

Vrij Aldert (Orcid ID: 0000-0001-8647-7763)

Jupe Louise (Orcid ID: 0000-0002-1912-8579)

**Analysing openly recorded pre-interview deliberations  
to detect deceit in collective interviews**

Aldert Vrij

Louise Marie Jupe

Sharon Leal

Zarah Vernham

Department of Psychology, University of Portsmouth

Galit Nahari

Department of Criminology, Bar Ilan University

Author notes

Correspondence concerning this article 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](mailto:aldert.vrij@port.ac.uk)

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.3601

## Abstract

Sham marriages occur frequently and, to detect them, partners are sometimes interviewed together. We examined an innovative method to detect deceit in such interviews. Fifty-three pairs of interviewees, either friends (truth tellers) or pretended to be friends (liars), were interviewed about their friendship. Just before the interview, they received the questions that would be asked in the interview and were invited to prepare the answers. We told them that these pre-interview deliberations would be recorded. Based on the transcripts we analysed cues to truthfulness (cues expected to be expressed more by truth tellers) and cues to deceit (cues expected to be expressed more by liars). Truth tellers and liars differed from each other, particularly regarding expressing cues to truthfulness. Pre-interview deliberations that are recorded with awareness of the interviewees can be used for lie detection purposes. We discuss further venues in this new line of research.

Accepted Article

A sham marriage is a marriage purely for the purpose of gaining an advantage associated with that status. Common reasons for sham marriages are to gain immigration, residency, work or citizenship rights for one of the spouses. Sham marriages may occur frequently. The estimated numbers are between 4,000 and 10,000 in the UK each year (Mowat, 2016), with arrests rising from 137 in 2010 to 1,545 in 2014, an increase of more than 840% (Mowat, 2016). In the USA it is estimated that between 5% and 30% of all marriages between American residents and foreign nationals were fraudulent, but that very few of these marriages were detected or acted upon (Seminara, 2008). We are aware of countries where in an effort to detect sham marriages the couples are interviewed together, so-called collective interviewing (Vernham & Vrij, 2015). In the current experiment we tested a new method to detect deceit in such interviews, by giving the couples the opportunity to discuss the questions that will be asked during the interview, whilst they are aware that these pre-interview deliberations are recorded and observed.

Collective interviewing has advantages for lie detection. First, it enables interviewers to not only analyse the nonverbal and verbal responses of each interviewee (which could also be assessed when interviewing people separately), but also to analyse how these interviewees interact and communicate with each other (which only can be assessed in collective interviews). Second, interventions could be introduced that are only suitable in collective interviewing, which could facilitate truth telling or hamper lying. An example is forced turn-taking, whereby the interviewer asks one interviewee in the pair to start answering the question and then intervenes every 20 seconds by asking the other interviewee in the pair to continue from the point at which the partner was stopped (Vernham, Vrij, Mann, Leal, & Hillman, 2014). In such research, truth telling pairs were more likely to continue on from one another, whereas lying pairs were more likely to wait and repeat what their partner last said before continuing (Vernham et al., 2014b). In the current experiment, we examined a new

factor: Informing interviewees just before the interview which questions will be asked, allowing them to discuss their answers, and informing them that such discussions would be recorded. We examined verbal differences between pairs of truth tellers and pairs of liars in these pre-interview deliberations. Because this study was the first of its kind, we analysed many verbal cues. The risk of analysing a small set of cues is that some verbal cues would remain undetected.

We grouped the dependent variables into six clusters we thought may reveal differences between truth tellers and liars, based on theory and previous research. First, we examined memory search, because this is relevant to truthful and deceptive story telling (Sporer & Schwandt, 2016; Vrij, 2014; Walczyk, Igou, Dixon, & Tcholakian, 2013). Second, we examined coming up with alternatives, because this would occur naturally in truth tellers but would go against the strategies liars typically employ (Hartwig, Granhag, & Strömwall, 2007). Third, we examined cognitive load, because in story telling lying is considered to be more mentally taxing than truth telling (Sporer, 2016; Vrij, Fisher, Mann, & Leal, 2006; Walczyk, Roper, Seemann, & Humphrey, 2003). Fourth, we examined specific strategies and impression management cues, thought to be prevalent in liars (Granhag & Hartwig, 2008; Hartwig, Granhag, & Strömwall, 2007). Fifth, we measured consistency because this is a cue likely to be at the forefront of liars' minds (Fisher, Vrij, & Leins, 2013; Vredeveldt, van Koppen, & Granhag, 2014). Finally, we examined giveaway cues of truth telling and lying which may occur naturally in the pairs' discussions.

### **Memory search**

When individuals collectively attempt to retrieve information from their shared memories, they typically communicate considerably with one another to retrieve as much information as possible. According to Wegner (1987), groups develop 'transactive memory systems' for retrieving information. Hollingshead (1998) refers to a transaction memory

search as whereby group members with shared memories attempt to increase recall by posing questions to each other to check or find out information and by cuing each other for further information. In a previous deception study it has been found that this pattern of transaction memory search occurred more frequently amongst truth tellers than liars (Vernham, Vrij, Leal, & Mann, 2014). In addition, memory of real events are embedded in location and time (Johnson, 1988, 1996) and research has shown that truth tellers provide more temporal and spatial information than liars (Masip, Sporer, Garrido, & Herrero, 2005; Sporer, 2004).

### **Coming up with alternatives**

In their memory search of joint experiences, truth tellers may come up with more than one experience that would answer a particular question they are asked to prepare an answer for. For example, one person may remember something the other person initially did not think of or one person may believe they have a better example than the other person initially provides. Although liars may go through a similar process while thinking of answers, previous collective interviewing research suggests that this is less likely to happen (Vrij et al., 2009). Liars' dominant strategy during the actual interviews was that one person comes up with an idea and the other person just accepts and agrees (Vrij et al., 2009). We examined whether a similar pattern emerges in pre-interview deliberations.

### **Cognitive load**

The task for truth tellers is probably easier than for liars. Truth tellers can openly consult their memories, discuss their experiences, and select the best memories for their answers. Liars in the current experiment do not have memories to share and therefore have to fabricate their answers, which is typically a more difficult task than coming up with true answers (Parkhouse & Ormerod, 2018). Fabricating answers is made more difficult for liars in the current experiment because they need to keep up the premise that they know each other and thus cannot openly polish their fabrications. In their pre-interview deliberations, pairs

may comment on the difficulty of the task at hand. If they do so, the difference in cognitive load between truth tellers and liars may become transparent in the comments they make. We examined this.

### **Strategies and Impression Management**

When liars prepare themselves for interviews, they typically prepare nonverbal and verbal strategies to come across convincing (Hartwig, Granhag, & Strömwall, 2007). Examples of nonverbal strategies are planning to avoid showing behaviours they think appear nervous, such as gaze aversion or fidgeting, whereas examples of verbal strategies are thinking of details to report that sound convincing (Colwell, Hiscock-Anisman, Memon, Woods, & Michlik, 2006; Hartwig et al., 2007; Hartwig, Granhag, Strömwall, & Doering, 2010). Truth tellers are typically less inclined to prepare for interviews; they often indicate this not to be needed, because they just will tell the truth (Hartwig et al., 2007; Vrij, Mann, Leal, & Granhag, 2010). Planning could be beneficial for liars. Prison inmates believed that planning enhanced their performance in interviews (Granhag, Andersson, Strömwall, & Hartwig, 2004). In addition, two meta-analyses revealed that planning did improve the quality of speech (DePaulo et al., 2003; Sporer & Schwandt, 2006; see Chan & Bull [2014] for a more recent study). Finally, another meta-analysis showed that unplanned lies were more frequently detected than planned lies (Bond & DePaulo, 2006). We listened for comments in the interview that revealed the planning of nonverbal and verbal strategies.

Whilst truth tellers typically focus on their story and try to tell it as it happened, liars do not take their credibility for granted when telling their stories (DePaulo et al., 2003; Gilovich, Savitsky, & Medvec, 1998), and are more concerned than truth tellers about how to appear convincing (Vrij et al., 2008). This could mean that liars – more than truth tellers – discuss in advance whether or not a possible response would sound convincing. We examined whether such discussions took place.

## **Consistency**

Compared to truth tellers, liars are more concerned with consistency (Fisher, Vrij, & Leins, 2013; Strömwall, Granhag, & Hartwig, 2004; Vredeveldt, van Koppen, & Granhag, 2014). Consistency is amongst the strongest beliefs about deception and people typically think that inconsistency constitutes lying (Strömwall et al., 2004). Liars therefore attempt to be consistent when interviewed about the same event multiple times (Vredeveldt et al., 2014), which expresses itself, amongst other cues, in the inclination to avoid omissions and reminiscences (adding new information) (Fisher, Leins, & Vrij, 2013). Although in the present interview, interviewees were only interviewed once, this inclination to be consistent may also become apparent when comparing what the interviewees agreed in the pre-interview deliberations that they would say in the interview with what they actually said in the interview.

## **Giveaway cues**

In the pre-interview deliberations truth tellers and liars may implicitly make clear that they do or do not know each other. For example, truth tellers may refer to shared knowledge without making it explicit (e.g. PP1: “I am worried about John” PP2: “Yes, he never got over splitting up with Jane”), whereas liars may ask each other questions which reveals that they do not know each other well (“What do you study?”). Another possibility is that liars start openly fabricating possible answers to questions (“Let’s say that...”). The ‘lying’ examples may indicate that liars did not follow the instructions well. Indeed, such cues occurred in the group of liars that we discuss in the supplementary materials (see below). However, the truth telling example would be in alignment with the instructions. We investigated the occurrence of giveaway cues in the pre-interview deliberations.

## Hypotheses

We tested the following hypotheses:

Truth tellers will show more signs of memory search than liars (Hypothesis 1)

Truth tellers will come up with more alternatives than liars (Hypothesis 2)

Truth tellers will report that they found the task easier than liars (Hypothesis 3)

Liars will discuss more strategies than truth tellers (Hypothesis 4)

Liars will discuss more impression management issues than truth tellers (Hypothesis 5)

Liars will show more consistency than truth tellers between their pre-discussion answers and actual interview answers (Hypothesis 6)

Liars and truth tellers will show giveaway cues that will reveal their veracity status (Hypothesis 7)

## Method

### Participants

Participants were recruited via a participant data base, online advertising using both student and staff portals and via word of mouth. Participants were either recruited as pairs, with the requirement that they had known each other socially for a period of longer than 12 months, or as individuals and placed with a partner on the day of the experiment.

A total of 154 participants (77 pairs) took part in the experiment, comprising of 39 males (25.3%) and 112 females (72.7%). Three individuals did not provide their gender. The mean age of participants was 22.53 years ( $SD = 6.37$ ). Truth telling pairs ( $n = 26$ ) were those who had known each other socially for more than 12 months, whilst lying pairs ( $n = 51$ ) were those who were placed with a partner (unknown to them) on the day. The 51 lying pairs were comprised of two groups; pairs who were just told that their discussion would be recorded ( $n = 24$ ) and pairs who were told that apart from their interviews being recorded, someone was



watching their preparation discussions through a one-way mirror ( $n = 27$ ). We introduced this second group of liars after data collection had started, because after observing a few pre-interview deliberations with liars, we noticed that some of them gave giveaway cues or openly fabricated answers. We can only speculate why this happened, see the Discussion section. Since the manipulation for the first group of liars was probably not strong enough, we do not discuss their results, but they are available in the supplementary materials. A comparison of Tables 2 and 3 shows that more cues to deceit occurred in the truth tellers - first group of liars comparison (Table 3) than in the truth tellers - second group of liars condition (Table 2), particularly amongst verbal strategies, giveaway cues and openly fabricating answers. The final sample consisted of 106 participants (53 pairs), comprising of 30 males (28.3%) and 74 females (69.8%). Two individuals did not provide their gender. The mean age of participants was 22.41 years ( $SD = 5.61$ ).

### **Procedure**

Upon arrival at the Psychology Department, participants were informed that they would be interviewed about their friendship; lying pairs were told that they would need to feign this friendship. No further instructions were given in relation to the nature of the interview. Participants were not informed whether their interview would be together or separate. If participants asked the experimenter, she replied that she did not know. To motivate participants to perform well during the experiment, they were informed that if the interviewer believed their account, they would be entered into the prize draw to win £150 (£75 each). However, if the interviewer did not believe their account (i.e., did not believe that they were true friends), they would not be added to the prize draw and would be required to write a secondary statement pertaining to their friendship for the interviewer.

All participants were then given time to prepare for this interview about their alleged friendship by having coffee at a local café. Each pair was given £10 for coffee and asked to

return to the department after 30 minutes. First, they were then asked to complete a pre-interview questionnaire, which they were asked to complete truthfully, irrespective of their veracity status. The pre-interview questionnaire collected simple demographics such as age, gender and student/staff status. Participants were also asked to rate their level of motivation to perform well during the interview (ranging from [1] extremely unmotivated to [7] extremely motivated). Truth tellers, but not liars, were asked for how long they knew each other (in months) and how much time the pairs spent together, on average, per month (type of response ranged from hours to days, and hours were divided by 8 to provide overall 'days' spent together). They were also asked to rate their friendship on 7- point Likert scales for how well they knew one another (ranging from [1] strangers to [7] best friends), closeness (ranging from [1] distant to [7] intimate), level of importance (ranging from [1] unimportant to [7] important), and level of trust (ranging from [1] distrusting to [7] trusting). The answers were clustered into one variable, labelled friendship strength (Cronbach's alpha = .825). For truth tellers, the three variables length of friendship, overall days spent together, and friendship strength were then correlated with all 21 dependent variables listed in Table 2. Only four out of these 63 correlations (6% of the variables) were significant, which could be expected by chance alone. We can thus conclude that length of friendship, overall days spent together, and friendship strength had no effect on the dependent variables.

Participants were then informed that they would be given the questions that would be asked by the interviewer and that they had ten minutes to prepare. They were told that this preparation time would be audio and video recorded (truth tellers and initial group of liars whose data are not reported) or they were told that this preparation time would be audio and video recorded and that the experimenter would also be watching their preparation time through a one-way mirror (liars whose data are reported). Audio and video recording and observing through a one-way mirror all actually took place. The four questions provided were

the exact questions that the interviewer would ask during the interview and describe a wide range of issues that we thought real friends should be able to talk about in some detail: (1) “Please can you describe how you became friends?”; (2) “Please can you describe the funniest event that has happened when you have been together?”; (3) “Please describe the most stupid thing you have ever witnessed your friend saying or doing?”, and (4) “Please describe the most recent activity you carried out together before today?”

The participants were given ten minutes for these audio and video recorded pre-interview deliberations. After ten minutes, an interviewer, blind to the veracity status of the interviewees entered the room, introduced him/herself and then asked the four questions the participants had prepared to answer. When the interview was over, the interviewer thanked the participants and left the room. The audio and video recording equipment were then turned off.

Following the interview, participants were asked to complete a post-interview questionnaire. Again, they were reminded to answer truthfully, irrespective of their veracity status. In this questionnaire, participants were asked to rate the extent to which they told the truth in the interview (as a percentage) on an 11- point Likert scale ranging from [0] 0% to [10] 100%. In addition, they were asked to rate on 7- point Likert scales, how likely they thought they (1) would be entered into the prize draw and (2) would have to provide a written statement (ranging from [1] not at all to [7] very likely). We also asked participants the extent to which they anticipated that (1) their ten minutes preparation time was going to be audio and video recorded and (2) that they were interviewed as a pair (ranging from [1] anticipated to [7] unanticipated). Note that these two questions refer to their anticipation that their conversations would be monitored and collective interviewing would take place, not to the fact that this monitoring and collective interviewing actually took place after being informed

about it. Finally, participants were asked how difficult they found it being interviewed as a pair (ranging from [1] easy to [7] difficult).

When the post-interview was complete, participants were debriefed about the nature of the study and also paid £10 for their participation. All participants were told that their account had been believed by the interviewer and that they would be entered into the prize draw.

### **Coding**

One coder coded all the 77 interviews, whereas a second coder coded a random sample of 20 interviews. Both coders were blind to the Veracity status and hypotheses. The first coder coded the first ten interviews and discussed her coding with the first author. The first coder then coded the remaining 67 interviews and a second coder the random sample of 20 interviews. The presence or absence of a verbal cue was recorded for each question and the results of the four questions were averaged. Each score could thus vary from 0 (never present) to 4 (present in all questions). The exception to this was the joint experience variable, for which the number of reported joint experiences were counted. Table 2 provides an overview of the variables that were coded.

For the Memory variables, we coded whether a participant reported an experience in terms of time (e.g. “Remember when you slipped in the snow last Christmas”) or location (e.g., “Remember when we walked to Fratton”), provided cues to each other (e.g., PP1: “We went to the library”, PP2: “True!, on Friday”) and finished each other’s sentence (e.g., PP1: “We had six pints each”, PP2” “Six pints of beer”). The Cambridge English dictionary defines the word remember as ‘to be able to bring back a piece of information into your mind’. Since this activity should reflect the act of truth tellers more than liars, we examined whether truth tellers also actually said the word ‘remember’ more than liars.

For the Alternatives variables, we coded whether more than one answer alternative was discussed between the pair. For questions 2 and 3, we also coded how many different experiences the participants reported.

For the Cognitive load variables, we coded whether participants explicitly mentioned that a question would be difficult or easy to answer. For Impression Management, we coded whether they discussed whether a given answer would sound credible (E.g., “If you say that, it is going to confuse the interviewer”) or embarrassing (E.g., “It makes us sound like such children”). We also coded swearing and making jokes (E.g., PP1: “What’s the most stupid thing you’ve ever witnessed me doing?” PP2: “That’s like every day”). For Strategies, we made a distinction between verbal strategies (E.g., “Maybe we should give lots of details when answering the questions”) and nonverbal strategies (E.g., “Make sure we are smiling and laughing to make it real”).

For the Consistency variables, we compared what the participants in the pre-interview deliberations agreed to say in the interview with what they actually said during the interview. Elaboration within a topic means that they discussed what to say in the pre-interview deliberations but said more about it in the actual interview. Elaboration new means that they said something in the interview they did not discuss in the pre-interview deliberations. Omission means that they discussed something in the pre-interview deliberations but left out of the interview.

For Giveaway cues, we coded whether they said something that reveals that they knew each other (E.g., “But you always say that I cannot talk about that”) or that they did not know each other (E.g., “My heater, I live in student halls, my heater doesn’t really work”).

To measure inter-rater reliability between the two coders, the two-way random effects model measuring consistency was used (Single Measures ICCs). They are reported in the

second column of Table 2 and revealed at least satisfactory reliability for each of the variables.

## Results

### Questionnaire results

Seven ANOVAs were carried out with Veracity as the only factor and the seven variables presented in Table 1 as dependent variables.

Truth tellers rated their truthfulness as significantly higher than liars, thought it would be more likely they would be entered in the prize draws and thought it less likely that they would have to write a statement. This indicates that the manipulation check was successful.

Truth tellers also anticipated less than liars that the pre-interview deliberations would be recorded and found it easier than liars to be interviewed together.

No Veracity differences emerged for motivation and anticipation of being interviewed together. The average mean scores indicated that participants were highly motivated and somewhat surprised to be interviewed together.

### Hypotheses-testing

A MANOVA was carried out with Veracity as factor and the 21 variables listed in Table 2 as dependent variables. We also report the Cohen  $d$ 's and Bayes Factors,  $BF_{10}$ . A  $BF_{10}$  smaller than 1 indicates evidence for the absence of an effect (support of the null hypothesis);  $BF$ s between 1 and 3 suggest weak evidence,  $BF$ s between 3 and 10 suggest strong evidence and  $BF$ s  $> 10$  very strong evidence for the alternative hypothesis (Jeffreys, 1961). We used the default Cauchy's prior of .707 for the Bayesian  $t$ -tests (Lakens, 2016).

The multivariate effect was significant,  $F(20,32) = 3.12, p = .002, \eta_p^2 = .66^i$ . At a univariate level, six significant differences emerged and they are indicated with '\*' in Table 2. Liars, compared to truth tellers, provided more cues to each other. Truth tellers, compared to liars, (i) finished each other sentence more often, (ii) discussed more than one alternative per

question, (iii) discussed more experiences in the pre-interview deliberations, (iv) expressed more frequently that reporting a specific event would be embarrassing, and (v) gave more cues that they knew each other. This means that Hypotheses 1, 3, 4 and 6 were rejected, that we found limited support for Hypothesis 5, and 7 and that we found strong support for Hypothesis 2.

The Bayes Factor analyses showed weak evidence for the finishing sentences and embarrassment effects but strong to very strong evidence for the remaining four effects. The Bayes Factors further showed mostly absence of evidence for the non-significant results.

### **Discussion**

In the current experiment, truth tellers' and liars' pre-interview deliberations were recorded and analysed. We examined cues to truthfulness (cues expected to be expressed more by truth tellers) and cues to deceit (cues expected to be expressed more by liars). Fewer cues to deceit (one) than cues to truthfulness (five) emerged in the experiment. In other words, liars were able to avoid showing cues to deceit, but could not fully replicate the verbal cues truth tellers did report. This reflects verbal lie detection in general, where in research many more cues to truthfulness have emerged than cues to deceit. For example, Criteria-Based Content Analysis -a widely researched verbal veracity assessment tool- contains 19 criteria, all of them cues to truthfulness (Amado, Arce, Fariña, & Vilarino, 2016). In addition, Reality Monitoring -another widely researched verbal veracity assessment tool- contains only one cue to deceit, cognitive operations, but that cue does not seem to discriminate truth tellers from liars (Masip, Sporer, Garrido, & Herrero, 2005). The absence of verbal cues to deceit is unfortunate because a mixture of cues to truthfulness and deceit creates the opportunity to create within-subjects measures, e.g., the proportion of cues to truthfulness (Vrij, Leal, Jupe, & Harvey, 2018). Within-subjects measures are desirable (Vrij, 2016, 2019) because they can control for individual differences, such as how talkative an interviewee is (Nahari & Pazuelo,

2015; Nahari & Vrij, 2014). Within-subjects comparisons may also help investigators to make judgements in individual cases, when cut-off scores can be established (Vrij, 2016).

Unsurprisingly, researchers recently pleaded for the search of verbal cues to deceit (Nahari et al., 2019).

The only cue to deceit we found -liars providing more cues to each other than truth tellers- was unexpected as we predicted the opposite to occur. In general, we expected memory cues to occur more often in discussions amongst truth tellers than amongst liars –as was found before in deception research (Masip et al., 2005; Vernham et al., 2014)- but the findings were erratic. The only support we found was a weak effect that truth tellers finished each other sentences more than liars. Perhaps differences between truth tellers and liars may emerge if different variables are taken into account, such as repeating or summarising each other (Vredeveldt, 2018). In addition, since collaborative remembering is facilitated through effective communication styles between the group members (e.g., elaborations, explanations, repetitions, reinstatements, Harris, Keil, Sutton, Barnier, & McILwain, 2011; Meade, Nokes, & Morrow, 2009), encouraging pairs to use such strategies may reveal differences between truth tellers and liars. Although Vredeveldt (2018) found no beneficial effect on memory recall of instructing truth telling partners to repeat each other or to elaborate upon each other's statements, this does not rule out that such an instruction may elicit differences between truth tellers and liars.

Two cues in particular revealed very strong Veracity effects: In the pre-discussion truth tellers (i) discussed more alternatives per question, and (ii) discussed more experiences. Unlike liars, who had to fabricate an answer, truth tellers could tap into their memories and they apparently showed that in the pre-discussion. Thus, rather than the way how they discussed their alleged memories (the memory variables), it was evidence of the richness of these alleged memories that discriminated truth tellers from liars. Thus, interview protocols



that focus on the richness of the to-be-discussed memories may be particularly successful in distinguishing truth tellers from liars. It may hereby be important not to prompt the pairs to recall as many memories as possible as this may give hints to liars what the interviewer is looking for, which could diminish the effect. That is, open-ended requests “Tell me all you remember” are more effective for lie detection purposes than specific requests “Describe the man you saw”, because to the latter specific requests both truth tellers and liars may provide similarly detailed answers (Vrij, Leal, Mann, Vernham, & Brankaert, 2015).

Two more cues to truthfulness occurred: Truth tellers expressed more frequently than liars that reporting a specific event would be embarrassing, and gave more cues to each other that they knew each other. The ‘embarrassment cue’ effect was weak, which means that replication of the effect is required before any conclusions can be drawn. The ‘know each other cues’ effect was stronger and it is worth examining whether this effect becomes stronger when using interview protocols focusing on the richness of alleged memories.

Many cues were not related to deception. For example, comments about how easy or difficult the task was did not distinguish truth tellers from both groups of liars, although the group liars, more than truth tellers, reported in the post-interview questionnaire that they found the interview setting difficult. In other words, liars may have experienced more cognitive load than truth tellers, but did not say so in the pre-interview deliberations. In addition, liars typically take their credibility less for granted than truth tellers (DePaulo et al., 2013; Gilovitch et al., 1998), but liars did not discuss more than truth tellers how to phrase answers and how credible such answers would sound. Also, liars typically think that their nonverbal behaviour gives away that they are lying, but liars did not discuss their nonverbal strategies more than truth tellers. Despite the absence of effects, we still believe that liars were more concerned with impression management than truth tellers. We believe that they did not reveal it because they were aware that their discussions were recorded and that such

discussions would be a giveaway sign of their deception. Finally, the consistency findings did not reveal differences between truth tellers and liars. Although this goes against the widespread belief that truth tellers are more consistent than liars (Strömwall et al., 2014) it is in alignment with reviews that showed that truth tellers and liars are often equally consistent [Fisher et al., 2013; Vredeveldt et al., 2014]).

● We initially told liars only that their pre-interview deliberations would be recorded. This resulted in many differences between truth tellers and liars particularly amongst verbal strategies, giveaway cues and openly fabricating answers (see Table 3 in the Supplementary Materials). We can only speculate why this happened. One reason is that liars thought that the recordings would not be analysed until after the interview. Such a thought would have been reasonable given that liars were aware that the interview would take place directly after the pre-interview deliberations. They may thus have thought that it may not be too damaging for them to give away their veracity status in the pre-interview deliberations.

When we told liars that their pre-interview deliberations would be observed also through a one-way mirror the obvious cues to deception were no longer present. In that scenario, only relatively few cues we examined (six out of 21) discriminated truth tellers from liars. This took us by surprise. It turned out to be less effective than we initially thought. This underlines the common finding in deception research that differences between truth tellers and liars are typically subtle (DePaulo et al., 2003) and difficult to spot (Bond & DePaulo, 2006). In hindsight, it provides additional justification to our decision to examine so many cues, because if we would have focused on a specific smaller set of cues, we may have found fewer differences or no differences at all. We are aware that obtaining relatively few significant effects and having relatively many hypotheses unsupported, may raise doubts about the validity of the findings. Such doubts may be further strengthened because several cues we examined lacked a strong theoretical foundation and were selected based on previous

research. We thus believe that replication is required before recommendations can be made to practitioners about introducing the tested manipulation in the field.

### **Methodological Considerations**

Four methodological considerations are worth discussing. First, in the current experiment, we had to strengthen the instruction provided to liars by informing them that the pre-interview deliberations would also be observed through a one-way mirror. As a result, we gave thus different instructions to liars and truth tellers, creating a confound. The results need to be replicated in an experiment without such a confound. Second, we could have chosen other categories of question, such as questions about personality traits. Third, in real life being caught out in a sham marriage could result in serious consequences, including deportation, whereas our experimental situation was a low-stakes scenario. At present, we have no convincing theoretical argument as to why a high-stakes situation would strengthen or weaken the results or leave them unchanged. This thus remains an empirical question.

Fourth, liars in our experiment could only briefly prepare themselves (they were sent to a local café and asked to return to the department after 30 minutes) and they probably have more preparation time in real life. A longer preparation time only benefits liars if their preparations are related to the questions asked in the interview (Vrij, Fisher, & Blank, 2017). Investigators can ask questions about all sorts of areas including questions that liars in all likelihood have not anticipated. Fifth, participants in the experiment self-selected to participate either as pairs of friends or as individuals. This creates, in theory, the possibility of group non-equivalence. This friend –no friend confound should be avoided in a future replication study.

### **Future Research**

This is the first experiment in which pre-interview deliberations between pairs of truth tellers and liars were recorded and coded. We believe that the findings show potential for use

in interviews aimed at detecting sham marriages, although there is a need for the present set of findings to be replicated, as discussed above. We therefore encourage researchers to carry out such a replication. We also hope that this experiment will make researchers enthusiastic about this new domain of research and make them to decide carrying out their own research in this area. To stimulate such research, we give a wide range of possible research venues.

First, perhaps interview protocols could be designed further highlighting the differences in richness of memory between truth tellers and liars. Second, through specific instructions perhaps differences in how truth tellers and liars discuss their alleged memories could be elicited. Third, it would be useful to search for cues that liars display, so that within-subjects comparisons can be made (e.g., proportion of cues to truthfulness), which may help investigators in making decisions in individual cases. Fourth, in the current experiment we focused on the pre-interview deliberations and on the consistency between these discussions and the answers given in the actual interview. A logical next step would be to broaden the horizon and examine the effect of such pre-interview deliberations on the quality of the actual interviews. Fifth, in the current experiment, an experimenter was watching the pre-interview deliberations in the liars condition. In future research, the actual interviewer could watch the pre-interview deliberations, which we think would increase pressure on liars. Sixth, future research could examine the effectiveness of interview questions other than about experiences, for example about personality traits. Finally, perhaps in future research a control group could be included in the design. One such a control group could be truth-tellers and liars who are asked to prepare themselves in recorded pre-interview deliberations but without having the questions given to them. Another control group could be truth tellers and liars whose interviews are secretly recorded. This probably would lead to two groups of liars, naïve liars who would have no idea that they would be recorded and more suspicious liars who may have thought that recordings took place. We expect the first naïve group of liars to give

obvious cues to deception, similar or even stronger cues than the group of liars in the present experiment whose results are discussed in the supplementary materials. The more suspicious group of liars are in our opinion more likely to show responses similar to the group of liars whose results are discussed in the main text.

#### **Acknowledgement**

This work was funded by the Centre for Research and Evidence on Security Threats (ESRC Award: ES/N009614/1)

#### **Availability of data**

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

#### **Conflict of interest**

The authors have no conflict of interest to declare.

Accepted Article

## References

- 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.org/10.1016/j.ijchp.2016.01.002.
- Bond, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgements. *Personality and Social Psychology Review, 10*, 214-234. Doi: 10.1207/s15327957pspr1003\_2
- Chan, S., & Bull, R. (2014). The effect of co-offender planning on verbal deception. *Psychiatry, Psychology, and Law, 23*, 457-464. Doi: 10.1080/13218719.2013.835703
- Colwell, K., Hiscock-Anisman, C., Memon, A., Woods, D., & Michlik, P. M. (2006). Strategies of impression management among deceivers and truth tellers: How liars attempt to convince. *American Journal of Forensic Psychology, 24*, 31-38.
- 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
- 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)*. Springer: New York.
- Gilovich, T., Savitsky, K., & Medvec, V. H. (1998). The illusion of transparency: Biased assessments of others' ability to read one's emotional states. *Journal of Personality and Social Psychology, 75*, 332-346. Doi: 10.1037/0022-3514.75.2.332
- Granhag, P. A., Andersson, L. O., Strömwall, L. A., & Hartwig, M. (2004). Imprisoned knowledge: Criminals' beliefs about deception. *Legal and Criminological Psychology, 9*, 103-119. Doi: 10.1348/135532504322776889

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

Harris, C. B., Keil, P. G., Sutton, J., Barnier, A. J., & McIlwain, D. J. F. (2011). We remember, we forget: Collaborative remembering in older couples. *Discourse*

*Process, 48*, 267–303. Doi: 10.1080/0163853X.2010.541854

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

Hartwig, M., Granhag, P. A., Strömwall, L, & Doering, N. (2010). Impression and information management: On the strategic self-regulation of innocent and guilty suspects. *The Open Criminology Journal, 3*, 10-16 (special issue on deception research).

Hollingshead, A. B. (1998). Retrieval processes in transactive memory systems. *Journal of Personality and Social Psychology, 74*, 659-671. Doi: 10. 1037/0022-3514.74.3.659

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.org/10.7771/1932-6246.1167

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

Johnson, M. K. (1988). Reality Monitoring: An experimental phenomenological approach. *Journal of Experimental Psychology: General, 117*, 390-394.

Johnson, M. K. (2006). Memory and reality. *American Psychologist, 61*, 760-771.

10.1037/0003-066X.61.8.760.

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>

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

Meade, M. L., Nokes, T. J., & Morrow, D. G. (2009). Expertise promotes facilitation on a collaborative memory task. *Memory*, *17*, 39–48. Doi: 10.1080/09658210802524240

Mowat, L., (2016). <https://www.express.co.uk/news/uk/648025/Sham-marriages-increased-by-almost-850-and-authorities-are-overwhelmed-warns-MP>

Nahari, G., Ashkenazi, T., Fisher, R. P., Granhag, P. A., Hershkovitz, I., Masip, J., Meijer, E., Nisin, Z., Sarid, N., Taylor, P. J., Verschuere, B., & Vrij, A. (2019). Language of Lies: Urgent issues and prospects in verbal lie detection research. *Legal and Criminological Psychology*, *24*, 1-23. Doi:10.1111/lcrp.12148

Nahari, G., & Pazuelo, M. (2015). Telling a convincing story: Richness in detail as a function of gender and priming. *Journal of Applied Research in Memory and Cognition*, *4*, 363-367. Doi: 10.1016/j.jarmac.2015.08.005.

Nahari, G. & Vrij, A. (2014). Are you as good as me at telling a story? Individual differences in interpersonal-Reality Monitoring. *Psychology, Crime, & Law*, *20*, 573-583. Doi:10.1080/1068316X.2013.793771

Parkhouse, T., & Ormerod, T. C. (2018). Unanticipated questions can yield unanticipated outcomes in investigative interviews. *PLoS ONE*, *13*: e0208751. Doi: 10.1371/journal.pone.0208751

Seminara, D. (2018). <https://cis.org/Report/Hello-I-Love-You-Wont-You-Tell-Me-Your-Name-Inside-Green-Card-Marriage-Phenomenon>



Sporer, S. L. (2004). Reality monitoring and detection of deception. In P. A. Granhag & L. A. Strömwall (Eds.), *Deception detection in forensic contexts* (pp. 64-102). Cambridge, England: Cambridge University Press.

Sporer, S. L. (2016). Deception and cognitive load: Expanding our horizon with a working memory model. *Frontiers in Psychology: Hypothesis and Theory*, 7, article 420. Doi: 10.3389/fpsyg.2016.00420.

Sporer, S. L., & Schwandt, B. (2006). Paraverbal indicators of deception: A meta-analytic synthesis. *Applied Cognitive Psychology*, 20, 421-446. Doi: 10.1002/acp.1190

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.

Vernham, Z., & Vrij, A. (2015). A review of the collective interviewing approach to detecting deception in pairs. *Crime Psychology Review*, 1, 43-58. Doi: 10.1080/23744006.2015.1051756

Vernham, Z., Vrij, A., Leal, S., & Mann, S. (2014a). Collective Interviewing: A transactive memory approach towards identifying signs of truthfulness. *Journal of Applied Research in Memory and Cognition*, 3, 12-20. Doi: 10.1016/j.jarmac.2014.01.001

Vernham, Z., Vrij, A., Mann, S., Leal, S., & Hillman, J. (2014b). Collective interviewing: Eliciting cues to deceit using a turn-taking approach. *Psychology, Public Policy and Law*, 20, 309-324. Doi: 10.1037/law0000015.

Vredeveltdt, A., & van Koppen, P. J. (2018). Recounting a common experience: On the effectiveness of instructing eyewitness pairs. *Frontiers in Psychology*, 09 March 2018. Doi: 10.3389/fpsyg.2018.00284

Vredeveltdt, 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. (2014). Interviewing to detect deception. *European Psychologist, 19*, 184-195 Doi: 10.1027/1016-9040/a000201

Vrij, A. (2019). Deception and truth detection when analysing nonverbal and verbal cues. *Applied Cognitive Psychology, 33*, 160-167. Doi:10.1002/acp.3457

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

Vrij, A., Fisher, R., Mann, S., & Leal, S. (2006). Detecting deception by manipulating cognitive load. *Trends in Cognitive Sciences, 10*, 141-142. Doi: 10.1016/j.tics.2006.02.003

Vrij, A., Leal, S., Jupe, L., & Harvey, A. (2018). 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., Vernham, Z., & Brankaert, F. (2015). Translating theory into practice: Evaluating a cognitive lie detection training workshop. *Journal of Applied Research in Memory and Cognition, 4*, 110-120. doi:10.1016/j.jarmac.2015.02.002

Vrij, A., Mann, S., Fisher, R., Leal, S., Milne, B., & Bull, R. (2008). Increasing cognitive load to facilitate lie detection: The benefit of recalling an event in reverse order. *Law and Human Behavior, 32*, 253-265. Doi: 10.1007/s10979-007-9103-y.

Vrij, A., Mann, S., Leal, S., & Granhag, P. A. (2010). Getting into the minds of pairs of liars and truth tellers: An examination of their strategies. *The Open Criminology Journal, 3*, 17-22. Doi: 10.2174/1874917801003010017 (special issue on deception research).

Walczyk, J. J., Igou, F. P., Dixon, A. P., & Tcholakian, T. (2013). Advancing lie detection by inducing cognitive load on liars: A review of relevant theories and techniques guided by

lessons from polygraph-based approaches. *Frontiers in Psychology*, 4, 14. Doi:

10.3389/fpsyg.2013.00014

Walczyk, J. J., Roper, K. S., Seemann, E., & Humphrey, A. M. (2003). Cognitive mechanisms underlying lying to questions: Response time as a cue to deception. *Applied Cognitive Psychology*, 17, 755-774. Doi: 10.1002/acp.914

Wegner, D. M. (1987). Transactive memory: A contemporary analysis of the group mind. In B. Mullen & G. R. Goethals (Ed.), *Theories of group behaviour* (pp. 185-208). New York: Springer-Verlag.

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

Accepted Article

Table 1

*Questionnaire Variables as a Function of Veracity*

	Truth tellers			Liars			<i>F</i>	<i>p</i>	$\eta_p^2$
	<i>M</i>	<i>SD</i>	<i>95% CI</i>	<i>M</i>	<i>SD</i>	<i>95% CI</i>			
Motivation	5.94	0.78	5.70, 6.19	5.79	0.93	5.56, 6.03	.80	.372	.01
Truthfulness	98.17	3.84	91.60, 104.74	19.43	32.01	13.19, 25.67	310.12	< .001	.75
Likelihood prize	6.48	0.83	6.16, 6.80	4.94	1.34	4.64, 5.25	49.99	< .001	.33
Likelihood statement	1.74	1.22	1.32, 2.14	3.54	1.73	4.14, 3.96	38.55	< .001	.27
Anticipation collective interview	3.54	2.37	2.90, 4.18	3.25	2.32	2.61, 3.88	.41	.523	.00
Anticipation recording	3.48	1.73	3.02, 3.94	4.13	1.45	3.70, 4.57	4.36	.039	.04
Difficulty	1.63	1.16	1.22, 2.05	2.45	1.64	2.07, 2.84	8.74	.004	.08

Table 2

*Dependent Variables as a Function of Veracity: ANOVA Results*

	ICC	Truth tellers			Liars			<i>F</i>	<i>p</i>	$\eta_p^2$	<i>d</i>	<i>BF</i> <sub>10</sub>
		<i>M</i>	<i>SD</i>	95% CI	<i>M</i>	<i>SD</i>	95% CI					
<b>MEMORY</b>												
Said remember	.88	1.23	1.21	0.72,1.75	1.59	1.65	1.09,2.10	0.83	.368	.02	.25 (-.30, .78)	0.388
Time context	.81	0.81	1.06	0.46,1.16	1.00	0.78	0.66,1.34	0.57	.455	.01	.21 (-.35, .75)	0.349
Location context	.63	1.23	0.99	0.82,1.65	1.78	1.05	1.37,2.19	3.79	.057	.07	.54 (-.02, 1.08)	1.290
*Providing cues	.72	0.19	0.63	-0.12,.51	0.85	0.95	0.55,1.16	8.79	.005	.15	.82 (.24, 1.36)	8.886
*Finish sentence	.76	0.54	0.81	0.32,0.75	0.15	0.36	-.06,0.36	4.57	.013	.11	.63 (.07, 1.17)	2.230
<b>ALTERNATIVES</b>												
*More than one alternative discussed	.69	2.54	1.03	2.15,2.93	1.11	0.98	0.73,1.50	26.92	< .001	.35	1.43 (.80, 2.01)	408.312
*Experiences reported	.86	6.73	3.91	5.72,7.74	2.96	1.37	1.97,3.95	22.12	< .001	.30	1.29 (0.67, 1.86)	928.055
<b>COGNITIVE LOAD</b>												
Difficult to answer	.64	0.12	0.43	-.04,0.27	0.25	0.44	-.04,0.26	.002	.967	.00	.30 (0.25, .84)	0.276
Easy to answer	1.00	0.15	0.37	0.02,0.29	0.15	0.37	0.02,0.28	.003	.955	.00	.00, -.54, .54)	0.267
<b>STRATEGY</b>												
Verbal strategy	1.00	0.00	0.00	-.16,0.16	0.15	0.46	-.01,0.31	2.74	.104	.05	.46 (-.10, 1.00)	0.847
Nonverbal strategy	1.00	0.00	0.00	-.09,0.09	0.00	0.00	-.09,0.09					
<b>IMPRESSION MANAGEMENT</b>												
Perceived credibility	.64	0.08	0.27	-.14,0.29	0.11	0.42	0.10,0.32	.121	.729	.00	.08 (-.46, .62)	0.290
*Perceived embarrassment	.88	0.38	0.64	0.22,0.55	0.07	0.27	-.09,0.24	5.43	.024	.10	.64 (.07, 1.18)	2.456
Swearing	.88	1.23	1.03	0.86,1.60	0.70	0.91	0.34,1.07	3.89	.054	.07	.55 (-.01, 1.09)	1.341
Jokes	.77	0.65	0.89	0.37,0.94	0.41	0.74	0.13,0.69	1.19	.280	.02	.38 (-.17, .92)	0.451
<b>CONSISTENCY</b>												
Elaboration within	.62	2.85	0.97	2.46,3.23	3.19	0.88	2.81,3.56	1.79	.187	.03	.37 (-.18, .90)	0.575
Elaboration new	.61	1.96	0.92	1.58,2.34	1.52	0.98	1.14,1.89	2.90	.095	.05	.46 (-.09, 1.00)	0.903
Omission	.77	1.85	1.26	1.41,2.28	1.37	0.79	0.94,1.80	2.75	.104	.05	.42 (-.13, .96)	0.849

**GIVEAWAY  
CUES**

*Know each other	.74	0.92	0.93	0.69,1.15	0.11	0.32	-.11,0.34	18.18	< .001	.26	1.17 (.57, 1.74)	3.156
Did not know each other	.86	0.08	0.27	-.20,0.35	0.26	0.59	-.01,0.53	2.04	.160	.04	.39 (-.16, .93)	0.637
Fabrication	.59	0.08	0.27	-.19,0.35	0.19	0.48	-.08,0.45	1.00	.332	.02	.28 (-.26, .82)	0.417

---

<sup>i</sup> Discussing nonverbal strategies did not occur, and this variable did therefore not contribute to the multivariate effect.