% CI [-.03, 1.24]. Bayesian analysis showed our data were more in support of the alternative hypothesis ($BF_{10} = 2.55$). These results support Hypothesis 3.

A 2 (Veracity: Truth-teller vs. Liar) x 2 (Interview Time: Immediate versus Delayed) x 2 (Reporting Instruction: Model Statement-present versus Model Statement-absent) ANOVA with total unique detail (detail reported in the initial statement plus new detail reported in the final statement) as the dependent variable revealed a significant main effect emerged for Reporting Instruction, $F(1,148) = 13.31$, $p <.001$, $d =.55$, *95% CI* [.22, .86], such that more details were provided in the Model Statement-present condition than in the Model Statement-absent condition. The Veracity x Reporting Instruction interaction was significant, $F(1,147) = 7.81$, $p = .006$, $\eta_p^2 = .05$. As Table 3 shows, truth tellers in the Model Statement-present condition and truth tellers in the Model Statement-absent condition reported similar amounts of unique details. As Table 3 also shows, liars in the Model Statement-present

condition reported more unique details than liars in the Model Statement-absent condition. The Veracity x Interview Time interaction was significant, $F(1,147) = 9.53$, $p = .002$, $\eta_p^2 = .06$. As Table 2 shows, truth tellers in the Immediate interviewing condition reported more unique details than truth tellers in the Delayed interviewing condition. Bayesian analysis showed our data were more in support of the alternative hypothesis ($BF_{10} = 293.60$). As Table 2 also shows, liars in the Immediate interviewing condition and liars in the Delayed interviewing condition reported similar amounts of unique details. Bayesian analysis showed our data were more in support of the null hypothesis ($BF_{10} = 0.31$). No other effects were significant, all F's < 3.67, all p's > .057. These findings support Hypothesis 2.

Intelligence Recording: Accuracy rates for reported information

We conducted analyses of detail quantity and quality (accuracy) for the intelligence video, but as the results were similar to those of the social interaction, we report only analyses of accuracy rates here. [These additional analyses are presented in the supplementary materials.]

A 2 (Veracity: Truth-teller vs. Liar) x 2 (Interview Time: Immediate versus Delayed) ANOVA with accuracy rate for detail reported during the Initial statement as dependent variable revealed a significant main effect for Veracity, $F(1,151) = 859.18$, $p < .001$, $d = 4.50$, *95% CI* [3.85, 5.02], such that the information provided by truth tellers was more accurate than that provided by liars. A significant main effect for Interview Time also emerged, $F(1,151) = 10.73$, $p = .001$, $d = .20$, 95% CI [-.11, .51], such that participants were more accurate in the Immediate condition than in the Delayed condition. The Veracity x Interview Time interaction effect was also significant, $F(1,151) = 4.45$, $p = .036$, $\eta_p^2 = .03$. As Table 2 shows, accounts provided by truth tellers in the Immediate condition had a higher accuracy rate than accounts provided by truth tellers in the Delayed condition. In contrast, the

accounts provided by liars in the Immediate condition and the accounts provided by liars in the Delayed condition had similar accuracy rates.

A 2 (Veracity: Truth-teller vs. Liar) x 2 (Interview Time: Immediate versus Delayed) x 2 (Reporting Instruction: Model Statement-present versus Model Statement-absent) ANOVA with accuracy rate for unique detail (accuracy rate for detail reported in the initial statement plus new details) as dependent variable revealed a significant main effect for Veracity, $F(1,148)= 13.81$, $p <.001$, $d = 4.34$, 95% CI [3.71, 4.85], such that truth tellers were more accurate than liars. A significant main effect also emerged for Interview Time, $F(1,148)= 11.41$, $p=.001$, $d= .21$, *95% CI* [-.10, .53], such that participants were more accurate in the Immediate condition than in the Delayed condition.

The Veracity x Interview Time interaction was significant, $F(1,148)=12.30$, $p = .001$, $\eta_p^2 = .08$. As Table 2 shows, accounts provided by Liars in the Immediate condition and accounts provided by liars in the Delayed condition had similar accuracy rates for unique detail. In contrast, accounts provided by truth tellers in the Immediate condition had a higher accuracy rate for unique detail than accounts provided by truth tellers in the Delayed condition. No other interaction effects were significant, all F's < 1.05, all p's > .308.

Discussion

The current study tested the hypothesis that liars' flawed metacognition can be amplified by a Model Statement encouraging interviewees to report more details. We found evidence supporting this hypothesis; after a delay, liars failed to accurately feign the forgetting observed for truth tellers, and when provided with a Model Statement, reported *more* detailed statements than truth tellers.

For the social interaction, a dissociation was observed for the effects of delay on truth tellers' and liars' statements. Liars in the Immediate and Delayed conditions reported similar

amounts of unique detail in their statements. This insensitivity to the effects of forgetting across time replicates previous research showing liars display a stability bias when lying after delays (e.g. Harvey et al., 2017a). This stability bias error is at odds with genuine memory performance (Murre et al., 2015; Wixted et al., 2007), and hence with the patterns of forgetting obtained for honest interviewees in our study. Truth tellers in the Immediate condition reported significantly more detail in their initial statement, and more unique detail overall, compared to truth tellers in the Delayed condition. Importantly, truth tellers in the Delayed condition who were provided a Model Statement reported similar quantities of detail compared to truth tellers in the Delayed condition who were not provided a Model Statement. As such, the Model Statement had no mnemonic effect. As what truthful interviewees can report depends on successful retrieval (Fisher & Geiselman, 2010; Vrij, Hope, Fisher, 2014), these findings underscore how the forthcoming verbal strategies typically displayed by honest interviewees (e.g. Hartwig, Granhag, & Strömwall, 2007; Hartwig, Granhag, Strömwall & Doering, 2010) are constrained by forgetting (Harvey et al., 2017a; 2017b). However, liars are less constrained by the limits of successful retrieval; instead, lies involve information generation (e.g. Vrij, 2008, 2015).

Consistent with the finding that truth tellers typically provide more detailed statements than liars (Amado et al., 2016; Masip et al., 2005; Oberlader et al. 2016), truth tellers in the Immediate condition reported initial statements containing more total detail compared to liars. However, truth tellers and liars in the Delayed condition reported initial statements containing similar total detail. This replicated previous research demonstrating that the utility of 'richness of detail' veracity cue decreases as delay increases (Harvey et al., 2017a; 2017b). For practitioners, the results provide a caveat to the typical finding that truth tellers provide more detailed statements than liars.

In terms of new detail reported after the initial statement; truth tellers and liars in the Model Statement-present condition reported more new details (displaying lower between-statement consistency; Vredeveldt, et al., 2014) than truth tellers and liars in the Model Statement-absent condition. Thus, truthful and deceptive suspects in our study responded to the Model Statement in a similar way: by disclosing new details. Previous strategy research indicates that liars prefer to keep their stories simple (Granhag et al., 2002; Leins, Fisher & Ross, 2012; Stromwall, Hartwig & Granhag, 2006), presumably to avoid reporting contradictions that would risk their credibility. Liars also avoid providing as many new, previously unreported details in their statements compared to truth tellers (Colwell, Hiscock-Anisman, Memon, Rachel, & Colwell, 2007; Colwell, Hiscock-Anisman & Fede, 2013; also see Hinds, Colwell, Hiscock-Anisman, Anisman & Montalvo, 2010). Thus, in some respects it is surprising liars add new details because doing so increases the opportunities for contradictions to emerge in their statements. Furthermore, liars are also theoretically motivated to maintain between-statement consistency by *repeating* what they have previously disclosed in subsequent statements (Granhag et al., 1999, 2001, 2003). However, it appears that the Model Statement led to truth tellers and liars prioritizing the provision of detailed statements (Vrij et al., 2018). For liars, the motivation 'to be detailed' appeared to override the motivation 'to be consistent' or 'keep the story simple'.

In this experiment, we used two different target event formats, however, the results for the social interaction and intelligence recording target events were similar. For both target events, in the Delayed Model Statement-absent condition, truth tellers and liars reported similar amounts of total detail in their statements. This finding is consistent with previous research (Harvey et al., 2017a; 2017b). Remarkably, in the Delayed condition using the Model Statement, liars reported *more* total detail compared to truth tellers**.** By reversing the

typical relationship between veracity and reported detail (truth tellers typically report more detailed statements compared to liars; e.g. Amado et al., 2016), the Model Statement appeared to amplify **liars'** overestimation of the appropriate amount of detail to disclose. Although liars display a stability bias in both the Model Statement-present and Model Statement-absent condition (in neither condition do liars successfully feign the natural curve of forgetting), the metacognitive error is most pronounced in the Model Statement condition. Only in the Delayed condition using the Model Statement did liars provide statements containing significantly more details compared to truth tellers. It is important to emphasis the direction of the difference reported in this study is a reverse to the orthodox finding in the literature, namely, that the statements reported by truth tellers are more detailed compared to the statements reported by liars (Amado et al., 2016; Oberlader et al. 2016). Applying the 'richness of detail heuristic' – the decision rule that truthful statements are richer in detail compared to fabricated statements (e.g. Nahari & Vrij, 2015) – conventionally would actually *impair* lie detection performance in the model-statement condition after delays. Thus, we believe caution is warranted when conventional criteria are used to assessed credibility in statements elicited after delays (Harvey et al., 2017a; 2017b).

Theoretically, three explanations may account for liars' stability bias. According to the dual view of metacognitive judgements derived from the Cue Utilization Approach (Koriat, 1997; Koriat et al., 2004; Kornell et al., 2009), metacognitive judgements may be based upon two sources of information: (i) subjective experience (experienced-based cues) and (ii) domain-specific knowledge (theory-based cues). Firstly, liars can utilize experience-based cues (e.g. the experience of retrieving their actual memory of the event they lie about) to inform their metacognitive judgements regarding how much detail to provide. Liars may revisit their memory during the 3-week delay, with this repeated retrieval inoculating against

memory decay. Hence, a testing-effect phenomenon (Roediger & Butler, 2011) may underpin the stability bias because liars use their own (strong) memory for the past event as a cue to base their performance. However, this rationale cannot account for stability bias when liars fabricate 'past events' that did not occur, i.e. have no corresponding memory (e.g. Harvey et al., 2017b, Experiment 1).

Secondly, liars may utilize theory-based cues to inform their judgements regarding how much detail to provide. Individuals may hold inaccurate beliefs about memory (Legaut et al., 2007; Ost, et al., 2015; Simons et al., 2011) and these erroneous beliefs motivate over-performance. Alternatively, individuals may hold accurate beliefs about memory performance, but fail to employ them (Koriat et al., 2004). Theoretically, for liars to take into account the effect of delay on memory retention, the concept of forgetting as declining memory performance over time must be activated. However, it appears such beliefs are organized hierarchically; only once the concept of forgetting is activated can liars employ their beliefs regarding the effects of delay on memory when estimating performance (Koriat et al., 2004). Apparently, the between-subject delay manipulation used in the current study (also see Harvey et al., 2017a; 2017b) was insufficient to activate this concept of forgetting. Manipulations that successfully activate the notion of forgetting may eliminate the stability bias effect amongst deceivers.

Thirdly, liars may not prioritize taking into account the effect of delay upon memory performance, despite being capable of making such predictions accurately. Hence, liars do not take the perspective of truth tellers (and do not consider how much they are likely to recall after delays) due to taking the mental perspective of the interviewer and considering what they will find convincing. Accordingly, liars maybe strategically unwilling to leave out too many details due to concerns about not projecting a sufficiently convincing impression to

the interviewer (e.g. Köhnken, 1989, 1996, 2004). This strategy would make good sense as more detailed statements are more likely to be perceived as credible (Bell et al., 1989; Johnson, 2006). Our current study is limited insofar it does not allow us to determine which explanation is correct. Further research is needed to identify the mechanism(s) underpinning the stability bias effect amongst deceivers.

Plausibly, informing interviewees about the stability bias effect will result in liars accurately taking the effect of delay into account, limiting the utility of questioning designed to exploit their flawed metacognition. However, even with information about the stability bias, accurately taking the effects of forgetting into account may prove difficult for liars. Liars would have to judge how much detail to withhold from interviewers to feign forgetting, and so may over- or under-estimate the true extent of genuine recollection. It is thus unclear if informing interviewees about the stability bias would eliminate, attenuate or even amplify its effect. This issue should be tackled in future research.

The current experiment was conducted in the relatively under studied intelligence setting (Vrij & Granhag, 2014), and it is possible different results will be found in other domains, such as insurance claims or police-suspect interviewing. Theoretically, the difficulty of liars' task may influence their ability to fabricate large quantities of detail after a delay, and thus mediate the stability bias effect observed amongst deceivers. Therefore, different findings may emerge across domains in which lying after delays is more or less cognitively demanding. Future research should explore this possibility.

The Model Statement made *both* truth tellers and liars provide more information. However, because cues to deception are more likely to emerge in longer statements compared to shorter statements (Leal et al., 2014; Mann et al., 2013), it may be the case that if the use of Model Statements encourages liars to report more detailed statements including more

fabrications, this may be an advantage at a later stage after the interview when the veracity of the reported details can be checked.

For the intelligence recording target event, statements provided by liars had similar accuracy rates for the initial statement and for unique details in both the Immediate and Delayed conditions. In contrast, the statements provided by truth tellers in the Immediate condition had higher accuracy rates (and higher accuracy rates for unique detail), than the statements provided to truth tellers in the Delayed condition. As liars could manipulate *and* outright fabricate the information they disclosed, deceivers had to remember relatively less information (compared to truth tellers) to maintain stable accuracy rates across time. The results concerning truth tellers are consistent with the finding that statement accuracy decreases as the delay between encoding and retrieval increases (Wixted & Ebbesen, 1991, 1997). However, the accuracy rate for unique details did not vary for truth tellers as a function of whether the Model Statement was present or absent, in either the Immediate or Delayed conditions. Apparently, truth tellers successfully regulated their memory to also take into account the *correctness* of their statements (Ackerman et al., 2008; Koriat et al., 1996; Sauer et al., 2016), withholding details they considered inaccurate, thus maintaining their statement's accuracy.

In conclusion, unlike truth tellers who display forgetting across time, liars provide similar amounts of detail in Immediate and Delayed interviewing conditions (displaying a metacognitive stability bias effect). Our study investigated the hypothesis that liars' flawed metacognition can be amplified by using a Model Statement encouraging interviewees to disclose more detail. Although liars never successfully feigned the natural curve of forgetting after a delay (and thus always displayed a stability bias), the metacognitive error was most pronounced in the Model Statement condition (in which liars provided more detailed

statements than truth tellers). Thus, liar's flawed metacognition may be vulnerable to exploitation via proactive interviewing.

References

Ackerman, R., & Goldsmith, M. (2008). Control over grain size in memory reporting--With and without satisficing knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 1224-1245. DOI:10.1037/a0012938

Amado, B. G., Arce, R., Farina, F., & Vilariño, M. (2016). Criteria-Based Content Analysis (CBCA) reality criteria in adults: A meta-analytic review. *International Journal of Clinical and Health Psychology*, *16*(2), 201-210. DOI: 10.1016/j.ijchp.2016.01.002

Bell, B. E., & Loftus, E. F. (1989). Trivial persuasion in the courtroom: The power of (a few) minor details. *Journal of Personality and Social Psychology*, *56*, 669-679. DOI: 10.1037//0022-3514.56.5.669

Bogaard, G., Meijer, E. H., & Vrij, A. (2014). Using an example statement increases information but does not increase accuracy of CBCA, RM, and SCAN. *Journal of Investigative Psychology and Offender Profiling*, *11*(2), 151-163. Doi:10.1002/jip.1409

Brackmann, N., Otgaar, H., Roos af Hjelmsäter, E., & Sauerland, M. (2017). Testing a new approach to improve recall in different ages: Providing witnesses with a Model Statement. *Translational Issues in Psychological Science*, *3*(2), 131-142, http://dx.doi.org/10.1037/tps0000116

British Psychological Society (2009). Code of Ethics and Conduct. Guidance by the Ethics Committee of the British Psychological Society. Leicester, UK: British Psychological Society.

British Psychological Society (2014). Code of Ethics and Conduct. Guidance by the Ethics Committee of the British Psychological Society. Leicester, UK: British Psychological Society.

Cohen, J. (1992). A power primer. *Psychological bulletin*, *112*(1), 155-159.

Colwell, K., Fede, J., & Hiscock-Anisman, C. (2013). Assessment Criteria Indicative of Deception: an example of the new paradigm of differential recall enhancement. In B.S. Cooper et al. (Eds.) *Applied Issues in Investigative Interviewing, Eyewitness Memory, and Credibility Assessments*. New York: Springer Science and Business Media. https://doi.org/10.1007/978-1-4614-5547-9_11

Colwell, K., Hiscock-Anisman, C. K., Memon, A., Taylor, L., & Prewett, J. (2007). Assessment Criteria Indicative of Deception (ACID): An integrated system of investigative interviewing and detecting deception. *Journal of Investigative Psychology and Offender Profiling*, *4*(3), 167-180, DOI: 10.1002/jip.73

Cumming, G. (2014). The new statistics: Why and how. *Psychological science*, *25*(1), 7-29, DOI: 10.1177/0956797613504966

Dienes, Z. (2008). *Understanding psychology as a science: An introduction to scientific and statistical inference*. Palgrave Macmillan.

Ebbinghaus, H. (1885). *Memory: A contribution to experimental psychology*. Leipzig: Duncker & Humblot.

Eriksson, K., Strimling, P., & Coultas, J. C. (2015). Bidirectional associations between descriptive and injunctive norms. *Organizational Behavior and Human Decision Processes*, *129*, 59-69. http://dx.doi.org/10.1016/j.obhdp.2014.09.011

Etz, A., & Vandekerckhove, J. (2017). Introduction to Bayesian inference for psychology. *Psychonomic bulletin & review*, *25*, 5-34, DOI 10.3758/s13423-017-1262-3

Etz, A., Gronau, Q. F., Dablander, F., Edelsbrunner, P. A., & Baribault, B. (2016). How to become a Bayesian in eight easy steps: An annotated reading list. *Psychonomic bulletin & review*, *25*, 1-16, https://doi.org/10.3758/s13423-017-1317-5

Ewens, S., Vrij, A., Leal, S., Mann, S., Jo, E., Shaboltas, A., Ivanova, M., Granskaya, J., & Houston, K. (2016). Using the Model Statement to elicit information and cues to deceit from native speakers, non-native speakers and those talking through an interpreter. *Applied Cognitive Psychology*.

Ewens, S., Vrij, A., Mann, S., & Leal, S. (2015). Using the Reverse Order Technique with Non-Native Speakers or Through an Interpreter. *Applied Cognitive Psychology*.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, *39*(2), 175-191.

Fisher, R. P., & Geiselman, R. E. (2010). The cognitive interview method of conducting police interviews: Eliciting extensive information and promoting therapeutic jurisprudence. *International journal of law and psychiatry*, *33*(5), 321-328, https://doi.org/10.1016/j.ijlp.2010.09.004

Fisher, R. P., Vrij, A., & Leins, D. A. (2013). Does testimonial inconsistency indicate memory inaccuracy and deception? Beliefs, empirical research, and theory. In *Applied issues in investigative interviewing, eyewitness memory, and credibility assessment* (pp. 173-189). Springer New York.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American psychologist*, *34*(10), 906-911. DOI: 10.1037/0003-066X.34.10.906

Granhag, P. A., & Strömwall, L. A. (1999). Repeated interrogations—Stretching the deception detection paradigm. *Expert Evidence, 7,* 163–174. doi:10.1023/a:1008993326434.

Granhag, P. A., & Strömwall, L. A. (2001). Deception detection based on repeated interrogations. *Legal and Criminological Psychology, 6,* 85–101. doi:10.1348/135532501168217.

Granhag, P. A., & Strömwall, L. A. (2002). Repeated interrogations: verbal and non-verbal cues to deception. *Applied Cognitive Psychology*, *16*(3), 243-257, https://doi.org/10.1002/acp.784

Granhag, P. A., Strömwall, L. A., & Jonsson, A.C. (2003). Partners in crime: How liars in collusion betray themselves. *Journal of Applied Social Psychology, 33,* 848–868. doi:10.1111/j.1559-1816.2003.tb01928.x.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics 3: Speech acts* (pp. 41–58). New York: Academic Press.

Hartwig, M., Granhag, P. A., & Strömwall, L. A. (2007). Guilty and innocent suspects' strategies during police interrogations. *Psychology, Crime & Law*, *13*(2), 213-227. doi:10.1080/10683160600750264

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

Harvey, A. C., Vrij, A., Hope, L., Leal, S., & Mann, S. (2017a). A Stability Bias Effect Among Deceivers. *Law and human behavior*. DOI:10.1037/lhb0000258.

Harvey, A. C., Vrij, A., Leal, S., Hope, L., & Mann, S. (2017b). Deception and Decay: Verbal Lie Detection as a Function of Delay and Encoding Quality. *Journal of Applied Research in Memory and Cognition*. https://doi.org/10.1016/j.jarmac.2017.04.002

Harvey, A. C., Vrij, A., Leal, S., Lafferty, M., & Nahari, G. (2017c). Insurance based lie detection: Enhancing the verifiability approach with a Model Statement component. *Acta psychologica*, *174*, 1-8. https://doi.org/10.1016/j.actpsy.2017.01.001

Hines, A., Colwell, K., Hiscock-Anisman, C., Garrett, E., Ansarra, R., Montalvo, L., 2010. Impression management strategies of deceivers and honest reporters in an investigative interview. European Journal of Psychology Applied to Legal Context, *2*, 73–90.

Jeffreys, H. (1961). Theory of probability (3rd ed.). New York: Oxford University Press.

Johnson, M. K. (2006). Memory and reality. *American Psychologist*, *61*(8), 760-771, DOI: http://dx.doi.org/10.1037/0003-066X.61.8 .760.

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological review*, *88*(1), 67-85, doi: 10.1037//0033-295X.88.1.67

Johnson, M. K., Foley, M. A., Suengas, A. G., & Raye, C. L. (1988). Phenomenal characteristics of memories for perceived and imagined autobiographical events. *Journal of Experimental Psychology: General*,*117*(4), 371. Doi: 10.1037/0096-3445.117.4.371

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

Köhnken, G. (1996). Social psychology and the law. *Applied social psychology*, 257-282.

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

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of experimental psychology: General*, *126*(4), 349. DOI: 10.1037/0096-3445.126.4.349

Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological review*, *103*(3), 490-517. doi:10.1037//0033-295X.103.3.490

Koriat, A., Bjork, R. A., Sheffer, L., & Bar, S. K. (2004). Predicting one's own forgetting: The role of expereicne-based and theory-based processes. *Journal of Experimental Psychology: General*, *133*, 4, 643-656. DOI: 10.1037/0096-3445.133.4.643.

Kornell, N., & Bjork, R. A. (2009). A stability bias in human memory: overestimating remembering and underestimating learning. *Journal of Experimental Psychology: General*, *138*(4), 449. DOI: 10.1037/a0017350

Kornell, N., Rhodes, M. G., Castel, A. D., & Tauber, S. K. (2011). The ease-of-processing heuristic and the stability bias dissociating memory, memory beliefs, and memory judgments. *Psychological Science*. DOI: 10.1177/0956797611407929

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

Legault, E., & Laurence, J. R. (2007). Recovered memories of childhood sexual abuse: Social worker, psychologist, and psychiatrist reports of beliefs, practices, and cases. *Australian Journal of Clinical and Experimental Hypnosis*, *35*, 111-133.

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

Mann, S., Vrij, A., Shaw, D. J., Leal, S., Ewens, S., Hillman, J., ... & Fisher, R. P. (2013). Two heads are better than one? How to effectively use two interviewers to elicit cues to deception. *Legal*

and Criminological Psychology, *18*(2), 324-340, https://doi.org/10.1111/j.2044-8333.2012.02055.x

Marsman, M., & Wagenmakers, E. J. (2017). Bayesian benefits with JASP. *European Journal of Developmental Psychology*, *14*(5), 545-555, https://doi.org/10.1080/17405629.2016.1259614

Masip, J., & Herrero, C. (2013). 'What Would You Say if You Were Guilty?'Suspects' Strategies During a Hypothetical Behavior Analysis Interview Concerning a Serious Crime. *Applied cognitive psychology*, *27*(1), 60-70. DOI: 10.1002/acp.2872

Masip, J., Sporer, S. L., Garrido, E., & Herrero, C. (2005). The detection of deception with the reality monitoring approach: A review of the empirical evidence. *Psychology, Crime & Law*, *11*(1), 99-122. http://dx.doi.org/10.1080/10683160410001726356

McCornack, S. A. (1997). The generation of deceptive messages: Laying the groundwork for a viable theory of interpersonal deception. In J. O. Greene (Ed.), Message production: Advances in communication theory (pp. 91-126). Mahway, NJ: Lawrence Erlbaum

McCornack, S. A., Morrison, K., Paik, J. E., Wisner, A. M., & Zhu, X. (2014). Information manipulation theory 2: a propositional theory of deceptive discourse production. *Journal of Language and Social Psychology*, *33*(4), 348-377, DOI: 10.1177/0261927X14534656

Murre J., & Dros J. (2015). Replication and analysis of Ebbinghaus' forgetting curve. *PLoS One,* 10:e0120644 10.1371/journal.pone.0120644

Nahari, G. (2017). Verifiability approach: Applications in different judgmental settings. In T. Docan-Morgan (Ed.). *The Handbook of Deceptive Communication*. Hampshire: Palgrave Macmillan.

Nahari, G., & Ben-Shakhar, G. (2011). Psychophysiological and behavioral measures for detecting concealed information: The role of memory for crime details. *Psychophysiology*, *48*(6), 733-744, DOI: 10.1111/j.1469-8986.2010.01148.x

Nahari, G., & Vrij, A. (2015). Systematic errors (biases) in applying verbal lie detection tools: richness in detail as a test case. *Crime Psychology Review*, *1*(1), 98-107, https://doi.org/10.1080/23744006.2016.1158509

Nahari, G., Vrij, A., & Fisher, R. P. (2012). Does the truth come out in the writing? Scan as a lie detection tool. *Law and Human Behavior*, *36*(1), 68. http://dx.doi.org/10.1037/h0093965

Oberlader, V. A., Naefgen, C., Koppehele-Goseel, J., Quinten, L., Banse, R., & Schmidt, A. F. (2016). Validity of Content-Based Techniques to Distinguish True and Fabricated Statements: A Meta-Analysis. *Law and human behavior*. DOI: 10.1037/lhb0000193

Ost, J., Easton, S., Hope, L., French, C. C., & Wright, D. B. (2015). Latent variables underlying the memory beliefs of Chartered Clinical Psychologists, Hypnotherapists and undergraduate students. *Memory*, 1-12. http://dx.doi.org/10.1080/09658211.2015.1125927

Otgaar, H., Howe, M. L., Smeets, T., & Wang, J. (2016). Denial-induced forgetting: False denials undermine memory, but external denials undermine belief. *Journal of Applied Research in Memory and Cognition*, *5*(2), 168-175, https://doi.org/10.1016/j.jarmac.2016.04.002

Otgaar, H., Romeo, T., Ramakers, N., & Howe, M. L. (2017). Forgetting having denied: The "amnesic" consequences of denial. *Memory & cognition*, 1-10, https://doi.org/10.3758/s13421-017-0781-5

Roediger, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in cognitive sciences*, *15*(1), 20-27, https://doi.org/10.1016/j.tics.2010.09.003

Sauer, J., & Hope, L. (2016). The effects of divided attention at study and reporting procedure on regulation and monitoring for episodic recall. *Acta psychologica*, *169*, 143-156. http://dx.doi.org/10.1016/j.actpsy.2016.05.015

Shaw, D. J., Vrij, A., Leal, S., Mann, S., Hillman, J., Granhag, P. A., & Fisher, R. P. (2013). Expect the unexpected? Variations in question type elicit cues to deception in joint interviewer contexts. *Applied Cognitive Psychology*, *27*(3), 336-343. DOI: 10.1002/acp.2911

Simons, D. J., & Chabris, C. F. (2011). What people believe about how memory works: A representative survey of the US population. *PloS one*, *6*(8). http://dx.doi.org/10.1371/journal.pone.0022757

Sporer, S. L., & Schwandt, B. (2006). Paraverbal indicators of deception: A meta-analytic synthesis. *Applied Cognitive Psychology*, *20*(4), 421-446.

Stromwall, L. A., Hartwig, M., & Granhag, P. A. (2006). To act truthfully: Nonverbal behaviour and strategies during a police interrogation. Psychology, Crime, & Law, 12, 207-219, https://doi.org/10.1080/10683160512331331328

Vredeveldt, A., van Koppen, P. J., Granhag, P. A. (2014). The inconsistent suspect: A systematic review of different types of consistency in truth tellers and liars. In R. Bull (Ed.), Investigative Interviewing (pp. 183-207). New York: Springer Science+Business Media. doi: 10.1007/978-1-4614-9642-7_10

Vrij, A. (2008). Detecting lies and deceit: Pitfalls and opportunities. Chichester, UK: John Wiley & Sons.

Vrij, A. (2015). Verbal Lie Detection tools: Statement validity analysis, reality monitoring and scientific content analysis. In P. A. Granhag, A. Vrij, & B. Verschuere. (Eds.) *Detecting Deception Current Challenges and Cognitive Approaches*. Chichester: John Wiley & Sons.

Vrij, A., & Granhag, P. A. (2014). Eliciting information and detecting lies in intelligence interviewing: An overview of recent research. *Applied Cognitive Psychology*, *28*(6), 936-944.

Vrij, A., Fisher, R. P., & Blank, H. (2017). A cognitive approach to lie detection: A meta-analysis. *Legal and Criminological Psychology*, *22*(1), 1-21. DOI: 10.1111/lcrp.12088

Vrij, A., Granhag, P. A., Mann, S., & Leal, S. (2011). Outsmarting the liars: Toward a cognitive lie detection approach. *Current Directions in Psychological Science*, *20*(1), 28-32, DOI: 10.1177/0963721410391245

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

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

Vrij, A., Leal, S., Granhag, P.A., Mann, S., Fisher, R.P., Hillman, J., & Sperry, K. (2009). Outsmarting the liars: The benefit of asking unanticipated questions. *Law and Human Behavio*r, 33, 159-166. DOI: 10.1007/s10979-008-9143-y

Vrij, A., Leal, S., Mann, S., Dalton, G., Jo, E., Shaboltas, A., ... & Houston, K. (2017). Using the Model Statement to elicit information and cues to deceit in interpreter-based interviews. *Acta psychologica*, *177*, 44-53, https://doi.org/10.1016/j.actpsy.2017.04.011

Vrij, A., Mann, S. A., Fisher, R. P., Leal, S., Milne, R., & Bull, R. (2008). Increasing cognitive load to facilitate lie detection: The benefit of recalling an event in reverse order. *Law and human behavior*, *32*(3), 253-265, DOI 10.1007/sl 0979-007-9103-y

Wagenmakers, E. J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., ... & Meerhoff, F. (2017). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review*, 1-19, https://doi.org/10.3758/s13423-017-1323-7

Wagenmakers, E. J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., ... & Matzke, D. (2017). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic bulletin & review*, 1-23, https://doi.org/10.3758/s13423-017-1343-3

Wagenmakers, E. J., Morey, R. D., & Lee, M. D. (2016). Bayesian benefits for the pragmatic researcher. *Current Directions in Psychological Science*, *25*(3), 169-176, doi: 10.1177/0963721416643289

Wagenmakers, E. J., Morey, R. D., & Lee, M. D. (2016). Bayesian benefits for the pragmatic researcher. *Current Directions in Psychological Science*, *25*(3), 169-176, DOI: 10.1177/0963721416643289

Wagenmakers, E. J., Wetzels, R., Borsboom, D., & van der Maas, H. L. J. (2011). Why psychologists must change the way they ana- lyze their data: The case of psi. *Journal of Personality and Social Psychology*, *100*, 426–432.

Wixted, J. & Ebbesen, E. B. (1991). On the form of forgetting. *Psychological Science*, 2, 409- 415. DOI: 10.1111/j.1467-9280.1991.tb00175.x

Wixted, J. & Ebbesen, E. B. (1997). Genuine power curves in forgetting: Quantitative analysis of individual subject forgetting functions. *Memory & Cognition*, 23, 731-739. DOI: 10.3758/BF03211316.

Wixted, J. T., & Carpenter, S. K. (2007). The Wickelgren power law and the Ebbinghaus savings function. *Psychological Science*, *18*(2), 133-134. DOI: 10.1111/j.1467-9280.2007.01862.x

Table 1.

Classification Scheme for the Bayes Factor, as Proposed by Jeffrey's (1961) and Adapted from Wagenmakers et al. (2011)

Bayes Factor, BF_{10}	Interpretation
>100	Extreme evidence for H_1
30-100	Very strong evidence for H_1
10-30	Strong evidence for H_1
3-10	Substantial evidence for H_1
1-3	Anecdotal evidence for H_1
1	No evidence
1/3-1	Anecdotal evidence for H_0
1/10-1/3	Substantial evidence for H_0
1/30-1/10	Strong evidence for H_0
1/100-1/30	Very strong evidence for H_0
<1/100	Extreme evidence for H_0

Table 2.

Interviewee verbal behaviour as a function of Interview Time and Veracity

	Immediate interview		Delayed interview		*t*	*df*	*p*	Cohens *d* (95% CI)
	M (*SD*)	95% CI	M (*SD*)	95% CI				
Initial statement detail (Social interaction)								
Truth-tellers	193.26 (53.04)	[177.27, 209.68]	152.43 (38.98)	[140.09, 165.77]	3.87	76	<.001***	.87 [.40, 1.32]
Liars	132.15 (51.12)	[116.81, 147.77]	154.26 (59.00)	[136.63, 173.27]	1.77	76	.081	.40 [-.05, .84]
Unique detail (Social interaction)								
Truth-tellers	251.77 (72.07)	[229.30, 274.00]	197.30 (43.06)	[184.64, 212.16]	3.97	63.55	<.001***	.92 [.44, 1.37]
Liars	196.69 (85.38)	[171.81, 225.68]	211.74 (84.36)	[184.93, 240.45]	.78	76	.436	.18 [-.27, .62]
Initial statement accuracy rate (Video)								
Truth-tellers	.92 (.05)	[.91, .94])	.81 (.18)	[.75, .86]	3.10	43.70	< .001***	.83 [.36, 1.28].
Liars	.25 (.16)	[.20, .30]	.22 (.11)	[.19, .26]	.81	76	.424	.22 [-.23, .66].
Unique detail accuracy rate (Video)								
Truth-tellers	.92 (.05)	[.90, .93]	.77 (.18)	[.71, .82]	4.82	43.50	< .001***	1.14 [.64, 1.60]
Liars	.25 (.13)	[.21, .29]	.25 (.12)	[-.42, .47]	.11	76	.910	.03, [-.42, .47]

Significance levels $p<0.05$ = *; $p<0.01$ = **; $p< 0.001$ = ***

Table 3.

Interviewee verbal behaviour as a function of Reporting Instruction and Veracity

	Model Statement-present		Model Statement-absent		*t*	*df*	*p*	Cohen's *d* (95% CI)
	M (*SD*)	95% CI	M (*SD*)	95% CI				
Unique detail (Social interaction)								
Truth-tellers	229.43 (73.57)	[208.57, 252.47]	219.89 (56.37)	[44.01, 65.88]	.64	76	.524	.14 [-.30, .59]
Liars	239.20 (83.30)	[215.59, 265.33]	167.39 (69.89)	[145.60, 189.39]	4.11	76	<.001***	.93 [.45, 1.38]
New reported detail (Social interaction)								
Truth-tellers	64.55 (28.50)	[55.78, 73.97]	38.65 (17.45)	[33.50, 44.28]	4.87	65.16	<.001***	1.09 [.60, 1.55]
Liars	89.00 (55.88)	[73.34, 107.12]	31.55 (20.55)	[25.63, 38.17]	5.97	49.83	<.001***	1.35 [.84, 1.82]
New detail accuracy rate (Video)								
Truth-tellers (Immediate)	.86 (.12)	[.70, .92]	.83 (.24)	[.70, .92]	.59	37	.559	.16 [-.47, .79]
Liars (Immediate)	.25 (.20)	[.16, .34]	.25 (.33)	[.11, .41]	.05	37	.963	.04 [-.59, .66]
Truth-tellers (Delayed)	.54 (.31)	[.41, .68]	.81 (.31)	[.66, .93]	2.73	37	.010*	.87 [.20, 1.51]
Liars (Delayed)	.26 (.25)	[.16, .38]	.29 (.14)	[.23, .35]	.40	37	.694	.15 [-.48, .77]
Unique detail accuracy rate (Video)								
Truth-tellers (Immediate)	.92 (.05)	[.86, .98]	.91 (.04)	[.86, .97]	.27	37	.786	.22 [-.41, .85]
Truth-tellers (Delayed)	.80 (.22)	[.74, .86]	.75 (.12)	[.69, 81]	.81	37	.423	.28 [-.35, .42]

Significance levels $p<0.05$ = *; $p<0.01$ = **; $p< 0.001$ = ***

APPENDIX 1. The social exchange in the waiting room

Once the participant enters the waiting room they will find another participant already sat there, in reality it will be a confederate B (CB). After one minute another participant (confederate A, CA) enters the room and an exchange between the two confederates will occur, their conversation will be briefly interrupted by another confederate (confederate C, CC) as follows:

CA) "Hiya, how are you?"

CB) "Ah not too bad thanks you?"

CA) "Yeah all good, are you still doing chemistry? I haven't seen you for a while."

CB) "No I do physics so I only share the 'thermodynamics' module from chemistry"

CA) "Oh, that's why then, I was ill and missed the last lecture on that …don't suppose you have notes do you?"

CB) "Yeah sure, in fact I may have them here (rummages in bag) but I'll need them back? (At this point confederate C enters, looks around the room and says " Oh I'm sorry I was looking for Zarah, ill see if she's in the other lab" and then leaves) CB continues talking "Perhaps you can photo copy them after this?" Finds and gives notes to CA.

CA) Great! Thanks, I'll photocopy them as soon as I've finished this and bring them back to you is that OK?

CB) Yeah fine

APPENDIX 2. The interview protocol

1. Please tell me everything that happened while you were waiting to be interviewed. Please start from the moment you entered the waiting room.
2. Now, I'd like you to focus upon telling me what the other people looked like. Attempt to describe them in enough detail so I could recognize them, but remember, *do not guess any information*. What did the others look like?
3. OK, still focusing upon the other people's appearance, can you describe for me in as much detail as possible what the other people were wearing?
4. Great. Now, I'm interested in what the other people said to each other. Try and remember, in as much detail as possible, what they said to each other whilst you waited. Even fragments of their conversation can be valuable so don't leave out any detail, even it appears irrelevant.'
5. So, I need you describe for me where the others were sat in the waiting room. Please describe where the others were sat relative to where you were sitting.

I'm now going to ask you questions about the surveillance device in the video

6. Please tell me, in as much detail as possible, what the device in the video looked like.
7. Now I need you to recall for me everything you can remember that the device could do. Try and remember, in as much detail as possible, what they said its functions where.
8. Finally, I need you to tell me where the device is going to be planted; please give me as much information about this location as you can recall.

APPENDIX 3. Condition Instructions

All liars

You are in the 'red' team but you will be interviewed by a member of the opposing blue team and as such it is important that you mislead the interviewer about certain details of the video, including what the device looked like and could do and also the location that is chosen to plant the device. The interviewer knows that the device would be placed somewhere, but does not know where. So, above all, you must *not* reveal the location that was selected to hide the spy device and your objective is to mislead the investigator by using the third location mentioned in the video as the location that was selected to plant the device.

You also need to mislead the interviewer about the device. The interviewer knows something about the device but does not have all the details, and it is not clear what the interviewer knows. Because of this, you need to provide *some truthful* and *some false* information about the device. This will help you to appear cooperative without having to tell the interviewer everything.

In addition, you should be mindful as to what other Red team members may be doing. If you encounter anyone that uses the word 'thermodynamics' this means that those people are also on the Red team. In all likelihood, the interviewer is aware that they are around, so no point to deny that if you see them. However, you should protect them by not telling the truth about what they looked like and what they said. It is important to appear cooperative. If the interviewer believes you are cooperative you earn [£10/ £15]. In addition, you will be entered into a draw to win up to £150 in prize money. If you do not appear cooperative, you will be asked to write a statement about what happened today.

After watching the video (in the Immediate condition) or after returning after three-weeks (in the Delayed condition), the participants will be told:

You are in the 'red' team but you will be interviewed by a member of the opposing blue team about the video and the exchange between the two people in the waiting room. As such it is important that you mislead the interviewer about certain details of the video, including what the device looked like and could do and also the location that is chosen to plant the device.

The interviewer knows that the device would be placed somewhere, but does not know where. So, above all, you must *not* reveal the location that was selected to hide the spy device and your objective is to mislead the investigator by using the third location mentioned in the video as the location that was selected to plant the device.

You also need to mislead the interviewer about the device. The interviewer knows something about the device but does not have all the details, and it is not clear what the interviewer knows. Because of this, you need to provide *some truthful* and *some false* information about the device. This will help you to appear cooperative without having to tell the interviewer everything.

In addition, you witnessed in the waiting room an encounter with two members of the Red team prior to watching the video. In all likelihood, the interviewer is aware that they were around, so no point to deny that you saw them. However, you should protect them by not telling the truth about what they looked like and what they said. It is important to appear

cooperative. If the interviewer believes you are cooperative you earn [£10/£15]. In addition, you will be entered into a draw to win up to £150 in prize money. If you do not appear cooperative, you will be asked to write a statement about what happened today.

All truth-tellers

You are in the 'Blue' team and you will be interviewed by a member of your team so it is important that you are totally truthful to the interviewer and provide as much information as you can recall. Note that there are also participants on the Red team about today and you should be mindful as to what other teams' members may be doing. If you encounter anyone that uses the words 'Rocket Science' this means that those people are also on your 'Blue' team, if you do not hear those words then they are members of the opposing team and you need to watch out for anything they do as that may be useful to your Blue team. It is important to appear cooperative. If the interviewer believes you are cooperative you earn [£10/£15]. In addition, you will be entered into a draw to win up to £150 in prize money. If you do not appear cooperative, you will be asked to write a statement about what happened today.

After watching the video (in the Immediate condition) or after returning after three-weeks (in the Delayed condition), the participants will be told:

You are in the 'Blue' team and you will be interviewed by a member of your team so it is important that you are totally truthful to the interviewer and provide as much information as you can recall about the exchange between the two people in the waiting room and about the video. It is important to appear cooperative. If the interviewer believes you are cooperative

you earn [£10/£15]. In addition, you will be entered into a draw to win up to £150 in prize money. If you do not appear cooperative, you will be asked to write a statement about what happened on that day.

Supplementary materials

Intelligence Video Event Reporting

Reported detail

Since the Reporting Instruction (Model Statement-present or Model Statement-absent) was introduced after the initial statement, this factor is not introduced in the analyses regarding the initial statement. Moreover, new details reported in the final statement relate necessarily to the amount of details reported in the initial statement. That is, the more details reported in the initial statement, the fewer new details are likely to be available for subsequent reporting. We therefore included the number of details reported in the initial statement as a covariate in the analyses of new details.

A 2 (Veracity: Truth-teller vs. Liar) x 2 (Interview Time: Immediate versus Delayed) ANOVA with the total number of details reported during the initial statement as the dependent variable revealed a significant main effect of Veracity, $F(1,152) = 13.14$, $p < .001$, Cohen's $d = .56$, *95% CI* [.23, .87], such that more details were provided by truth tellers than liars. The main effect for Interview Time was not significant, $F(1,152)= 1.31$, $p = .254$, $d = .17$, 95% CI [-.15, .48]. However, a significant Veracity x Interview Time interaction effect emerged, $F(1,152)= 14.81$, $p < .001$, $\eta_p^2 = .09$.

Regarding the Veracity x Interview Time interaction effect, liars in the Immediate condition (M=50.28, SD= 16.41, 95% CI [45.03, 55.49]) and liars in the Delayed condition (M=56.59, SD= 26.77, 95% CI [48.71, 65.41]) reported similar amounts of detail during the initial statement, $t(63.02)=1.26$, p=.214, d= .28, 95% CI [-.17, .73]. In contrast, truth tellers in

the Immediate condition (*M*= 62.56, *SD*= 16.32, 95% CI [57.49, 67.16]) reported more detail than truth tellers in the Delayed condition during the initial statement (*M*= 48.28, *SD*= 19.49, 95% CI [42.05, 54.28]), *t*(76)= 3.51, *p*=.001, *d*=.79 [.32, 1.24].

A 2 (Veracity: Truth-teller vs. Liar) x 2 (Interview Time: Immediate versus Delayed) x 2 (Reporting Instruction: Model Statement-present versus Model Statement-absent) ANCOVA with the total number of new details reported in the final statement as dependent variable, and the total amount of detail reported during the initial statement as covariate revealed a significant main effect for Veracity, $F(1,147) = 5.13$, $p = .025$, $d = .30$, 95% CI [-.02, .61], with liars (*M* = 19.69, *SD* = 22.12, 95% CI [14.86, 24.57] reporting more new details than truth tellers (*M* = 14.35, *SD* = 12.32, 95% CI [11.59, 22.89]). A significant main effect also emerged for Reporting Instruction, $F(1,147) = 19.95$, $p < .001$, $d = .72$, 95% CI [.39, 1.03], with interviewees in the Model Statement-present condition (*M* = 22.99, *SD* = 20.76, 95% CI [18.44, 27.34]) reporting more new details than interviewees in the Model Statement-absent condition (*M* = 10.74, *SD* = 11.88, 95% CI [8.15, 13.72]). All other effects were not significant, all *F*'s < 3.31, all *p*'s > .071.

A 2 (Veracity: Truth-teller vs. Liar) x 2 (Interview Time: Immediate versus Delayed) x 2 (Reporting Instruction: Model Statement-present versus Model Statement-absent) ANOVA with total unique detail (detail reported in the initial statement plus new detail reported in the final statement) as dependent variable revealed a significant main effect for Reporting Instruction, $F(1,148) = 9.60$, $p = .002$, $d = .49$, 95% CI [.16, .81], with interviewees in the Model Statement-present condition (*M* = 78.84, *SD* = 34.56, 95% CI [71.75, 86.46]) reporting more unique detail than interviewees in the Model Statement-absent condition (*M* = 63.67, *SD* = 26.36, 95% CI [57.66, 70.10]). The main effect for Veracity was

not significant, $F(1,148) = .44$, $p = .510$. Other than the Veracity x Interview Time interaction effect, $F(1,148) = 6.10$, $p = .015$, $\eta_p^2 = .04$, no other effects were significant, all F's < 1.00, all p's > .319.

Regarding the Veracity x Interview Time interaction effect, liars in the Immediate condition (M=68.13, SD= 27.68, 95% CI [59.44, 76.90]) and liars in the Delayed condition (M=78.13, SD= 45.39, 95% CI [64.18, 93.51) reported similar amounts of unique detail, $t(62.84)= 1.18$, p=.245, d= .27, 95% CI [-.18, .71]. In contrast, truth tellers in the Immediate condition (M=76.95, SD= 22.55, 95% CI [69.92, 83.54]) reported more unique detail than truth tellers in the Delayed condition (M=62.59, SD=24.23, 95% CI [55.03, 70.92]), $t(76)= 2.71$, p=.008, d=.61 95% CI [.15, 1.06].