Lying about Flying:

The Efficacy of the Information Protocol and Model Statement for Detecting Deceit

Aldert Vrij[12]

Sharon Leal

Haneen Deeb

Department of Psychology, University of Portsmouth

Stephanie Chan

Majeed Khader

Whistine Chai

Jeffery Chin

Home Team Behavioural Sciences, Ministry of Home Affairs, Singapore

[1] Correspondence concerning this report should be addressed to Aldert Vrij, Department of Psychology, University of Portsmouth, King Henry Building, King Henry 1 Street, PO1 2DY, Hants, United Kingdom. Email: aldert.vrij@port.ac.uk

2 The authors have no conflict of interest to declare

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/acp.3614

Abstract

Due to time constraints interviews aimed to detect deception in airport settings should be brief and veracity assessments should be made in real time. In two experiments carried out in the departure hall of an international airport, truth tellers were asked to report truthfully their forthcoming trip, whereas liars were asked to lie about the purpose of their trip. In Experiment 1 we examined five verbal cues we thought had potential to discriminate truth tellers from liars in short airport interviews: Elaboration in disclosing information, checkable details, how many people are aware of the trip, complications, and plausibility. In Experiment 2 we attempted to improve the interview protocol by adding a second interview phase in which we introduced an Information Protocol and Model Statement. All five cues differentiated truth tellers from liars in both experiments but the Information Protocol and Model Statement did not enhance these differences.

Lying about Flying:

The Efficacy of the Information Protocol and Model Statement for Detecting Deceit

To keep air travel safe from terrorist attacks Government agencies sometimes employ lie detection techniques. For example, in the USA the Transportation Security Administration (TSA) has employed thousands of trained behaviour detection officers (BDO) over the past ten years to identify passengers exhibiting nonverbal behaviours indicative of stress, fear, or deception at airport screening checkpoints. According to TSA, certain behavioural indicators may indicate mal-intent, such as the intent to carry out a terrorist attack. The USA Government Accountability Office (2017) reviewed the evidence behind the TSA program of spotting possible wrongdoers at American airports. It concluded that the TSA has no evidence that most indicators it uses to identify wrongdoers at airports are actually suitable for that purpose.

In this article, we focus on detecting lies through analysing speech. Verbal lie detection in airport settings is challenging. Since there are so many passengers at airports, the interviews should be brief. Short interviews hamper verbal lie detection as more speech cues to deception occur when the interviews become longer. This is unsurprising as words are the carriers of verbal cues to deceit (Vrij, Mann, Kristen, & Fisher, 2007). This implies that speech related cues found to be diagnostic of deceit in longer interviews, do not automatically discriminate truth tellers from liars in shorter interviews.

The present project consisted of two experiments. In Experiment 1 we tested whether five verbal cues could be used to discriminate truth tellers from liars in short airport interviews: Elaboration in disclosing information, checkable details, how many people are aware of the trip, complications, and plausibility. We selected these cues for two reasons. First, there is empirical evidence that they are reliable verbal cues in discriminating truth tellers and liars investigative interview settings and we thought they had the potential to also work in airport settings. Second, we thought that in case they did discriminate truth tellers from liars, immigration officers would be able to code them easily in real time in their interviews. Although we did not test this second point in the current two experiments, it is an important requirement, because only verbal lie detection techniques that can be used in real time have a chance to be implemented at airports. In Experiment 2 we attempted to improve the diagnostic value of the interview protocol in various ways: We examined the effect of an Information Protocol and Model Statement on truth tellers' and liars' statements. We start this article by introducing the five verbal cues.

Elaboration in disclosing information

Truth tellers typically report more details than liars (Amado, Arce, Fariña, & Vilarino, 2016). First, liars are unable to provide as many details as truth tellers because they lack the imagination to fabricate details that sound plausible (Köhnken, 2004; Leal, Vrij, Warmelink, Vernham, & Fisher, 2015). Second, liars provide fewer details than truth tellers for strategic reasons. They are unwilling to provide many details out of fear that those details give leads to investigators that they are lying (Nahari, Vrij, & Fisher, 2014a) or out of fear that they cannot

remember them anymore when asked the same question again (Vrij, 2008). We expect that practitioners in real life would have difficulty counting in real time the number of details someone provides in an interview. Also, this counting may come at the expense of actually listening to the answers an interviewee gives. We thought it may be easier for practitioners to judge whether someone elaborates when answering the questions. The reasoning is that longer detailed answers require more elaboration from the interviewee than shorter answers that lack detail.

The Verifiability Approach

Truth tellers and liars employ different strategies in interviews. To explain these differences, Granhag and Hartwig (2008) refer to self-regulation theory (Fiske & Taylor, 1991). They claim that both liars and truth tellers view the upcoming interview as a potential threat. Both run the risk of not being believed by the investigator, which is a stressful experience that can have far-reaching negative consequences. However, the threat between liars and truth tellers differ: Liars may be afraid that investigators could discover all they know about their own activities, whereas truth tellers may be afraid that investigators may *not* come to know all they know about their own activities.

This different type of fear between truth tellers and liars leads to different strategies (Granhag & Hartwig, 2008): Truth tellers are inclined to be open and to tell all they remember about their activities. In contrast, liars are motivated not to tell the investigator all they know, particularly not the information they fear incriminates them. A possible strategy for liars to use to avoid detection is to provide information that cannot incriminate them. That is, when attempting to make an honest impression, liars may choose to provide details that are difficult to verify (e.g., 'I'm heading into Auckland in New Zealand to meet up with some music producers in a hotel down there') and may avoid providing details that are easy to verify (e.g., 'I phoned this producer in New Zealand on my mobile this morning to confirm the tour schedule'). The core of the Verifiability Approach is that liars provide fewer details that can be checked than truth tellers. This is a common finding in the Verifiability Approach literature to date, which exists of more than ten studies (Vrij & Nahari, 2019).

The Verifiability Approach has been tested in an airport setting before (Jupe, Leal, Vrij, & Nahari, 2017) and also in that study it was found that truth tellers provided significantly more verifiable details than liars. Unlike Jupe et al. (2017), who examined the Verifiability Approach in only one way, we examined the Verifiability Approach in three different ways. First, similar to Jupe et al. (2017), we examined the checkable details reported by truth tellers and liars. This means that we sought to replicate their findings, which provides insight into the robustness of the Verifiability Approach findings. Second, we explicitly asked participants how many people were aware of the trip they were going to make. The advantage of this question is that immigration officers can easily count the number of people mentioned. The Verifiability Approach predicts that liars will be less keen to mention people than truth tellers out of fear that these people will then be consulted. Third, in

a post-interview questionnaire, we asked participants about the underlying assumptions of the approach, particularly about the motivation to include checkable details.

Although in the original Verifiability Approach research verifiable details were examined, a recent development is to examine verifiable sources (Leal et al., 2018). To explain the difference, the sentence "I spoke with my friend Fred at 10.30 this morning" contains five verifiable details (the underlined words) and one verifiable source ('Fred'). We counted verifiable sources rather than verifiable details for applied reasons, because the number of verifiable sources is easier to count in real time than the number of verifiable details. Verifiable sources and verifiable details are related to each other because a verifiable source leads to verifiable details.

The diagnostic value of the Verifiability Approach will lessen if liars make up details that can be checked, which we will call bluffing. We measured in the post-interview questionnaire the extent to which liars were inclined to bluff.

Complications

A complication is "an occurrence that makes a situation more difficult to report than necessary" (Vrij, Leal, & Fisher, 2018a; Vrij, Leal, Jupe, & Harvey, 2018c). For example, if someone says that someone had to change his ticket because his grandson was born earlier than expected, he describes a complication. If the person then adds that the grandson is now in a hospital far away from the airport, he adds another complication. Complications occur more often in truthful statements than in deceptive statements (Amado, Arce, & Fariña, 2015; Vrij, Leal, Jupe, & Harvey, 2018c). In interviews, liars prefer to keep their stories simple (Hartwig, Granhag, & Strömwall, 2007), but adding complications makes the story more complex. Complications have been measured before (Amado et al., 2015; Vrij et al., 2017, 2018b, 2018c) but always in longer interviews. We will test for the first time whether complications is a diagnostic cue in short interviews.

Plausibility

Plausibility refers to the extent to which a statement seems consistent with what the observer knows about how the real world works. If someone says that during the first day in London he would visit Buckingham Palace, Windsor Castle and Tate Gallery, the story sounds implausible because this is just too much to do in one day: The distances between these three sites are substantial and a visit to each site takes a few hours at least. Plausibility is a somewhat subjective cue, as it leaves room for individual interpretation, but different judges appear to agree to a satisfactorily extent how plausible a story is (Leal, Vrij, Warmelink, Vernham, & Fisher, 2015; Vrij, Fisher, Blank, Leal, & Mann, 2016). In a meta-analysis concerning cues to deceit, plausibility emerged as a diagnostic cue with truthful stories being judged as more plausible than deceptive stories (DePaulo et al., 2003).

Hypotheses

We tested the following eight hypotheses:

- Truth tellers will elaborate more when providing their statements than liars (Hypothesis 1) (measured through verbal coding of statement)

- Truth tellers will be more motivated than liars to provide lots of detail during the interview (Hypothesis 2) (measured in post-interview questionnaire)

- Truth tellers will include more checkable details (coded as checkable sources) into their account than liars (Hypothesis 3) and will mention more people who are aware of their trip than liars (Hypothesis 4) (both measured through verbal coding of statement)

- Truth tellers will be more motivated than liars to provide details the investigator can check, whereas liars will be more motivated than truth tellers to include details the investigator cannot check (Hypothesis 5) (measured in post-interview questionnaire)

- Truth tellers will find it easier to include checkable details into their accounts than liars (Hypothesis 6) (measured in post-interview questionnaire)

- Truth tellers will include more complications in their stories than liars (Hypothesis 7) (measured through verbal coding of statement)

- Truth tellers' statements will sound more plausible than liars' statements (Hypothesis 8) (measured through verbal coding of statement)

Method

Participants

A total of 80 participants took part (57 males) and their age ranged from 18 to 70 years old (M = 38.98, SD = 13.45).

Procedure

The study took place airside at Singapore Changi Airport. Passengers who were not traveling in groups were approached directly after passport control and were first asked whether they were going home. Those who were going home were thanked and told that they could not take part. Those who did not travel home were asked to take part in a short study testing airport security interview protocols. Those who did not want to take part did so for various reasons, including that they had no time, wanted to go shopping or having something to eat or drink, because they were tired, or because their English was not good enough. Those who took part (although we did not record it we estimated that around 25% of the people being asked took part) were taken to a large room where the study took place.

The experimenter introduced the study to the participants. They first read and signed an informed consent form. They were then asked where they were traveling to and what the reason was for their trip. They were then randomly allocated to the truth or lie condition.

Truth tellers (n = 41) were told that they would be interviewed in English by an immigration officer about their trip and were requested to answer each question truthfully. They were also told that they had to pretend that they just arrived at the airport of their final destination, so that the interview would reflect an immigration interview. Liars (n = 39) were also told that an immigration officer would interview them in English about their trip. They were requested to be honest about their final destination but to lie about their activities at that destination. They were also told that they had to pretend that they just arrived at the airport of their final destination.

The experimenter asked the participants whether they had any questions about the Procedure and then offered them time to prepare themselves for the interview. When they said they were prepared (preparations never took long, just a few minutes) participants were brought to an interview-counter where the interview took place. Professional immigration officers, who were blind to the veracity status of the participants, carried out the interviews. We used numerous officers and each officer interviewed truth tellers and liars. They were instructed to keep themselves to the standard interview protocols without leaving out questions or adding any questions. The protocol consisted of four questions, which were always asked in the same order. The first question focused on the activities during the trip and the second on the planning of the trip. We hereby explicitly encouraged participants to provide information that the investigator could check. The third trip focused on people that would be aware of the trip and the fourth questions asked for reassurances that the participants were truthful. A copy of the exact interview protocol requires written approval from the authors of the Home Team Behavioural Sciences Centre.

The interviews were audiotaped. The participants were made aware of this when they were introduced to the experiment and have given us permission to audiotape their interviews. Reading the transcribed interviews revealed that investigators kept themselves well to the instructions to follow the interview protocol and that all participants understood what was expected from them.

The participants were asked to describe their experiences in a post-interview questionnaire. First, they were asked how motivated they were to do well in the interview. Answers were given on a 7-point Likert scale ranging from [1] extremely unmotivated to [7] extremely motivated. They then were asked to what extent they were motivated to provide (ii) lots of details; (iii) details the investigator can check and (iv) details the investigator cannot check, again on 7-point Likert scales ranging from [1] extremely unmotivated to [7] extremely motivated. Next they were asked to what extent they found it easy to provide checkable details, which the answer given on a 7-point Likert scale ranging from [1] extremely difficult to [7] extremely easy.

The final two questions examined (v) To what extent the participant did bluff (defined in the questionnaire as providing false information that can be checked in the responses) (7-point Likert scale ranging from [1] not at all to [7] to a very great extent) and (vi) To what extent the participant did tell the truth (11-point scale ranging from 0% to 100%). After

completing the questionnaire participants received a small present, were thanked and debriefed.

Coding

The interviews were transcribed. Two coders who were blind to the Veracity conditions coded the interviews independently from each other for five cues. Participant's *Elaboration* in answering the two questions in which they were asked to describe their trip in detail was measured with two dichotomous yes/no scales: Example of two answers that revealed elaboration are: "I'll be arriving in Taipei this evening, linking up with my colleague at the airport arriving on flight UA871, from the US, and we'll be travelling to (city mentioned). We'll be doing the dinner there and we'll go to sleep" and "Well this trip just came up, it's a business trip, so it- it came up from a meeting last week, I had with a customer in senior management, and they wanted to go down to our factory in (city mentioned), to check out some of those production delays. That's what the meeting is about tomorrow"

Example of two answers that revealed no elaboration are: "We are going to Taipei first and then we're gonna take a coach to somewhere near (city mentioned). And then we are going to our accommodation" and "I didn't do much because it's all planned for us as a group"

Checkable details are i) Activities with identifiable or named persons who the interviewer can consult ("My sister's friend, Fred, will pick me up from the airport", ii) Activities that have been witnessed by identifiable or named persons who the interviewer can consult ("My neighbour was with me when I booked the safari"); iii) Activities that the interviewee believes may have been captured on CCTV ("I was in the travel agency to discuss this trip as their CCTV footage will demonstrate"); and iv) Activities that may have been documented or recorded through technology (other than CCTV), such as using debit cards, mobile phones, or computers (Harvey, Vrij, Nahari, & Ludwig, 2017; Vrij & Nahari, 2019). Note that details need to have a certain specificity to count as verifiable. Thus, 'neighbour' is specific and counts as a verifiable detail but 'friend' or 'colleague' is not. For the same reason 'booked the safari' does not count as a verifiable detail because it does say where it is booked or where the booking actually is. In the sentence above the booking can only be checked via the neighbour. Checkable details were coded for all four questions using dichotomous yes/no scales. Four examples of checkable details are: 1) "I have the hotel booking on my phone"; 2) "I link up with the president of the Vietnam traditional martial arts group, Mr (name of person); 3) I'll be picked up by my business partner in Taiwan, Mrs (name of person); and I've booked everything with my credit card, so if I give you my credit card statement you'll get the bills of the stuff that I've done.

There is a difference between labs in coding verifiable details. Although in our coding system 'friend' does not count as a verifiable detail, it counts as a verifiable detail in Nahari's lab (Nahari, 2018), which as argument used that an interviewee in all likelihood will expect further questions about a 'friend'. Nahari and Nisin (2019) refer in this context to differences in resolution levels of verifiability and argue that further research into the resolution of verifiability may strengthen the Verifiability Approach.

We further coded the *number of people* the participant said were aware of the details of the trip. For example the sentence "The office assistant (name of the assistant) is aware of my trip, my boss is aware, and of course my colleagues at the factory" contains three people who could be consulted (colleagues was counted as a single entity) and the answer: "My HR in (name of company), and my boss and colleagues know that I'm going for this business trip" also contains three people who could be consulted.

Complications were defined as "an occurrence that makes a situation more difficult than necessary" (Vrij, Leal, & Fisher, 2018a; Vrij, Leal, Jupe, & Harvey, 2018c). This definition differs somewhat from how a complication is defined in the CBCA literature (a reported activity or event that was not expected or planned (Steller & Köhnken, 1989). Information which is considered to be a complication according to the CBCA definition is also a complication according to this definition, but our definition includes more information as complications. For example, if someone says that when driving from their hometown A to their holiday destination B, they also paid a short visit to town C, this short visit to C is considered a complication according to our definition but not according to the CBCA definition (Vrij, Leal, Jupe, & Harvey, 2018c). A Complication in the answer was coded for the two questions in which participants were asked to describe their trip in detail on dichotomous yes/no scales. Two examples of complications are: 1) "I'll be arriving in Taipei this evening, linking up with my colleague at the airport arriving on flight UA871 from the US" and 2) "Tonight I'll go straight to the hotel, it's actually about two hours outside of Beijing.

Plausibility was defined as: 'How likely is it that this person is going to make the trip as described in the interview?' Answers were given on a 7-point Likert scale ranging from [1] very implausible to [7] very plausible.

Inter-rater reliability (Intraclass Correlation Coefficient, ICC) was calculated using a two-way mixed effects model using the average measures coefficients. These coefficients were good for all five variables: elaboration [ICC = .77), checkable details [ICC = .92], number of people [ICC = .96], complications [ICC = .80], and plausibility [ICC = .80]. For elaboration, checkable details, number of people and complications, the two coders discussed their disagreements and came to an agreement about the final coding. For plausibility, the two scores of the coders were averaged.

Results

Questionnaire Results

A one-way MANOVA with Veracity as the only factor and the seven questions from the questionnaire depicted in Table 1 as dependent variables revealed at a multivariate level a significant Veracity main effect, $F(7, 72) = 46.72$, $p < .001$, $\eta_p^2 = .82$.

Table 1 about here

Table 1 shows that five out of seven variables were significant at a univariate level. Truth tellers were more motivated than liars to provide details the investigator can check,

whereas liars were more motivated than truth tellers to provide details the investigator cannot check. This supports Hypothesis 5. Truth tellers also found it easier than liars to provide details someone can check, supporting Hypothesis 6. In addition, liars were more engaged than truth tellers in bluffing and truth tellers were more truthful in their answers than liars. Note that liars were not entirely deceptive (they reported that 32.05% of what they said was true). That liars reported also information that was truthful is a typical finding in deception research, a phenomenon called embedded lies (Leins, Fisher, & Ross, 2013; Vrij, 2008). Regarding bluffing, the mean score of 5.69 obtained by liars is high and showed a considerable inclination to provide false information that someone can check. For all these significant findings the effect sizes were substantial.

Regarding the non-significant findings, truth tellers and liars were equally and both highly motivated to do well in the interview (both mean scores are 5.85 on a 7-point Likert scale). Truth tellers and liars were also equally and highly motivated to provide details (scores were around 5.50), but the absence of a Veracity effect means that Hypothesis 2 is rejected.

We also calculated the Bayes factors using JASP software. Bayes factors allow for further inferences to be drawn regarding the evidence in an experiment (Jarosz & Wiley, 2014). Bayes factor analysis is a method to test the probability of the observed data under the null hypothesis compared to the alternative hypothesis (Wetzels & Wagenmakers, 2012). A BF_{10} smaller than 1 indicates evidence for the absence of an effect (support of the null hypothesis). According to the cut-off thresholds provided by Jeffreys (1961), BF values between 1 and 3 suggest weak evidence for the alternative hypothesis, values between 3 and 10 suggest substantial evidence, values between 10 and 20 constitute strong evidence, and values above 20 constitute very strong evidence. We used the default Cauchy's prior of .707 for the Bayesian *t*-tests (Lakens, 2016). The Bayes Factor analyses confirm the analyses of variance reported above: They revealed substantial to very strong evidence for the five significant effects and support for the null-hypothesis for the two non-significant effects.

Verbal cues

A MANOVA was carried out with Veracity as the only factor and the five verbal cues depicted in Table 2 as dependent variables. The multivariate effect was significant, $F(5, 74) = 9.08$, $p < .001$, $\eta_p^2 = .38$. The univariate results are depicted in Table 2.

Table 2 about here

Table 2 shows that all variables measured in this experiment significantly distinguished truth tellers from liars. Truth tellers elaborated more when answering the questions than liars, which supports Hypothesis 1, and truth tellers included more checkable sources in their answers than liars, which supports Hypothesis 3. In addition, truth tellers included more complications in their answers, which supports Hypothesis 7. Finally, truthful stories sounded more plausible than deceptive stories, supporting Hypothesis 8. The effect sizes were high and the Bayes Factor analyses revealed strong to very strong evidence for these hypotheses. Although truth tellers mentioned significantly more people than liars, the

Bayes Factor analyses showed only weak evidence for this effect, which means we only received weak support for Hypothesis 4.

Plausibility Prediction

The effect size for plausibility was very large (d = 1.42) and the BF_{10} score revealed very strong evidence, indicating a clear difference between truth tellers and liars in terms of plausibility. Since plausibility is a subjective cue we examined which elements of a statement contributed to plausibility by carrying out a regression analysis, using the 'enter' method. We entered as predictors the four variables other than plausibility which are depicted in Table 2. The model could explain 81% of the variance, r^2 = .65, $F(4, 752) = 35.16$, $p < .001$. Two significant predictors emerged: Verifiable details was the strongest predictor, *beta* = .49, t = 4.94, $p < .001$, followed by elaboration, *beta* = .29, t = 3.47, p = .001. Complications, *beta* = .05, t = 0.65, p = .518, and number of people mentioned, *beta* = .15, t = 1.72, p = .090, did not contribute to the model.

Discussion

Truth tellers and liars differed significantly from each other regarding each of the five verbal cues we examined. First, truth tellers elaborated more when volunteering information than liars. Although truth tellers and liars thought they put similar effort into their stories according to post-interview questionnaire self-reports (rejecting Hypothesis 2), this became evident in their answers during the interview, supporting Hypothesis 1. The combination of these two findings suggests that liars thought they put enough effort into their stories, whereas in fact they did not.

We also found empirical support for the Verifiability Approach. The post-interview self-reports showed that truth tellers were more motivated than liars to include details an investigator can check, whereas liars were more motivated than truth tellers to include details an investigator cannot check, supporting Hypothesis 5. The post-interview questionnaire also showed that truth tellers found it easier to include checkable details than liars, supporting Hypothesis 6. In the actual interview truth tellers included more checkable details (coded as checkable sources) than liars in their answers (supporting Hypothesis 3). Finally, when asked who is aware of the details of their trip, truth tellers mentioned more people than liars, but the Bayes Factor analysis revealed only weak support for this (Hypothesis 4). Note that the support for the Verifiability Approach occurred despite liars' strong inclination to bluff by making up verifiable details.

We also found that truth tellers included more complications in their answers than liars, supporting Hypothesis 7 and found that the stories of truth tellers sounded more plausible than those of liars, supporting Hypothesis 8. The effect size for plausibility was the largest effect obtained in this experiment. A disadvantage of plausibility is that it is a more subjective cue than any of the other verbal cues we measured, which means that plausibility scores are open for interpretation and perhaps differences between immigration officers may emerge in assessing plausibility. Promising in this respect is that in the present experiment the two plausibility coders obtained good reliability suggesting that plausibility coding in

short airport interviews perhaps could be reliably assessed. Because plausibility is subjective, we examined which cues that are more objective could explain it. Plausibility was positively linked with verifiable details and elaboration. The reference to verifiable details suggests that the coders took a credibility indicator into account when determining plausibility. This is understandable given that plausibility is defined in the deception literature as 'degree to which the message seems plausible, likely or believable' (DePaulo et al., 2013, p.113).

Out of the five verbal cues we examined, 'elaboration' had not been examined before, but the remaining four cues already have been examined in deception research, albeit not in short interviews in an airport setting. Since the diagnostic value of verbal cues to deceit may vary across interview settings (Vrij et al., 2007), it is important to test the diagnostic value in different settings. Experiment 1 demonstrated the utility of these verbal cues to distinguish between truth tellers and liars in short airport interviews.

Experiment 2

Experiment 2 had three aims. First, we sought to replicate the findings of Experiment 1. Phase 1 of Experiment 2 was largely identical to the interview of Experiment 1 and we examined the eight hypotheses examined in Experiment 1. We believe the ability to replicate is necessary for any lie detection tool before it can be recommended to be used in real life. We made two changes compared to Experiment 1. First, we examined one additional verbal cue to the list of verbal cues reported in Experiment 1 -total details. We measured total details for two reasons. First, although we realise that immigration officers in all likelihood cannot count in real time the number of details a passenger provides, we thought it worthwhile to examine this cue because we expected it to be related to the variable elaboration which we believe can be easily coded in real time: The more details someone provides, the more the person gives the impression to elaborate when providing details. Second, in Experiment 2 the interviewees were interviewed twice which gave us the opportunity to measure consistency in their answers. However, to measure consistency, counting details is necessary.

The second aim of Experiment 2 was to improve the diagnostic value of the interview protocol in various ways: We examined consistency between the two phases and the effect of an Information Protocol and Model Statement (in Phase 2 of Experiment 2).

The third aim was that, unlike in Experiment 1, we investigated participant's first language and whether this affected the results. There is evidence of differences in skin conductance responses of liars speaking in their first language or second language (Caldwell-Harris, 2015) and observers seem to evaluate first-language and second-language speakers differently (Da Silva & Leach, 2013). It is therefore worthwhile examining whether verbal lie detection tools are sensitive to first or second language speaking.

Two interviews - Consistency

In the present experiment passengers were interviewed twice - by different interviewers - and were asked the same questions in both interviews. This enables us to measure consistency. There is a strong belief amongst laypersons and professionals that

consistency indicates truthfulness and inconsistency indicates deception (Deeb, Vrij, Hope, Mann, Granhag, & Strömwall, 2018; Strömwall, Granhag, & Hartwig, 2004; Vredeveldt, van Koppen, & Granhag, 2014). This is actually not the case, as liars can be equally consistent or even more consistent than truth tellers (Vredeveldt et al., 2014). This can be explained via the repeat versus reconstruct hypothesis (Granhag, & Strömwall, 1999). Liars are probably aware that observers associate inconsistencies with deceit and therefore will try to avoid those to remain consistent. When interviewed multiple times liars therefore tend to repeat what they have said before. Truth tellers are less concerned with impression management than liars. They therefore will go back to their memory when asked about the same event again and will tell their story as they remember it then, which can easily result in inconsistencies.

The concept consistency entails four elements (Fisher, Vrij, & Leins, 2013): Repetitions (repeat what has been said before), reminiscences (new information not reported before), omissions (information left out that has been reported before) and contradictions (contradict what has been said before). A perfectly consistent story just contains repetitions, the other three elements lead to inconsistency. Since liars tend to focus on consistency we can predict liars to repeat themselves more than truth tellers. As truth tellers rarely report all they know in the first instance, a second recall may make truth tellers, more than liars, report new information (reminiscences) and this may occur at the expense of information they said before (omissions). Contradictions occur when someone misremembered what s/he said before, which is more likely to occur amongst liars.

Information Protocol

The information protocol is part of the Verifiability Approach and refers to explicitly instructing the interviewee to include details the investigator can check. This instruction typically makes the difference in reporting verifiable details between truth tellers and liars more pronounced as the instruction typically results in truth tellers adding more verifiable details than liars (Harvey et al., 2017; Nahari, Vrij, & Fisher, 2014b). The instruction makes truth tellers think about verifiable details they had not thought about before, whereas liars will find it difficult to report details that support their fabricated stories. Four types of verifiable detail are typically distinguished: Activities: (i) carried out with or (ii) witnessed by a named or otherwise identifiable person; iii) the interviewee believes may have been captured on CCTV; and iv) that may have been documented or recorded through technology (other than CCTV), such as using debit cards, mobile phones, or computers. In the present experiment we examined whether explicitly informing interviewees about these verifiable activities would enhance the efficacy of the Verifiability Approach. We thought it would because it will give truth tellers a better understanding of what verifiable details exactly are and this better understanding may enhance their ability to report them.

The four types of verifiable detail discussed in the previous paragraph are always clustered into one overall ‘verifiable details’ category. In the present experiment, however, we split up this single category into two separate categories: verifiable witness details (categories i and ii above) and verifiable digital details (categories iii and iv above). Such a distinction may reveal liars’ bluffing strategies (providing false verifiable details). When liars bluff, do they

have a preference which type of verifiable details (witness or digital) they fabricate? Insight into which type of verifiable details liars tend to fabricate may benefit investigators who use the Verifiability Approach. We explored the difference between truth tellers and liars regarding reporting verifiable witness and verifiable digital sources.

Model Statement

A model statement (MS) is an example of a detailed report about an event unrelated to the topic of investigation. A MS is usually presented in an audio-format at some time during an interview (Leal et al, 2015). Interviewees typically underestimate how much information is required from them in an interview (Vrij, Hope, & Fisher, 2014). The detailed MS works as a social comparison (Cialdini, 2007; Festinger, 1954) and raises the expectations amongst both truth tellers and liars about how much information they should provide (Ewens et al., 2016). As a result, both truth tellers and liars provide more details after being exposed to a MS (Ewens et al., 2016; Leal et al., 2015). A MS may work better than an instruction to 'report all detail someone can remember' because a model statement is an example, and examples are easier to follow than instructions.

In all seven deception studies described by Vrij et al. (2018a) in their MS overview article, the MS facilitated the elicitation of information. However, the MS technique has generally been unsuccessful at enhancing differences in reporting details between truth tellers and liars, as both groups tend to provide a similar amount of additional details when exposed to a MS. However, a different finding emerged when examining the type of detail truth tellers and liars reported. In one study it was found that the difference in plausibility between the stories of truth tellers and liars was more pronounced in the MS-present condition than in the MS-absent condition (Leal et al., 2015), and in another study it was found that a MS enhanced the differences between truth tellers and liars in reporting complications (Vrij et al., 2017). Complications are the kind of detail that interviewees are likely to leave out when describing key aspects of activities as a story often makes perfect sense and is easily understood without reporting complications. Since liars either lack the creativity to make up complications or are reluctant to provide them to keep their story simple, the increase in complications as a result of a MS is more likely to occur in truth tellers than in liars. This results in more pronounced differences between truth tellers and liars in the MS-present condition than in the MS-absent condition.

Hypotheses

We tested, apart from the same eight hypotheses as in Experiment 1, the following hypotheses. Truth tellers will report more details than liars (Hypothesis 9). In terms of consistency, we predicted that, compared to liars, truth tellers will have more reminiscences (Hypothesis 10a) and more omissions (Hypothesis 10b) but fewer repetitions (Hypothesis 10c) and fewer contradictions (Hypothesis 10d).

In terms of verifiable details, we predicted a Veracity x Information Protocol interaction effect and expected that in Phase 2 truth tellers would report more verifiable sources than liars, particularly in the Information Protocol present condition (Hypothesis 11).

In terms of the Model Statement, we hypothesised that in Phase 2 interviewees will report more details (Hypothesis 12a) and will appear to elaborate more when telling their stories (Hypothesis 12b) in the MS-present than in the MS-absent condition (MS main effect). We further predicted that in Phase 2 truth tellers will report more complications than liars (Hypothesis 13a), and that their stories will sound more plausible (Hypothesis 13b), particularly in the MS-present condition (Veracity x MS interaction effect).

Procedure

Participants

A total of 208 passengers took part in the study, of whom 127 were male and 78 were female (three unknown). Their average age was $M = 42.59$ ($SD = 14.29$). There were 28 different nationalities in the sample and the vast majority of participants ($n = 189$) spoke English as their first language.

Procedure

The study took place at the same airport where Experiment 1 took place. Phase 1 of Experiment 2 was largely identical to Experiment 1. The only difference was that in Experiment 2 participants were encouraged to report details the investigator could check (by adding the words 'try to include details that we can check' to each question). Unlike Experiment 1, Experiment 2 also included a Phase 2. After the Phase 1 interview, the truth tellers ($n = 101$) and liars ($n = 107$) were brought to another counter where a different interviewer carried out the Phase 2 interview. Participants were randomly allocated to one of the three Phase 2 conditions. In the control condition ($n = 70$) the interviewer said "I'm going to have to ask you again to tell me about your trip" and then asked the same questions as in the Phase 1 interview and in the same order. In the Information Protocol condition ($n = 73$) the instruction was as follows: 'I'm going to have to ask you again to tell me about your trip. However, before you provide your statement, I want you to listen carefully to the following information. We know from research that liars prefer <u>*to avoid*</u> providing details that can be verified (i.e. details that can be checked) whereas truth tellers prefer to provide details that can be verified". After this, the four types of activities that constitute verifiable details were given.

The procedure in the Information Protocol plus Model Statement condition ($n = 65$) was almost identical to the procedure in the Information Protocol condition. Only one difference occurred at the beginning of the interview. The participants received the following instructions: 'I'm going to have to ask you again to tell me about your trip. Before I do so I would like you to read a model statement to give you an idea about how much detail I like you to include in your responses.' The model statement, a transcribed account of someone attending Formula 2 motor racing -as used in Leal et al. (2015)- was then given to the participant to read. After reading the model statement the interviewer introduced the Information Protocol in the same way as in the Information Protocol condition.

After Phase 1, the participants were asked to describe their experiences in a post-interview questionnaire. The questionnaire was almost the same as was used in Experiment 1. The only difference was that the questions were asked for each phase of the interview. After completing the questionnaire participants received a small present, were thanked and debriefed.

Coding

The two interviews (Phases 1 and 2) were audiotaped and videotaped. The same coders as in Experiment 1 were used and they were blind to the Veracity conditions. The coding for elaboration, complications, verifiable details, number of people mentioned and plausibility was almost identical to Experiment 1, except for verifiable details. They were coded as verifiable witness sources (activities 1 and 2) and verifiable digital sources (activities 3 and 4). In Phase 2, only the new (not reported in Phase 1) complications and verifiable details were coded.

In addition, the coder coded the number of details in each phase of the interview. The following statement contains 19 details: "I'm here to meet some friends, they live just outside New Delhi in a place called Rajinder Nagar. So I'll be meeting them and then we'll be going in a car to their home and have some quality time. I'll stay overnight and fly back home tomorrow night". In Phase 2, the coder split the details into repetitions (details mentioned in Phase 1 and in Phase 2), reminiscences (details mentioned in Phase 2 only) and contradictions. Based on the number of repetitions and contradictions, omissions (details mentioned in Phase 1 only) could be calculated. To give an example of a contradiction, someone said in Phase 1 "I have talked to my friends about trying to find a right time" and in Phase 2 "Firstly agreed with my friend on the timing of the trip". In the second sentence *friend* is a contradiction because it was *friends* (plural) before.

A second coder, also blind to the Veracity condition, coded these variables in 50 transcripts. ICCs (using the single measures scores) were good to satisfactory: elaboration (.82), unique verifiable witness sources (verifiable sources in Phase 1 and new verifiable sources in Phase 2 combined): 58; unique verifiable digital sources: .58; unique complications: .78; unique total of people mentioned: .94; Plausibility: .78: repeated verifiable witness sources in Phase 2 (sources also reported in Phase 1): 61; repeated verifiable digital sources in Phase 2: .60; repeated complications: .74; repeated people mentioned: .95.

In real time, immigration officers will not be able to code the four consistency variables objectively, as we did in the current experiment. We therefore also carried out a subjective coding. A coder coded the presence of repetitions etc. on 3 point scales ranging from [1] absent to [3] present. A second coder carried out the same coding on a sample of 50 transcripts. Both coders were blind to the Veracity condition. Intraclass Correlation Coefficients (ICCs) were poor for the four subjective consistency measures: Repetitions .10; omissions .23; reminiscences: .51 and contradictions (ICC could not be calculated because one coder did not notice any contradictions). The subjective consistency coding was subsequently dropped from the analyses.

Results

The Effect of Language

We ran two MANCOVAs with Veracity as factor and Language (English as first language or not) as covariate. The first MANCOVA included the seven verbal cues of Phase 1 (Table 4) and the second MANCOVA included the 11 verbal cues of Phase 2 (Table 5). In the Phase 1 analysis, the covariate was not significant at a multivariate level, $F(7, 196) = 0.51$, $p = .829$, $\eta_p^2 = .02$. In the Phase 2 analysis, the covariate was not significant at a multivariate level either, $F(10, 193) = 0.94$, $p = .496$, $\eta_p^2 = .05$. This is perhaps due to the small number of non-native English speakers in the sample (n = 19). This Language factor is therefore ignored in the hypotheses-testing analyses.

Strategies and Experiences

A MANOVA with Veracity (truth vs lie) and Type of Interview (control, Information Protocol or Information Protocol + MS) as factors was carried out on the questionnaire variables listed in Table 3.

Table 3 about here

At a multivariate level, only the Veracity effect was significant, $F(13, 190) = 50.69$, $p < .001$, $\eta_p^2 = .78$. The results showed that truth tellers and liars were both highly (around 5.65 on a 7-point scale) and equally motivated to do well, but that truth tellers were more motivated than liars to provide details in Phase 1. This latter finding supports Hypothesis 2 for Phase 1. Truth tellers also found it easier than liars to provide checkable details in Phases 1 and 2 (supporting Hypothesis 6). In addition, truth tellers were also more motivated than liars to provide verifiable details in Phases 1 and 2, whereas liars were more motivated than truth tellers to provide details someone cannot check in Phases 1 and 2 (supporting Hypothesis 5). Liars were more inclined than truth tellers to bluff in Phases 1 and 2 and liars' inclination to do so was considerable in both phases (mean average scores around 5.00 on a 7-point scale). Truth tellers reported that they overwhelmingly told the truth in both phases, whereas liars reported that they mostly lied in both phases, with about 30% of the information to be truthful. The Bayes Factor analyses revealed support for the null-hypothesis for the non-significant findings and substantial to very strong support for the alternative hypotheses for the significant findings, with an exception for details that an investigator cannot check. Only weak evidence was found for these variables, which means only weak support was obtained for Hypothesis 5.

The findings are virtually identical to those found in Experiment 1. The only difference is that in Experiment 2 truth tellers were more motivated than liars to provide details (in Phase 1), whereas no difference was found between truth tellers and liars in Experiment 1.

At a multivariate level, the Type of Interview main effect, $F(26, 378) = 1.16$, $p = .266$, $\eta_p^2 = .07$, and the Veracity X Type of Interview interaction effect, $F(26, 378) = 1.08$, $p = .366$, $\eta_p^2 = .07$, were not significant, suggesting that the manipulations in Phase 2 had no impact on the interviewees' strategies and experiences.

Phase 1

A one-way MANOVA regarding the Phase 1 data (first interview) was carried out with Veracity as the only factor and the seven variables presented in Table 4 as dependent variables. Type of Interview was not included as a factor because the different interview protocols were only introduced in Phase 2 of the interview.

Table 4 about here

The MANOVA revealed a significant multivariate Veracity main effect, $F(7, 200) = 9.74$, $p < .001$, $\eta_p^2 = .25$. At a univariate level, all Veracity effects were significant. Compared to liars, truth tellers appeared to elaborate more, and reported more details, witness sources, digital sources, complications and people. Truth tellers' statements also sounded more plausible than liars' statements. Table 4 shows that all effect sizes were medium to large, suggesting that differences between truth tellers and liars should be noticeable. The Bayes Factor analyses revealed very strong support for the alternative hypotheses. This supports Hypotheses 1, 3, 4, 7, 8 and 9 and replicates the findings of Experiment 1.

Phase 2

A MANOVA with Veracity (truth vs lie) and Type of Interview (control, Information Protocol or Information Protocol + MS) as factors was carried out on the Phase 2 data. The dependent variables were the eleven variables listed in Table 5.

Table 5 about here

At a multivariate level, both the Veracity main effect, $F(10, 193) = 5.91$, $p < .001$, $\eta_p^2 = .23$ and the Type of Interview main effect were significant, $F(20, 386) = 1.94$, $p = .01$, $\eta_p^2 = .09$. The univariate Veracity effects are presented in Table 5. Truth tellers reported more details than liars in Phase 2, supporting Hypothesis 9. Regarding consistency, compared to liars, truth tellers repeated more details that they already reported in Phase 1, rejecting Hypothesis 10c, but also omitted more details in Phase 2 that they did report in Phase 1, supporting Hypothesis 10b. No differences emerged for reminiscences and contradictions, rejecting Hypotheses 10a and 10d.

As in Phase 1, also in Phase 2, truth tellers elaborated more when telling their stories than liars (supporting Hypothesis 1) and reported more new verifiable witness sources than liars (supporting Hypothesis 3 for witness sources). Also, the statements of truth tellers sounded more plausible than those of liars (supporting Hypothesis 8). Unlike in Phase 1, the effects for digital sources, complications and number of people were not significant, rejecting Hypotheses 3 (for digital sources), 4 and 7. Table 5 shows that all significant effects, except for witness sources, were accompanied with effect sizes ranging from medium to large, suggesting that differences between truth tellers and liars should be noticeable. The Bayes Factor analyses revealed very strong support for the alternative hypotheses in case the effect was significant, except for witness sources, which only obtained weak support. The Bayes Factor analyses further revealed evidence for the null-hypotheses for the non-significant effects.

The univariate results for the Type of Interview main effect showed a clear pattern with all variables related to 'reporting details' being significant. That is, significant effects occurred for total details, $F(2, 202) = 5.66$, $p = .004$, $\eta_p^2 = .05$, new details, $F(2, 202) = 13.15$, $p < .001$, $\eta_p^2 = .12$, and elaboration, $F(2, 202) = 3.32$, $p = .038$, $\eta_p^2 = .03$. Regarding total details, Tukey post-hoc tests showed that in the Information Protocol + MS condition more details were provided ($M = 36.83$, $SD = 18.08$, 95% CI [32.72,41.33]) than in the control condition ($M = 27.00$, $SD = 21.94$, 95% CI [22.86,31.14]), $p = .004$, $d = .49$, 95% CI [.14, .82] which supports Hypothesis 12a. The difference between the control and Information Protocol condition ($M = 30.15$, $SD = 13.70$, 95% CI [26.22,34.34]), was not significant, $p = .553$, $d = .17$ 95% CI [-.16, .50]. The difference between the Information Protocol + MS condition and Information Protocol condition was not significant either, $p = .069$, $d = .42$, 95% CI [.08, .75].

Regarding new details, Tukey post-hoc tests showed that in the Information Protocol + MS condition more new details were provided ($M = 17.37$, $SD = 13.31$, 95% CI [14.88,19.95]) than in the control ($M = 8.43$, $SD = 10.29$, 95% CI [5.99,10.87]), $p < .001$, $d = .76$, 95% CI [.39, 1.09] and Information Protocol conditions ($M = 11.22$, $SD = 6.92$, 95% CI [8.87,13.65]), $p = .002$, $d = .59$, 95% CI [.24, .92]. This supports Hypothesis 12a. The difference between the control and the Information Protocol condition was not significant, $p = .243$, $d = .32$ 95% CI [-.02, .64].

Regarding elaboration, the pattern of results was the same as for total details. Tukey post-hoc tests showed that in the Information Protocol + MS condition, participants elaborated more when providing information ($M = 1.02$, $SD = 0.78$, 95% CI [0.86,1.19]) than in the control condition ($M = 0.73$, $SD = 0.74$, 95% CI [0.57,0.89]), $p = .042$, $d = .50$, 95% CI [-1.80,2.78] supporting Hypothesis 12b. The difference between the control and Information Protocol condition ($M = 0.92$, $SD = 0.70$, 95% CI [0.77,1.09]) was not significant, $p = .228$, $d = .17$ 95% CI [-.16, .50], neither was the difference between the Information Protocol and Information Protocol + MS conditions, $p = .682$, $d = .42$, 95% CI [.08, .75].

The multivariate Veracity X Type of Interview interaction effect was not significant, $F(28, 376) = 1.36$, $p = .108$, $\eta_p^2 = .09$. Since we predicted significant univariate interaction effects for verifiable details, complications and plausibility, we examined these effects, despite the multivariate effect not being significant. The interaction effects for verifiable details and plausibility were not significant, all $F's < 2.61$, all $p's > .112$, rejecting Hypotheses 11 and 13b. The interaction effect for new complications was significant, $F(2,202) = 3.70$, $p = .026$, $\eta_p^2 = .04$. However, simple effect tests showed that truth tellers and liars did not differ significantly from each other in reporting new complications in the control $F(1,68) = 1.94$, $p = .169$, $\eta_p^2 = .03$, $d = .31$, 95% CI [-.16,.78], Information Protocol, $F(1,71) = 1.67$, $p = .201$, $\eta_p^2 = .02$, $d = .30$, 95% CI [-.17,.75] and Information Protocol plus MS conditions, $F(1,63) = 3.63$, $p = .061$, $\eta_p^2 = .06$, $d = .46$, 95% CI [-.04,.95]. These findings reject Hypothesis 13a.

Plausibility Prediction

In Phases 1 and 2, plausibility emerged as a strong indicator of truthfulness. To examine which elements of a statement contributed to plausibility we carried out a linear regression analysis, using the 'enter' method. We ran two regression analyses, one to predict plausibility in Phase 1 based on the variables we examined in Phase 1 (see Table 4) and one to predict plausibility in Phase 2 based on the variables we examined in Phase 2 (see Table 5).

In the Phase 1 analysis, the model could explain 77% of the variance (r^2 = .59, F(6, 201) = 48.47, p < .001). Five significant predictors emerged, with elaboration being the strongest predictor, *beta* = .45, t = 7.34, p < .001, followed by verifiable digital sources, *beta* = .18, t = 3.34, p = .001, verifiable witness sources, *beta* = .13, t = 2.47, p = .014, complications, *beta* = .12, t = 2.23, p = .027, and total number of details, *beta* = .12, t = 1.99, p = .048. Only the number of people mentioned did not contribute to the model, *beta* = .06, t = 1.15, p = .252.

In the Phase 2 analysis, the model could explain 74% of the variance (r^2 = .55, F(13, 194) = 17.94, p < .001). Three significant predictors emerged, with, again, elaboration being the strongest predictor, *beta* = .45, t = 8.16, p < .001. This was followed by new complications, *beta* = .14, t = 2.61, p = .010, and people already mentioned in Phase 1, *beta* = .15, t = 2.10, p = .037.

Phases 1 and 2 combined

A MANOVA with Veracity (truth vs lie) and Type of Interview (control, Information Protocol or Information Protocol + MS) as factors was carried out on the unique details reported in Phases 1 and 2, that is Phase 1 details and the new details in Phase 2 combined. The dependent variables were the five variables listed in Table 6. At a multivariate level, the Veracity main effect was significant, F(5, 198) = 8.63, p < .001, η_p^2 = .18, whereas the Type of Interview main effect, F(10, 394) = 1.41, p = .170, η_p^2 = .04, and the Veracity X Type of Interview interaction effect, F(10, 394) = 1.24, p = .262, η_p^2 = .03, were not significant.

At a univariate level, all five Veracity effects were significant. Compared to liars, truth tellers provided more details, more verifiable witness and digital sources, more complications and more people who could be contacted. The effect sizes were somewhat small for complications but medium for the other variables. The Bayes Factor showed substantial to very strong support for the effects.

Correlational Analysis

A correlational analysis showed that, as expected, the variable elaboration correlated positively with the variable number of details in both Phase 1, r(208) = .573, p < .001 and Phase 2, r(208) = .759, p < .001.

Discussion

Phase 1: A Replication of Experiment 1

In Phase 1 we attempted to replicate the findings obtained in Experiment 1. We succeeded in this. Being able to replicate findings in a different, independent, sample is important as it shows the robustness of the findings. The effects we found varied from medium to strong in size, which means that they should be noticeable. This is a good achievement given that the interviews were short, which typically hampers verbal lie detection: Verbal cues to deception are more likely to occur when the interviews become longer.

From the seven variables examined in Phase 1, the most substantial differences between truth tellers and liars emerged in 'elaborating more when providing information' and 'plausibility', with plausibility being the strongest indicator of the two. This plausibility result replicated the findings of Experiment 1. Since plausibility is a subjective cue, we examined what constituted plausibility. We found that plausibility was largely determined by whether or not the interviewees appeared to elaborate when providing information.

The variable elaboration was measured in deception research for the first time and yielded large effect sizes and strong Bayes Factor results in both experiments. It was related to the number of details provided. We believe that these findings are encouraging. Number of details is amongst the strongest indicators of veracity in verbal deception research (Amado et al., 2015) and elaboration is probably easier to code in real life time than counting the frequency of details.

Phase 2: Consistency, Information Protocol and Model Statement

In Phase 2 we attempted to improve the ability to distinguish between truth tellers and liars in three different ways: (i) by asking the same questions as in Phase 1 to measure consistency in the answers, (ii) by informing interviewees through an Information Protocol what constitutes a verifiable source with the aim to enlarge the difference between truth tellers and liars in providing verifiable sources, and (iii) by providing a model statement with the aim to enlarge the difference between truth tellers and liars in reporting complications and providing plausible answers.

We need to dampen the expectations about how much improvement could be made. The differences between truth tellers and liars were quite pronounced in Phase 1, so not much improvement could be expected in the short Phase 2 interviews. Despite this, it is disappointing that no improvement was made. For example, examining the effect sizes shows that the results of Phases 1 and 2 combined (Table 6) were not better than the results of just Phase 1 (Table 4).

Regarding consistency, we expected truth tellers to provide more additional details than liars, but this did not happen. In fact, the number of additional details by truth tellers (M = 13.52) and liars (M = 10.95) were very similar. Truth tellers increased the number of new details compared to what they reported in Phase 1 by 30%, which is rather substantial and

perhaps as expected. Liars, however, increased the number of new details by 41%, which is very substantial and perhaps beyond expectations given that liars are concerned with consistency, and that additional details result in inconsistency. A possible explanation for liars' responses is that they were not much concerned with consistency given that they were interviewed in Phase 2 by a different interviewer than in Phase 1. Perhaps liars assumed that the two interviewers would not communicate with each other (e.g., Vrij, Leal, Mann, & Fisher, 2012). This would mean that different findings might emerge if they would be interviewed twice by the same interviewer.

Apart from not obtaining veracity effects for consistency, we did not succeed to code consistency subjectively (on a 3-point scale) in a reliable manner. This suggests clear individual differences between observers in what constitutes consistency (see also Granhag & Strömwall, 2001a, 2001b) and that measuring consistency reliably in a subjective coding system will be challenging. This has implications for measuring consistency in airport settings because objective coding is not possible due to time constraints. In other words, subjective coding of consistency is the only way to achieve consistency measures.

The Information Protocol was meant to give truth tellers better insight into what checkable sources are, so that they could search their memory better for possible verifiable information. Indeed, truth tellers reported in Phase 2 more new verifiable witness sources than liars, but this was across all conditions and not influenced by the Information Protocol. The most plausible explanation is that truth tellers already exhausted most possible verifiable information in the Phase 1 interview. This sounds reasonable because in Phase 1 interviewees were already asked to provide information the interviewer could check. Therefore, verifiable information was perhaps already on their minds and a different finding in Phase 2 would perhaps emerge if interviewees would not be instructed in Phase 1 to provide verifiable information. An alternative explanation is that interviewees did not fully comprehend the Information Protocol and thus did not fully understood what verifiable details are. A study examining interviewees' views about what in their view verifiable details are is in this respect welcome.

We made a distinction between verifiable witness sources and verifiable digital sources. The two variables were equally effective in distinguishing between truth tellers and liars in Phase 1 but witness sources was more effective than digital sources in Phase 2. Given the Phase 1 results, the best strategy for investigators probably is to take both variables into account.

The Model Statement was meant to increase the amount of details provided in Phase 2 and to enhance the differences between truth tellers and liars in complications and plausibility. It succeeded in eliciting more details only. An explanation as to why it did not enhance the differences between truth tellers and liars is that hardly any new complications were provided in Phase 2 (around 0.15 on average). This is not surprising given that interviewees only provided on average ten new details in Phase 2. The amount of new information was too little for a difference in complications to occur.

Despite the fact that the Information Protocol and Model Statement manipulations were not effective in enhancing differences between truth tellers and liars, the condition in which these two manipulations were combined resulted in most information. In other words, they had a positive effect on encouraging travellers to provide information. This is useful for border security because they can assist officers in their efforts to get travellers to talk more and reveal more information.

Strategies and Experiences

Interviewees' self-reported strategies and experiences were generally in alignment with our expectations. This explains why the predicted differences between truth tellers and liars emerged, particularly in Phase 1. The failure of Phase 2 to enhance the differences between truth tellers and liars could be explained –as we mentioned before– by having different interviewers carry out the interviews in Phases 1 and 2 (which could have affected consistency); by already asking in Phase 1 to provide verifiable information (exhausting the opportunity to provide verifiable information in Phase 2); and the failure to elicit enough new information in Phase 2.

Methodological issues

Four methodological issues in the two experiments merit attention. First, the majority of passengers we approached to take part declined, and we do not know how this has affected the results. Note, that a low response rate is not unusual in laboratory research. A typical experiment is carried out in a university setting where students are invited to take part through emails, flyers, announcements during lectures etc. Of course, only a small percentage of those who hear about the study will actually take part. In addition, we do not think that a low response rate matters on this occasion because the passengers who declined were hardly told anything about the study. In other words, they did not know what they actually declined to take part in. Our impression is that they had valid reasons to decline and that these reasons had nothing to do with the content of the study.

Second, participants were invited to prepare themselves but did not take up much time to do so. We do not know whether a more thorough preparation would have affected the results, but the effect of a more thorough preparation could be limited. People prepare answers to possible questions and people at airports overwhelmingly expect questions about the purpose of a trip only (Warmelink et al., 2012). Since only one question was related to the purpose of the trip, it is doubtful whether liars who prepare themselves more thoroughly would be more successful in answering the questions than the liars in the current experiment. In addition, some cues to deception may occur even with more thorough planning. For example, it should be difficult for thoroughly prepared liars to include truthful checkable details.

Third, the stakes (positive consequences for liars to remain undetected and negative consequences of being detected) were low in this experiment and much lower than they would be in real life. It is an empirical question how stakes would affect the interviewees' responses, but we believe that checkable details -already a strong diagnostic cue to deceit in

our experiment- could become a stronger diagnostic cue in real life interviews. A liar will probably become more reluctant to include false checkable details (e.g. bluffing) if i) s/he thinks that an investigator actually will check the details s/he provides and ii) the negative consequences are severe once the bluffs are detected. Both elements will be much stronger in real life interviews than in experimental studies, which means that bluffing will be less prominent in real life than in experimental studies. The lower the occurrence of bluffing, the more diagnostic the Verifiability Approach becomes.

Fourth, in future research it is important to examine the effect of talking in a first language on the results. We found not such an effect when we measured it in Experiment 2, but the number of participants who did not speak English as a first language was too small to properly examine this issue. A study that specifically examines the effect of first or second language speech on the results may be useful.

Fifth, participants were told that they had to pretend that they just landed at their destination. This means that truth tellers thus had to lie. We do not think it has influenced the results significantly, because the questions were not about this. That is, the questions were not about where the interviewees were at the moment of the interview, but they were about their activities at their destination. We included this instruction because interviewers took the role of immigration officers and felt more comfortable this way.

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Amado, B. G., Arce, R., Fariña, F. (2015). Undeutsch hypothesis and Criteria Based Content Analysis: A meta-analytic review. *The European Journal of Psychology Applied to Legal Context*, *7*, 3-12. Doi: 10.1016/j.ejpal.2014.11.002

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

Caldwell-Harris, C. L. (2015). Emotionality differences between a native and foreign language: Implications for everyday life. *Current Directions in Psychological Science*, *24*, 214–219. Doi: 10.1177/0963721414566268

Cialdini, R. B. (2007). *Influence: The psychology of persuasion*. New York, NY: William Morrow and Company.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (second edition)*. Hillsdale, NJ: Erlbaum.

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

Deeb, H., Vrij, A., Hope, L., Mann, S., Granhag, P. A., & Strömwall, L. (2018). Police officers' perceptions of statement inconsistency. *Criminal Justice and Behavior, 45,* 644-665. Doi: 0.1177/009385481875880

DePaulo, B. M., Lindsay, J. L., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin*, *129*, 74-118. Doi: 10.1037/0033-2909.129.1.74

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*, *30*, 854-862. Doi:10.1002/acp.3270

Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7,* 117-140. Doi:10.1177/001872675400700202

Fisher, R. P., Vrij, A., & Leins, D. A. (2013). Does testimonial inconsistency indicate memory inaccuracy and deception? Beliefs, empirical research and theory. In B. S. Cooper, D. Griesel, & M. Ternes (Eds.*) Applied issues in investigative interviewing, eyewitness memory, and credibility assessment* (pp. 173-190). New York: Springer.

Fiske, S. T., & Taylor, E. T. (1991). *Social cognition*. New York: McGraw-Hill.

Government Accountability Office (2017). *Aviation Security: TSA does not have valid evidence supporting most of the revised behavioral indicators used in its behavior detection activities.* Retrieved from http://www.gao.gov/assets/690/686001.pdf

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

Granhag, P. A., & 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. (2001a). Deception detection: Examining the consistency heuristic. In C. M. Breur, M. M. Kommer, J. F. Nijboer, & J. M. Reijntjes (Eds.), *New trends in criminal investigation and evidence, volume 2* (pp. 309-321). Antwerpen, Belgium: Intresentia.

Granhag, P. A., & Strömwall, L. A. (2001b). Deception detection: Interrogators' and observers' decoding of consecutive statements. *The Journal of Psychology, 135*, 603-620. Doi: 10.1080/00223980109603723

Hartwig, M., & Bond, C. F. (2014). Lie detection from multiple cues: A meta-analysis. *Applied Cognitive Psychology, 28*, 661-667. Doi: 10.1002/acp.3052.

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

Harvey, A., Vrij, A., Nahari, G,. & Ludwig, K. (2017). Applying the Verifiability Approach to insurance claims settings: Exploring the effect of the information protocol. *Legal and Criminological Psychology, 22*, 47-59. Doi: 10.1111/lcrp.12092.

Jarosz, A. F., & Wiley, J. (2014). What Are the Odds? A Practical Guide to Computing and Reporting Bayes Factors. *The Journal of Problem Solving, 7*(1). Doi: 10.7771/1932-6246.1167

Jeffreys, H. (1961). *Theory of Probability*. Oxford, UK: Oxford University Press.

Johnson, M. K. (2006). Memory and reality. *American Psychologist, 61*, 760-771. Doi: 10.1037/0003-066X.61.8.760.

Jupe, L., Leal, S., Vrij, A., & Nahari, G. (2017). Applying the verifiability approach in an international airport setting. *Psychology, Crime, & Law, 23*, 812-825. Doi:10.1080/1068316X.2017.1327584

Köhnken, G. (2004). Statement Validity Analysis and the 'detection of the truth'. In P. A. Granhag & L. A. Strömwall (Eds.), *Deception detection in forensic contexts* (pp. 41-63). Cambridge, England: Cambridge University Press.

Lakens, D. (2016, January 14). Power analysis for default Bayesian *t*-tests [Blog post]. Retrieved from http://daniellakens.blogspot.com/2016/01/power-analysis-for-default-bayesian-t.html

Leal, S., Vrij, A., Vernham, Z., Dalton, G., Jupe, L., Harvey, A., & Nahari, G. (2018). Cross-cultural verbal deception. *Legal and Criminological Psychology*, *23*, 192-213. Doi: 10.1111/lcrp.12131

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

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

Nahari, G. (2018). The applicability of the verifiability approach to the Real world. In J. P. Rosenfeld (Ed.), Detecting Concealed Information and Deception: Recent Developments (pp. 329 - 349). London: Elsevier. Doi: 10.1016/B978-0-12-812729-2.00014-8.

Nahari, G., & Nisin, Z. (2019). Digging deeper into the speech of liars: Future research prospects in verbal lie detection. *Frontiers in Psychiatry*, *10*, 56. Doi: 10.3389/fpsyt.2019.00056

Nahari, G., Vrij, A., & Fisher, R. P. (2014a). Exploiting liars' verbal strategies by examining the verifiability of details. *Legal and Criminological Psychology*, *19*, 227-239. Doi: 10.1111/j.2044-8333.2012.02069.x

Nahari, G., Vrij, A., & Fisher, R. P. (2014b). The Verifiability Approach: Countermeasures facilitate its ability to discriminate between truths and lies. *Applied Cognitive Psychology*, *28*, 122-128. Doi: 10.1002/acp.2974.

Strömwall, L. A., Granhag, P. A., & Hartwig, M. (2004). Practitioners' beliefs about deception. In P. A. Granhag & L. A. Strömwall (Eds.), *Deception detection in forensic contexts* (pp. 229-250). Cambridge, England: Cambridge University Press.

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, second edition*. Chichester: John Wiley and Sons.

Vrij, A., Fisher, R., Blank, H., Leal, S., & Mann, S. (2016). A cognitive approach to elicit nonverbal and verbal cues of deceit. In J. W. van Prooijen & P. A. M. van Lange (Eds.), *Cheating, corruption, and concealment: The roots of dishonest behavior* (pp. 284-310). Cambridge, England: Cambridge University Press.

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

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

Vrij, A., Leal, S., Fisher, R. P., Mann, S., Dalton, G. Jo, E., Shaboltas, A., Khaleeva, M., Granskaya, J., & Houston, K. (2018b). Sketching as a technique to elicit information and cues to deceit in interpreter-based interviews. *Journal of Applied Research in Memory and Cognition*, *7*, 303-313. Doi: 10.1016/j.jrarmac.2017.11.001

Vrij, A., Leal, S., Jupe, L., & Harvey, A. (2018c). Within-subjects verbal lie detection measures: A comparison between total detail and proportion of complications. *Legal and Criminological Psychology*, *23*, 265-279. Doi:10.1111/lcrp.12126

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

Vrij, A., Leal, S., Mann, S., & Fisher, R. (2012). Imposing cognitive load to elicit cues to deceit: Inducing the reverse order technique naturally. *Psychology, Crime, & Law*, *18*, 579-594. Doi: 10.1080/1068316X2010.515987

Vrij, A., Mann, S., Kristen, S., & Fisher, R. (2007). Cues to deception and ability to detect lies as a function of police interview styles. *Law and Human Behavior*, *31*, 499-518. Doi: 10.1007/s10979-006-9066-4

Vrij, A. & Nahari, G. (2019). The Verifiability Approach. In J. J. Dickinson, N. Schreiber Compo, R. N. Carol, B. L. Schwartz, & M. R. McCauley (eds.) *Evidence-Based Investigative Interviewing* (pp. 116-133). New York, U.S.A.: Routledge Press.

Warmelink, L., Vrij, A., Mann, S., Jundi, S., & Granhag, P. A. (2012). The effect of question expectedness and experience on lying about intentions. *Acta Psychologica*, *141*, 178-183. Doi: 10.1016/j.actpsy.2012.07.011

Wetzels, R., & Wagenmakers, E. J. (2012). A default Bayesian hypothesis test for correlations and partial correlations. *Psychonomic Bulletin & Review*, *19*, 1057-1064. Doi: 10.3758/s13423-012-0295-x

Table 1.

Results of the Post-Interview Questionnaire as a Function of Veracity: Experiment 1.

	Truth			Lie			*F*	*p*	Cohen's *d* [95% CI]	BF_{10}
	M	*SD*	[95% CI]	*M*	*SD*	[95% CI]				
Motivation to do well	5.85	0.79	[5.64, 5.62]	5.85	0.76	[5.62, 6.07]	0.01	.962	0.00 [-0.44, 0.44]	0.23
Motivation to provide many details	5.54	1.02	[5.23, 5.84]	5.44	0.91	[5.13, 5.75]	0.21	.645	0.10 [-0.34, 0.55]	0.26
Motivation to provide details investigator CAN check	5.71	1.01	[5.34, 6.08]	4.95	1.36	[4.57, 5.33]	8.13	.006	0.64 [0.18, 1.09]	7.20
Motivation to provide details investigator CANNOT check	4.05	1.70	[3.60, 4.50]	5.05	1.15	[4.59, 5.52]	9.45	.003	0.70 [0.23, 1.15]	12.33
Easiness of providing checkable details	5.44	1.31	[4.98, 5.90]	4.15	1.66	[3.68, 4.63]	14.87	<.001	0.87 [0.40, 1.33]	105.17
Extent of bluffing	1.42	0.99	[1.06, 1.79]	5.69	1.34	[5.32, 6.07]	263.72	<.001	3.65 [2.90, 4.35]	2.072e+23
Extent of telling the truth	98.05%	5.11	[91.96, 104.51]	32.05%	29.31	[25.43, 48.68]	201.62	<.001	3.14 [2.47, 3.80]	9.749e+19

Table 2.

Verbal Cues in the Interviews as a Function of Veracity: Experiment 1

	Truth			Lie			*F*	*p*	Cohen's *d* [95% CI]	BF_{10}
	M	*SD*	95% CI	*M*	*SD*	95%CI				
Elaboration	1.24	0.89	[1.00, 1.49]	0.67	0.66	[0.42, 0.92]	10.78	.002	0.73 [0.27, 1.19]	20.98
Checkable sources	1.51	0.95	[1.21, 1.81]	0.74	0.97	[0.44,1.05]	12.85	.001	0.80 [0.34, 1.26]	47.84
Complications	0.61	0.74	[0.43, 0.79]	0.18	0.39	[-0.01, 0.37]	10.50	.002	0.75 [0.27, 1.19]	18.79
Number of people mentioned	2.27	1.94	[1.74, 2.80]	1.46	1.41	[0.92, 2.00]	04.49	.037	0.48 [0.15, 0.79]	1.59
Plausibility	5.05	1.50	[4.57, 5.52]	2.88	1.55	[2.40, 3.37]	40.20	<.001	1.42 [0.92, 1.92]	714788.22

Table 3.

Results of the Post-Interview Questionnaire as a Function of Veracity: Experiment 2

	Truth			Lie			*F*	*p*	Cohen's *d*	BF_{10}
	M	*SD*	95% CI	*M*	*SD*	95% CI			[95% CI]	
Motivation to do well	5.68	0.97	[5.38, 5.88]	5.61	1.03	[5.41,5.80]	0.28	.600	0.07 [-0.20,0.34]	0.172
Motivation to provide many details in Phase 1	5.29	1.19	[5.04, 5.52]	4.85	1.23	[4.62,5.08]	6.62	.011	0.36 [0.08,0.63]	3.672
Motivation to provide many details in Phase 2	5.62	1.08	[5.39, 5.85]	5.35	1.24	[5.12,5.57]	2.73	.100	0.23 [-0.04,0.50]	0.535
Motivation to provide details investigator CAN check in Phase 1	5.35	1.24	[5.09, 5.59]	4.73	1.30	[4.48,4.97]	12.05	.001	0.49 [0.20,0.76]	41.653
Motivation to provide details investigator CAN check in Phase 2	5.61	1.13	[5.37, 5.86]	5.16	1.35	[4.91,5.39]	7.11	.008	0.34 [0.06,0.61]	3.605
Motivation to provide details investigator CANNOT check in Phase 1	3.85	1.65	[3.53, 4.12]	4.36	1.43	[4.08,4.65]	6.71	.010	0.33 [0.05,0.60]	2.246
Motivation to provide details investigator CANNOT check in Phase 2	3.86	1.70	[3.54, 4.14]	4.28	1.43	[3.99,4.57]	4.41	.037	0.27 [-0.01,0.54]	0.864

Easiness of providing checkable details in Phase 1	5.28	1.41	[4.98, 5.57]	4.56	1.60	[4.26,4.84]	11.82	.001	0.48 [0.19,0.75]	31.862
Easiness of providing checkable details in Phase 2	5.40	1.38	[5.09, 5.68]	4.64	1.62	[4.34 ,4.92]	13.22	< .001	0.50 [0.22,0.77]	66.342
Extent of bluffing in Phase 1	1.20	0.82	[00.92,1.47]	4.94	1.78	[4.66, 5.20]	367.83	< .001	2.67 [2.26, 3.00]	1.773e+44
Extent of bluffing in Phase 2	1.17	0.74	[00.90,1.43]	5.04	1.75	[4.77, 5.29]	418.58	< .001	2.85 [2.43, 3.19]	1.580e+49
Extent of telling the truth in Phase 1	97.13	11.52	[92.67,101.66]	31.31	30.13	[27.12, 35.95]	425.58	< .001	2.85 [2.43, 3.20]	1.491e+48
Extent of telling the truth in Phase 2	99.03	3.96	[83.96, 111.97]	28.96	29.93	[24.96, 33.40]	473.46	< .001	3.24 [2.78, 3.60]	1.430e+52

Table 4.

Verbal Cues in the Interviews as a Function of Veracity: Experiment 2, Phase 1

	Truth *M* (*SD*) (95% CI)	Lie *M* (*SD*) (95% CI)	*F*	*p*	Cohen's *d* (95% CI)	BF_{10}
Number of details	32.79 (18.32) (29.86,35.72)	23.04 (10.78) (20.19,25.88)	22.19	< .001	0.65 (0.36,0.92)	3400.012
Elaboration	00.95 (00.71) (00.83,01.07)	00.41 (00.53) (00.29,00.53)	38.59	< .001	0.87 (0.57,1.14)	3.364e+6
Verifiable witness sources	01.95 (01.75) (01.67,02.23)	01.33 (01.05) (01.05,01.60)	09.80	.002	0.43 (0.15,0.70)	13.969
Verifiable digital sources	00.59 (00.98) (00.45,00.74)	00.20 (00.46) (00.05,00.34)	14.21	< .001	0.51 (0.23,0.78)	101.327
Complications	00.45 (00.74) (00.33,00.56)	00.14 (00.37) (00.03,00.25)	12.58	< .001	0.53 (0.25,0.80)	48.915
Number of people mentioned	02.93 (01.97) (02.59,03.27)	01.64 (01.47) (01.32,01.98)	28.68	< .001	0.75 (0.45,1.01)	54805.053
Plausibility	03.91 (01.57) (03.64,04.18)	02.69 (01.14) (02.43,02.95)	41.54	< .001	0.89 (0.60,1.16)	1.114e+7

Table 5.

Verbal Cues in the Interviews as a Function of Veracity: Experiment 2, Phase 2

	Truth *M* (*SD*) (95% CI)	Lie *M* (*SD*) (95% CI)	*F*	*p*	Cohen's *d* (95% CI)	BF_{10}
Number of details	35.99 (20.81) (32.74,39.64)	26.64 (14.78) (23.33,30.03)	15.17	< .001	0.52 (0.24,0.79)	96.229
Repeated number of details	22.46 (13.12) (20.32,24.65)	15.68 (08.44) (13.55,17.75)	19.96	< .001	0.62 (0.33,0.89)	1244.644
Number of new details	13.52 (11.75) (11.67,15.74)	10.95 (10.07) (09.05,13.00)	03.47	.064	0.24 (-.04,0.50)	0.581
Number of omissions	10.33 (07.68) (09.11,11.65)	07.35 (05.01) (06.10,08.57)	11.51	.001	0.46 (0.18,0.73)	96.229
Number of contradictions	00.44 (00.81) (00.04,00.84)	00.99 (02.69) (00.61,01.38)	03.86	.051	0.27 (0.00,0.54)	0.964
Elaboration	01.17 (00.71) (01.04,01.31)	00.62 (00.68) (00.49,00.75)	34.04	< .001	0.79 (0.50,1.06)	301492.487
New verifiable witness sources	00.51 (01.05) (00.35,00.68)	00.22 (00.50) (00.07,00.38)	06.42	.012	0.36 (0.08,0.62)	2.591
New verifiable digital sources	00.32 (00.85) (00.18,00.46)	00.20 (00.57) (00.06,00.34)	01.50	.222	0.17 (-.11,0.44)	0.299
New complications	00.17 (00.48) (00.09,00.26)	00.17 (00.52) (00.08,00.27)	00.00	.952	0.00 (-.27,.0.27)	0.153
New number of people mentioned	00.56 (01.17) (00.37,00.76)	00.37 (00.80) (00.18,00.56)	01.88	.171	0.19 (-.08,0.46)	0.369

Plausibility	03.96 (01.70) (03.64,04.28)	02.66 (01.57) (02.36,02.98)	33.20	< .001	0.80 (0.50,1.07)	317438.976

Table 6.

Verbal Cues in the Interviews as a Function of Veracity: Experiment 2, Phases 1 and 2 combined

	Truth *M* (*SD*) (95% CI)	Lie *M* (*SD*) (95% CI)	*F*	*p*	Cohen's *d* (95% CI)	BF_{10}
Number of unique details	46.32 (24.95) (42.52,50.62)	33.99 (16.22) (30.08,37.94)	19.25	< .001	0.59 (0.30,0.86)	556.045
Number of unique verifiable witness sources	02.46 (02.00) (02.15,02.79)	01.55 (01.17) (01.25,01.87)	15.97	< .001	0.56 (0.27,0.83)	234.931
Number of unique verifiable digital sources	00.91 (01.43) (00.69,01.13)	00.39 (00.71) (00.18,00.61)	11.08	.001	0.46 (0.18,0.73)	25.816
Number of unique complications	00.62 (00.90) (00.47,00.77)	00.32 (00.80) (00.16,00.48)	6.59	.011	0.35 (0.07,0.62)	3.439
Number of people mentioned	03.50 (02.28) (03.11,03.91)	02.02 (01.82) (01.62,02.41)	27.28	< .001	0.72 (0.43,0.98)	24432.593