



The effectiveness of different model statement variants for eliciting information and cues to deceit

Sharon Leal*, Aldert Vrij , Charlotte Hudson, Pasquale Capuozzo and Haneen Deeb

Psychology Department, University of Portsmouth, UK

Background. According to previous research, the use of a model statement results in both truth tellers and lie tellers reporting a similar amount of extra information than the instruction to be detailed. We examined (1) whether level of engagement with attending to the model statement affects the veracity findings and (2) whether valuable details is a diagnostic veracity indicator.

Method. We created four model statement variants, two had lower levels of engagement and two had higher levels of engagement with attending to the model statement content. Participants were allocated to one of these four conditions or to the no model statement control condition.

Results. Participants in one of the higher engagement conditions recalled the model statement content significantly better than participants in one of the lower engagement conditions. The audio model statement and the face-to-face model statement resulted in more information than the no model statement control condition. In none of the model statement conditions did total details emerge as a veracity indicator; valuable details did not yield the expected effect either.

Conclusion. A model statement serves as an eliciting information tool and the amount of additional information is similar among truth tellers and lie tellers.

To encourage interviewees to provide more information researchers have developed the model statement tool: An audiotaped example of a detailed verbal account that is unrelated to the topic of investigation (Leal, Vrij, Warmelink, Vernham, & Fisher, 2015). Exposure to the model statement raises expectations among interviewees – both truth tellers and lie tellers – about how much information they are expected to provide (Ewens et al., 2016). A model statement leads to more information than an instruction to be detailed because a model statement is an example and examples are perhaps easier to follow than instructions (Vrij, Leal, & Fisher, 2018).

A model statement facilitates providing information among truth tellers and lie tellers to a similar extent, which means that the amount of additional reported information (details) after exposure to a model statement does not distinguish truth tellers from lie tellers (Vrij, Leal, & Fisher, 2018). However, previous research has shown that differences

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

*Correspondence should be addressed to Sharon Leal, Department of Psychology, University of Portsmouth, King Henry Building, King Henry 1 Street, Hants PO1 2DY, UK (email: sharon.leal@port.ac.uk).

between truth tellers and lie tellers emerge after exposure to a model statement in specific types of detail, such as complications (Deeb, Vrij, & Leal, 2020; Vrij, Leal, Jupe, Leal, Jupe, & Harvey, 2018; Vrij et al., 2017) and core details (Deeb et al., 2020; Leal, Vrij, Deeb, Vrij, Deeb, & Jupe, 2018; Sakrisvold, Granhag, & MacGiolla, 2017). In the present experiment, we continued the search for cues to deceit in model statement interviews. First, we examined whether differences between truth tellers and lie tellers in details emerge if the format in which the model statement is delivered (low vs. high engagement model statements) is taken into account. Second, we examined a type of detail not investigated before: valuable details.

Different types of model statement

In almost all research to date, the model statement is introduced in an audiotaped format whereby interviewees listen to an audiotaped verbal account. This gives lie tellers a good opportunity to disengage quickly: Instead of listening to the model statement they could think of what to report next. Differences in reporting details between truth tellers and lie tellers after exposure to a model statement may emerge if during exposure to the model statement interviewees have fewer opportunities to think of what they are going to say next. Since truth tellers can rely on their memory of the event, having no opportunity to think of what to say next should affect them less than lie tellers who cannot rely on such memory. A way to achieve this lack of opportunity is by changing the level of engagement with listening to the model statement. For example, providing interviewees with a written version of the model statement and asking them to read it aloud should give them less opportunity to disengage with the model statement than listening to the audiotaped version. Indeed, research has shown that listening leads to more mind wandering than reading aloud (Varao Sousa, Carriere, & Smilek, 2013).

In the present pre-registered experiment (<https://osf.io/2twdp>), we introduced two variants of the model statement that do not require high engagement with or attention to the model statement: the audiotaped variant (hereafter labelled Audio MS) and the written variant in which the interviewee is asked to read the text in silence (hereafter labelled Written MS). Varao Sousa et al. (2013) found that also reading a text in silence leads to more mind wandering than reading a text aloud. We also introduced two higher engagement variants: The variant where the interviewee reads the model statement aloud (hereafter labelled Read Aloud MS) and the variant where the interviewer reads the model statement aloud while maintaining eye contact with the interviewee (hereafter labelled Face-to-Face MS).

We are not aware of mind wandering research in the face-to-face variant and this variant may deserve a little introduction. Interviewees who want to think of what to say next should prefer to look away from the interviewer because eye contact with a conversation partner is distracting (Doherty-Sneddon & Phelps, 2005; Vrij, Mann, Leal, & Fisher, 2010). However, gaze aversion is difficult when the interviewer speaks and maintains eye contact because looking at the speaker and reciprocating eye contact is the norm in social interactions (Binetti, Harrison, Coutrot, Mareschal, & Johnston, 2015; Rogers, Speelman, Guidetti, & Longmuir, 2018). Displaying non-normative behaviour comes across as suspicious (Bond et al., 1992), something lie tellers typically want to avoid (Hocking & Leathers, 1980).

Valuable details

A distinction has been made in deception research between core and peripheral details (Leal et al., 2018). Details that, if changed, can result in changes in the basic and most important part of the story are considered core details; details that have no such impact are considered peripheral (Heuer & Reisberg, 1990). To explain this distinction, Leal et al. (2018) gave as example attending an Adele concert: All details that are about the actual concert are considered core details, whereas peripheral details are those such as the drinks in the pub before and after the concert.

In the present experiment participants recalled a videotaped secret meeting. All these details are core details, but we made a further distinction within this core details category and distinguished between more or less valuable core details. In the example above, recalling things that Adele said to the audience during the concert would be stronger evidence to demonstrate that someone actually attended the concert (more valuable details) than reporting that she sang a few of her well known songs (less valuable details), because the latter, less idiosyncratic information, can be easily made up without attending the conference. Lie tellers may realize that they are expected to provide more details after being exposed to a model statement, but they have to consider what details to add. To tell a successful lie they should avoid adding valuable details, because the more they say about the topic, the greater the opportunity they give to investigators to discover that these additional details are not true (see Hartwig, Granhag, & Strömwall, 2007). In contrast, for truth tellers this distinction between more and less valuable details seem less relevant.

Accuracy of truth tellers' accounts

In the present experiment we also measured amongst truth tellers the accuracy of the details reported in the control and model statement conditions. When truth tellers' expectations are raised about how much information is expected from them, the accuracy of the information they provide may decrease (Ackerman & Goldsmith, 2008; Koriat & Goldsmith, 1996). A previous experiment did not reveal a decrease in accuracy in model statement interviews when interviewees were interviewed immediately after an event (Ewens et al., 2016). In that experiment the Audio MS was used but it is important to examine whether this effect can be generalized to different Model Statement variants.

Hypotheses

The following pre-registered hypotheses were tested:

Both truth tellers and lie tellers will report more new details after exposure to a model statement than in the no model statement control condition (Model Statement main effect: Hypothesis 1).

The low engagement model statement conditions (Audio MS and Written MS) will yield more new details and new valuable details from truth tellers than the high engagement MS conditions (Face to Face MS and Read Aloud MS) as it will be easier for truth tellers to disengage and focus on their memory of the event (Model Statement main effect, truth tellers only: Hypothesis 2).

The low engagement MS conditions (Audio MS and Written MS) will elicit more new details from lie tellers than the high engagement MS conditions (Face to Face MS and Read Aloud MS) but not more new valuable details. In the low engagement MS conditions, lie tellers will be able to disengage and think of non-essential new details that they can add

whilst avoiding giving away additional valuable details (Model Statement main effect, lie tellers only: Hypothesis 3).

Following from Hypotheses 2 and 3, the smallest difference between truth tellers' and lie tellers' accounts will emerge in reporting new details in the two low engagement model statement conditions (Veracity \times Model Statement interaction effect: Hypothesis 4).

As reported in the pre-registration, we will explore the effect of a model statement on the accuracy of truth tellers' recall of the experienced event.

To measure the level of engagement with listening to the model statement, we asked participants in the model statement conditions not only to recall the experienced event but also to recall the content of the model statement. More engagement should result in a better recall of the model statement. We thus expected that participants in the high engagement model statement conditions (Face to Face MS and Read Aloud MS) will recall a higher proportion of accurate details about the content of the model statement than participants in the low engagement model statement conditions (Audio MS and Written MS; Hypothesis 5). This hypothesis was not pre-registered.

Method

Participants

A total of 230 students and University personnel took part in the study, 170 females and 58 males. Two participants did not reveal their gender. Their average age was $M = 24.57$ years ($SD = 9.97$). A large group of participants ($n = 115$) described themselves as White British. Others identified themselves as White European ($n = 42$), Asian ($n = 32$), Black British ($n = 14$), mixed ($n = 10$), African ($n = 3$), Black European ($n = 1$) or Arab ($n = 1$). Twelve participants did not answer the question.

Design

The experiment employed a 2 (Veracity: truth tellers, lie tellers) \times 5 (Model Statement: No MS, Audio MS, Written MS, Face to Face MS, Read Aloud MS) factorial design. Participants were randomly allocated to each of the ten cells which ranged in size from 22 to 23 participants. A total of 114 participants were allocated to the truth tellers condition and 116 to the lie tellers condition; 47 participants were allocated to the No MS condition, 46 to the Audio MS condition, 46 to the Written MS condition, 45 to the Read Aloud MS condition and 46 to the Face to Face MS condition.

The analyses were pre-registered. To investigate Hypotheses 1 and 4, analyses of variance were conducted utilizing a 2 (Veracity: truth tellers, lie tellers) \times 5 (Model Statement: No MS, Audio MS, Written MS, Face to Face MS, Read Aloud MS) factorial design with details in Phase 1, valuable details in Phase 1, new details in Phase 2 and new valuable details in Phase 2 as dependent variables. To investigate Hypothesis 2, only truth tellers were included in the analysis. To investigate Hypothesis 3, only lie tellers were included in the analysis.

Materials

Secret meeting recording

Participants watched a secret meeting recording whereby three individuals discussed where to plant a spy device. In the recording, the technical features (e.g., the device can

hack mobile phones) and physical features (e.g., the device looks like a smoke detector) were discussed before the three people voted on the best location to plant it. The same secret meeting scenario was used by Shaw et al. (2013), Shaw, Vrij, Leal, and Mann (2014).

Model statement conditions

Participants in the model statement present conditions were subjected to one of the following four model statement variations:

1. Audiotaped model statement (Audio MS)
2. The interviewee read the model statement in silence (Written MS)
3. The interviewer read the model statement aloud whilst maintaining eye contact with the interviewee (Face to Face MS)
4. The interviewee read the model statement aloud (Read Aloud MS)

For the Face to Face MS we checked before the experiment started that the interviewer complied with the instructions. Once the interviewer knew the content of the model statement well, s/he was able to maintain good eye contact while reading out the model statement.

The same Model Statement was used in each condition. We used the model statement introduced by Leal et al. (2015) in which someone recalls his experiences while attending a Formula 2 motor race event.

Procedure

Participants were recruited via the University intranet. The experiment was advertised as an investigation of credibility cues when individuals lie or tell the truth about a witnessed event. Everyone over 18 years old with a good grasp of English could take part. Participants were told that in the experiment they had to imagine that they are an undercover agent for the government assigned to either tell the truth or to lie about a secret film they would watch. Participants would receive one study credit or £10 for their time and, if they sounded convincing, they would also be entered in a prize draw with the possibility to earn £150, £75 or £50.

After arrival at the laboratory and completing the consent form, participants were told they would see a 'secret meeting recording', which they should try to remember. Truth tellers were instructed to tell the truth about everything they witnessed in the film, whereas lie tellers were instructed to lie to the interviewer about: (1) the appearance of the device, (2) the people they saw in the video, and (3) the location where the device would be planted. Participants were told that if the interviewer believed they were telling the truth they would be entered into a draw to win up to £150. If, however, the interviewer did not believe them they would be asked to write a statement about the film.

Participants were then offered as much time to prepare for the interview as they wished. Once participants said that they were ready for the interview they were asked to complete the pre-interview questionnaire measuring three background characteristics (gender, age, ethnicity). It also measured motivation: 'To what extent are you motivated to do well in the experiment' (on a 5-point Likert scale ranging from 1 = *not at all* to 5 = *very much*). It further measured preparation thoroughness: 'How would you describe your preparation for the interview'. Answers were given on three Osgood scales: 1 (*shallow*) to 7 (*thorough*), 1 (*insufficient*) to 7 (*sufficient*), and 1 (*poor*) to 7 (*good*).

The three answers were averaged (Cronbach's alpha = .92). Participants were also asked whether they thought they were given enough preparation time: 'Do you think the amount of time you were given to prepare was': 1 (*insufficient*) to 7 (*sufficient*).

After completing the pre-interview questionnaire, participants were brought to the interview room. The interview consisted of two phases. The first phase involved a free recall to the request: "Please tell me in as much detail as possible everything that happened in the secret meeting recording you witnessed from the very start to the very end of the meeting". After the free recall the interviewer left the room for a moment. After re-entering the room, the interviewer said: "Sorry but I am going to have to ask you again. Please tell me once more in as much detail as possible what you saw in the secret film clip" (control condition). In the four model statement conditions, the interviewer also said: "Before doing so I will expose you to a model statement to give you an idea of exactly how much detail I would like you to include in your response". Participants were then subjected to one of the four model statement variations.

After recalling the witnessed film for the second time (Phase 2), participants were asked what they thought the likelihood was that they had to provide a handwritten statement about the secret meeting on a 7-point Likert scale ranging from 1 (*not at all*) to 7 (*completely*). Participants in the four model statement conditions were additionally asked the following four questions about the model statement which could be answered on 7-point Likert scales ranging from 1 (*not at all*) to 7 (*definitely*): (1) The model statement made me give fewer details than I originally planned to provide; (2) The model statement made me give more details than I originally planned to provide; (3) During the model statement I was paying careful attention to the content of the statement; and (4) During the model statement I was thinking about what I would say in the interview. They were finally asked to write down in as much detail as they could remember what the man said in the model statement.

After completing the post-interview questionnaire participants were debriefed and given £10 or one course credit. All participants were entered in the prize draw.

Coding

The interviews were audiotaped and transcribed; and the transcripts were coded

An experienced coder, previously trained by us and blind to the veracity status and hypotheses, coded the number of details in each statement. Details were defined as a unit of information. For example, the following sentence has 15 details: "I saw two ladies sitting down at a table talking to each other. A man in a suit came in and put his black mug on a shelf and introduced himself". In Phase 2, only new details were counted (repetitions were ignored). The coder was then shown the videotaped scenario and asked to categorize the coded details into accurate and inaccurate details.

A second coder, also experienced and previously trained by us and also blind to the veracity status and hypotheses, coded a random sample of 51 transcripts (22% of the sample). Inter-rater reliability between the two coders, using the two-way random effects model measuring consistency, was very good for all measures: Details in Phase 1 (Single Measures ICC = .92), inaccurate details in Phase 1 (Single Measures ICC = .84), new details in Phase 2 (Single Measures ICC = .84), and inaccurate new details in Phase 2 (Single Measures ICC = .86).

Regarding valuable details coding, we prepared a list of 17 details considered valuable in the secret meeting. The list was derived from Shaw et al. (2014), see Appendix A.

Following Shaw et al. (2014), we instructed the coder, blind to the Veracity condition and hypotheses, to assess the extent to which these details occurred in Phase 1 using a 7-point scale ranging from 1 (*little or no valuable information*) to 7 (*highly valuable information*). We also asked the coder to assess the number of new valuable details in Phase 2 of the interview (valuable details mentioned in Phase 2 but not in Phase 1). A second coder followed the same procedure on 51 transcripts (22% of the sample). Inter-rater reliability between the two coders, using the two-way random effects model measuring consistency, was very good for valuable details in Phase 1 (Single Measures ICC = .89) and new valuable details in Phase 2 (Single Measures ICC = .94).

An alternative way of measuring valuable details is to count the number of details on the 17-point scale (Appendix A). We instructed the coder and second coder also to do this. Inter-rater reliability between the two coders, using the two-way random effects model measuring consistency, was very good for valuable details in Phase 1 (Single Measures ICC = .89) and new valuable details in Phase 2 (Single Measures ICC = .89). Since we included only the 7-points scale coding of valuable details in the pre-registration we report that scale coding in the main text and the 17-points scale coding in Appendix B. The two sets of analysis produced almost identical results.

Regarding accuracy of details in reporting the model statement, a coder, blind to the Veracity condition and hypotheses, coded the number of accurate and inaccurate details the participants wrote down when reporting the model statement content. A second coder followed the same procedure on 51 transcripts (22% of the sample). Inter-rater reliability between the two coders, using the two-way random effects model measuring consistency, was very good for both accurate (Single Measures ICC = .93) and inaccurate details (Single Measures ICC = .88).

The main coders and reliability coders rated the statements separately from each other and did not discuss their ratings. Only the ratings of main coders were used and, given the high inter-rater reliability, no changes were made.

Results

Motivation, preparation thoroughness, preparation time and having to write a statement

Four 2 (Veracity) \times 5 (Model Statement) ANOVAs were carried out with motivation, preparation thoroughness, preparation time; and having to write a statement as dependent variables. Motivation was measured on a 5-point Likert scale and the other variables on 7-point Likert scales. No significant main effects or interaction effects emerged for the first three variables, all F 's < 1.86, all p 's > .119. Participants were very well motivated ($M = 4.30$, $SD = 1.60$), thought that their preparation was reasonably good ($M = 4.73$, $SD = 1.33$), and thought they were given sufficient preparation time ($M = 5.36$, $SD = 1.41$).

Lie tellers ($M = 4.23$, $SD = 1.97$, 95% CI [3.89, 4.56]) more than truth tellers ($M = 3.63$, $SD = 1.86$, 95% CI [3.29, 3.98]) thought it to be likely that they had to write a statement, $F(1, 220) = 5.81$, $p = .017$, $d = 0.31$, 95% CI [0.05, 0.57]. The interaction effect was also significant, $F(4, 220) = 3.19$, $p = .014$, $\eta_p^2 = .06$. Only in the Written MS condition were lie tellers more worried than truth tellers that they had to write a statement, see Table 1.

Table 1. Statistics for having to write a statement as a function of veracity and model statement

	Truth tellers			Lie tellers			<i>F</i>	<i>p</i>	<i>d</i>	95% CI
	<i>M</i>	<i>SD</i>	95% CI	<i>M</i>	<i>SD</i>	95% CI				
No MS	3.65	2.04	2.86, 4.45	4.12	1.78	3.34, 4.91	0.72	.400	0.25	−0.33, 0.82
Audio MS	3.39	1.90	2.60, 4.19	4.22	1.88	3.42, 5.01	2.19	.146	0.44	0.01, 0.89
Written MS	3.26	1.54	2.42, 4.10	5.43	2.37	4.59, 6.28	13.58	.001	0.77	0.16, 1.36
Face to Face MS	3.96	2.33	3.04, 4.88	3.83	2.04	2.91, 4.75	0.04	.841	0.06	−0.52, 0.64
Read Aloud MS	3.91	1.41	3.34, 4.48	3.57	1.24	3.01, 4.12	0.76	.389	0.26	−0.33, 0.84

Impressions of the model statement

A 2 (Veracity) \times 4 (Model Statement) MANOVA was carried out with the four model statement variants included and the no model statement control condition excluded. The dependent variables were the four impressions of the model statement, which were measured on 7-point Likert scales. The analysis resulted in non-significant multivariate main effects for Veracity, $F(4, 172) = 1.11, p = .356, \eta_p^2 = .03$, and Model Statement, $F(12, 512) = 1.24, p = .252, \eta_p^2 = .03$. Also the Veracity \times Model Statement interaction effect was not significant, $F(12, 512) = 0.36, p = .975, \eta_p^2 = .01$. Participants disagreed that the model statement made them give fewer details ($M = 2.14, SD = 1.45$) but agreed that it made them give more details than originally planned ($M = 5.46, SD = 1.63$). The statements that the participants listened to the model statement ($M = 4.77, SD = 1.66$) and that they were thinking about what to say next ($M = 4.60, SD = 1.87$) received similar support.

Details in statements in phase 1

Two 2 (Veracity) \times 5 (Model Statement) ANOVAs were conducted with pre-MS details and pre-MS valuable details (7-point scale) as dependent variables. The details analysis resulted in a significant Veracity main effect, see Table 2. Truth tellers reported more details in Phase 1 than lie tellers. The Model Statement main effect, $F(4, 220) = 2.12, p = .080, \eta_p^2 = .04$, and the Veracity \times Model Statement interaction effect, $F(4, 220) = 0.75, p = .562, \eta_p^2 = .01$, were not significant.

The valuable details analysis also resulted in a significant Veracity main effect, see Table 2. Truth tellers also reported more valuable details in Phase 1 than lie tellers. The Model Statement main effect, $F(4, 220) = 0.99, p = .414, \eta_p^2 = .02$, and the Veracity \times Model Statement interaction effect, $F(4, 220) = 0.52, p = .718, \eta_p^2 = .01$, were not significant.

New details in statements in phase 2

Two 2 (Veracity) \times 5 (Model Statement) ANOVAs were conducted with post-MS new details and post-MS new valuable details (7-point scale) as dependent variables. In the new details analysis the Veracity main effect, $F(1, 220) = 0.16, p = .690, d = 0.05, 95\% \text{ CI } [-0.21, 0.31]$, and the Veracity \times Model Statement interaction effect, $F(4, 220) = 0.68, p = .604, \eta_p^2 = .01$, were not significant, but the Model Statement main effect was significant, $F(4, 220) = 3.01, p = .019, \eta_p^2 = .05$. The Audio MS and Face to Face MS conditions resulted in more Post-MS new details than the no model statement condition, see Table 3. The other comparisons were not significant, all p 's $> .134$. This partially supports Hypothesis 1.

Table 2. Statistics for details as a function of veracity

	Truth tellers				Lie tellers				F	p	d	95% CI
	M	SD	95% CI		M	SD	95% CI					
Pre-MS (Phase 1)												
Details	63.04	23.46	58.46, 67.58		50.32	26.22	45.76, 54.80	15.28	<.001	0.51	0.24, 0.77	
Valuable details	04.80	01.68	04.48, 05.12		03.57	01.73	03.26, 03.88	29.68	<.001	0.72	0.44, 0.98	
Post-MS new details (Phase 2)												
Details	22.45	15.80	19.65, 25.26		23.19	15.05	20.48, 26.04	00.16	.690	0.05	-0.21, 0.31	
Valuable details	01.98	01.46	01.66, 02.30		02.65	01.93	02.33, 02.97	08.56	.004	0.39	0.12, 0.65	

Table 3. Statistics for details and accuracy proportion as a function of model statement (5 levels)

Variable	No MS					Low engagement MS					High engagement MS																															
	M		SD		95% CI		Audio MS		SD		95% CI		Written MS		SD		95% CI		Face to face MS		SD		95% CI		Read aloud MS		SD		95% CI													
	M	SD	M	SD	95% CI	M	SD	M	SD	M	SD	95% CI	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD								
Pre-MS (Phase 1)																																										
Details	64.55	28.88	57.65	71.86	56.17	26.24	49.00	63.35	54.76	23.29	47.58	61.94	50.24	24.68	43.06	57.42	57.22	23.53	50.04	64.56	4.96	1.74	3.78	4.76	4.70	1.79	3.37	4.37	4.74	1.91	3.83	4.82	4.70	1.69	3.48	4.48	4.91	1.34	3.97	4.98		
Accuracy (truth tellers only)	0.90	0.13	0.83	0.96	0.87	0.19	0.81	0.94	0.91	0.13	0.85	0.98	0.89	0.17	0.82	0.95	0.90	0.14	0.83	0.96																						
Post-MS new details (Phase 2)																																										
Details	16.28 ^a	12.97	11.93	20.67	25.17 ^b	13.52	20.76	29.59	23.02 ^{ab}	16.18	18.61	27.44	26.09 ^b	16.72	21.67	30.50	23.71 ^{ab}	15.93	19.24	28.18																						
Valuable details	2.02	1.57	1.52	2.51	2.35	1.79	1.85	2.85	2.35	1.59	1.85	2.85	2.59	1.94	2.08	3.09	2.29	1.84	1.77	2.79																						
Accuracy (truth tellers only)	0.94	0.11	0.87	1.00	0.89	0.13	0.83	0.95	0.87	0.15	0.81	0.93	0.95	0.08	0.88	1.01	0.86	0.24	0.80	0.93																						

Note. Only mean scores with a different superscript differ significantly ($p < .05$) from each other. In the Audio-MS, the interviewee listens to the Model Statement. In the Written MS, the interviewee reads the Model Statement in silence. In the Face to Face MS, the interviewer reads the Model Statement aloud whilst maintaining eye contact with the interviewee. In the Read Aloud MS, the interviewee reads the Model Statement aloud. Valuable details were measured on a 7-point scale.

The analysis for new valuable details resulted in a significant Veracity main effect, see Table 2. Lie tellers reported more new valuable details in Phase 2 than truth tellers. The Model Statement main effect, $F(4, 220) = 0.65$, $p = .630$, $\eta_p^2 = .01$, and the Veracity \times Model Statement interaction effect, $F(4, 220) = 0.45$, $p = .772$, $\eta_p^2 = .01$, were not significant. The two nonsignificant interaction effects indicate that Hypothesis 4 was not supported.

To test Hypothesis 2, only truth tellers were selected. Two one-way ANOVAs with Model Statement (five levels) as the only factor and new details and new valuable details as dependent variables were carried out. The analyses for new details, $F(4, 109) = 1.49$, $p = .210$, $\eta_p^2 = .05$, and new valuable details, $F(4, 109) = 0.27$, $p = .899$, $\eta_p^2 = .01$, were both not significant. Hypothesis 2 was therefore not supported.

To test Hypothesis 3, only lie tellers were selected. Two one-way ANOVAs with the Model Statement (five levels) as the only factor and new details and new valuable details as dependent variables were conducted. The analyses for new details, $F(4, 111) = 2.27$, $p = .066$, $\eta_p^2 = .08$, and new valuable details, $F(4, 111) = 0.72$, $p = .583$, $\eta_p^2 = .03$, were both not significant. Hypothesis 3 was therefore not supported.

To check whether the non-significant results were due to a lack of statistical power, we ran a post-hoc power analysis using G*Power software. The input included an error probability of .05, a medium to large effect size ($f = 0.325$), and a numerator df of 4. The analysis revealed that the study achieved very high power (99%).

Proportion of accurate of details in statements amongst truth tellers

Accuracy proportions concerning the recall of the viewed secret meeting film are only relevant for truth tellers, so for the next set of analyses only truth tellers were included. Proportion scores were calculated by dividing accurately recalled details by the total of accurately and inaccurately recalled details. Two one-way ANOVAs with Model Statement (five levels) as the only factor and pre-MS and post-MS accuracy as dependent variables did not result in a significant effect for either pre-MS accuracy, $F(4, 109) = 0.21$, $p = .934$, $\eta_p^2 = .008$, or post-MS accuracy, $F(4, 109) = 1.44$, $p = .227$, $\eta_p^2 = .05$. Table 3 provides the accuracy rates.

Proportion of accurate details when recalling the model statement

An ANOVA was carried out utilizing a 2 (Veracity) \times 4 (Model Statement) design with the proportion of accurate details reported about the model statement as dependent variable. Proportion scores were calculated by dividing accurately recalled details by the total of accurately and inaccurately recalled details. The Veracity main effect, $F(1, 174) = 0.82$, $p = .368$, $d = 0.15$, 95% CI [-0.11, 0.40], and the Model Statement main effect, $F(3, 174) = 1.92$, $p = .129$, $\eta_p^2 = .03$, were both not significant, but the Veracity \times Model Statement interaction effect was significant, $F(3, 174) = 2.67$, $p = .049$, $\eta_p^2 = .04$.

Table 4 provides the accuracy proportion scores for truth tellers and lie tellers separately. For truth tellers a significant Model Statement main effect emerged, $F(3, 86) = 4.32$, $p = .007$, $\eta_p^2 = .13$. Accuracy rates were significantly higher in the Read Aloud MS than in the Audio MS condition, whereas the other comparisons were not significant, all p 's $> .195$. For lie tellers the Model Statement main effect was not significant, $F(3, 88) = 0.95$, $p = .442$, $\eta_p^2 = .03$, and none of the comparisons revealed a significant difference, all p 's $> .514$. Overall, the results on accuracy partially supported Hypothesis 5. See Appendix C for a similar set of analyses for the number of accurate details variable.

Table 4. Statistics for the proportion of accurate details in recalling the model statement as a function of model statement (4 levels)

Variable	Low engagement MS				High engagement MS							
	Audio MS		Written MS		Face to Face MS		Read aloud MS					
	M	SD	95% CI	M	SD	95% CI	M	SD	95% CI			
Proportion accurate details: Truth tellers	0.66 ^a	0.34	0.55, 0.77	0.78 ^{ab}	0.23	0.67, 0.89	0.80 ^{ab}	0.21	0.69, 0.91	0.92 ^b	0.08	0.81, 1.03
Proportion accurate details: Lie tellers	0.82 ^a	0.27	0.71, 0.93	0.71 ^a	0.32	0.60, 0.82	0.70 ^a	0.30	0.59, 0.81	0.79 ^a	0.29	0.68, 0.90

Note. Only mean scores with a different superscript differ significantly ($p < .05$) from each other. In the Audio-MS, the interviewee listens to the Model Statement. In the Written MS, the interviewee reads the Model Statement in silence. In the Face to Face MS, the interviewer reads the Model Statement aloud whilst maintaining eye contact with the interviewee. In the Read Aloud MS, the interviewee reads the Model Statement aloud.

Discussion

We examined the effectiveness of different model statements in eliciting information and cues to deceit. The results for the different model statement variants were similar. First, partially supporting Hypothesis 1, a model statement resulted in more reported details than the instruction to be detailed in every model statement condition (but the differences were only significant in two conditions). Second, in none of the conditions did ‘details’ become a diagnostic veracity cue, providing no support for Hypothesis 4. These two findings are typical for the classic model statement variant (audiotaped model statement) predominantly researched to date (Vrij, Leal, & Fisher, 2018).

Only two out of four model statement conditions resulted in significantly more new details than the control condition: The Audio MS and Face to Face MS. Previous research revealed that listening (Audio MS) in particular makes people’s minds wander (Varao Sousa et al., 2013). The finding that the Audio MS condition -amongst truth tellers- resulted in the lowest accuracy in recalling the model statement (supporting Hypothesis 5) suggests that mind wandering also occurred in this condition in the present experiment. Participants may have switched off listening to the Model Statement and started thinking what to say in the Phase 2 recall, which may have benefitted that recall.

The Face to Face MS condition resulted in significantly more new details than the control condition. We did not expect the Face to Face MS condition to stimulate mind wandering and the high accuracy rates in recalling the model statement in this condition showed no evidence that this was the case. We suggest another explanation. The Face to Face MS and the Audio MS have in common that the interviewee listens to the model statement. This is in contrast to the two other model statements conditions (Written MS and Read Aloud MS) where interviewees read. Perhaps there is a better match between the input ‘listening to an oral verbal statement’ and the output ‘producing an oral statement’ than between the input ‘reading a verbal statement’ and the output ‘producing an oral statement’. A better match between input and output facilitates recall (Fisher & Geiselman, 1992). However, the differences in elicited information between the four model statement conditions was very small (ranging from an average of 23.02 to an average of 26.99) and did not differ significantly from each other. This makes it somewhat difficult to conclude that one model statement is more effective in eliciting information than another. One of the two significant findings was obtained in the classic Audio MS condition, which provides support for continuing using this variant in future research for eliciting information purposes.

The self-reports from the post-interview questionnaire provide insight into why details did not become a veracity indicator in any of the model statement conditions. The model statement made both truth tellers and lie tellers equally aware that they were expected to report more details than they had anticipated. Both truth tellers and lie tellers further reported that, when exposed to the model statement, they were attending to the model statement but were also thinking of what they would say next. Regarding the latter finding, we hypothesized that participants in the two high engagement conditions (Face to Face MS and Read Aloud MS conditions) would have less opportunity to think of what to say next than participants in the low engagement conditions (Audio MS and Written MS conditions). Indeed, participants in one of the high engagement conditions (Read Aloud MS) did recall the Model Statement better than participants in one of the two low engagement conditions (Audio MS), suggesting that they paid more attention to the model statement than their counterparts. However, participants in the Read Aloud MS condition

still reported to be able to think about what to say next in the interview, suggesting that they were not totally engaged with listening to the model statement.

We examined a new verbal veracity cue: Valuable details. We expected lie tellers to avoid reporting such details more than truth tellers. This indeed happened in Phase 1, but not in Phase 2 when lie tellers reported more new valuable details than truth tellers, regardless of the model statement condition they were in. A combination of two factors may explain this. First, perhaps truth tellers already had reported most valuable information they could remember in Phase 1 and therefore had relatively little valuable information to add in Phase 2. The 17-point scale analysis (Appendix B) supports this. Truth tellers reported $M = 10.75$ ($SD = 3.88$) details in Phase 1 and added only $M = 1.66$ ($SD = 1.58$) details in Phase 2. Second, perhaps having to report the secret meeting for a second time may have put lie tellers under pressure to report new information. They may have struggled to add less valuable information and subsequently compromised themselves by adding valuable information. Whatever the reason, the distinction between more and less valuable information is worth exploring in future deception research.

The accuracy of the information truth tellers reported about the secret meeting was very high and did not differ between the model statement absent and model statement present conditions. This replicates previous research (Ewens et al., 2016) and suggests that the increase in informativeness does not affect accuracy. This applies to situations similar to those examined by Ewens et al. (2016) and in the present experiment where there is a short timespan between the experienced event and interview. The situation may be somewhat different when the timespan is longer (Harvey, Vrij, Leal, Hope, & Mann, 2019), which is worthwhile to investigate.

One limitation in particular is worth discussing. For reliability coding, the second rater coded 22% of the sample. This is not exceptional in deception research, for example, Granhag, Strömwall, Willén, and Hartwig (2013) and Shaw et al. (2014) coded 25% of the sample and Deeb et al. (2020) coded 15%. All ICCs in the present experiment were at least 0.84, so the coding appears reliable. However, it is likely that reliability would improve if a higher percentage of transcripts had been second coded. Indeed, Hauch, Sporer, Masip, and Blandón-Gitlin (2017) recommend all transcripts should be coded by at least two raters, as employed, for example, by Nahari, Vrij, and Fisher (2014) and Masip, Blandón-Gitlin, Martínez, Herrero, and Ibabe (2016).

In conclusion, the model statement variants were more beneficial for eliciting information than the instruction to be detailed, but this benefit was not evident for eliciting cues to deceit. The classical Audio MS that has been used in previous MS studies was one of the MS variants that elicited the most information in the current experiment. This sheds positive light on the validity of previous MS findings and encourages researchers to continue using this variant for effective information elicitation.

Acknowledgements

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

Conflicts of interest

All authors declare no conflict of interest.

Author contribution

Sharon Leal: Conceptualization (equal); Methodology (equal); Writing – original draft (equal). **Aldert Vrij:** Formal analysis (equal); Funding acquisition (equal); Writing – original draft (equal). **Charlotte Hudson:** Data curation (equal). **Pasquale Capuozzo:** Data curation (equal). **Haneen Deeb:** Formal analysis (equal); Writing – review & editing (equal).

Data availability statement

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

References

- Ackerman, R., & Goldsmith, M. (2008). Control over grain size in memory reporting—With and without satiating knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 1224–1245. <https://doi.org/10.1037/a0012938>
- Binetti, N., Harrison, C., Coutrot, A., Mareschal, I., & Johnston, A. (2015). Reciprocating the gaze of others: How we look and how long we like to be looked at. *Journal of Vision*, *15*(12), 172. <https://doi.org/10.1167/15.12.172>
- Bond, C. F., Omar, A., Pitre, U., Lashley, B. R., Skaggs, L. M., & Kirk, C. T. (1992). Fishy-looking liars: Deception judgment from expectancy violation. *Journal of Personality and Social Psychology*, *63*, 969–977. <https://doi.org/10.1037/0022-3514.63.6.969>
- Deeb, H., Vrij, A., & Leal, S. (2020). The effects of a model statement on information elicitation and deception detection in multiple interviews. *Acta Psychologica*, *207*, 103080. <https://doi.org/10.1016/j.actpsy.2020.103080>
- Doherty-Sneddon, G., & Phelps, F. G. (2005). Gaze aversion: A response to cognitive or social difficulty? *Memory and Cognition*, *33*, 727–733. <https://doi.org/10.3758/BF03195338>
- Ewens, S., Vrij, A., Leal, S., Mann, S., Jo, E., Shaboltas, A., . . . 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. <https://doi.org/10.1002/acp.3270>
- Fisher, R. P., & Geiselman, R. E. (1992). *Memory enhancing techniques for investigative interviewing: The cognitive interview*. Springfield, IL: Charles C. Thomas.
- Granhag, P. A., Strömwall, L. A., Willén, R., & Hartwig, M. (2013). Eliciting cues to deception by tactical disclosure of evidence: The first test of the Evidence Framing Matrix. *Legal and Criminological Psychology*, *18*, 341–355. <https://doi.org/10.1111/j.2044-8333.2012.02047.x>
- Hartwig, M., Anders Granhag, P., & Strömwall, L. A. (2007). Guilty and innocent suspects' strategies during police interrogations. *Psychology, Crime & Law*, *13*(2), 213–227. <https://doi.org/10.1080/10683160600750264>
- Harvey, A. C., Vrij, A., Leal, S., Hope, L., & Mann, S. (2019). Amplifying deceivers' flawed metacognition: Encouraging disclosures after delays with a model statement. *Acta Psychologica*, *200*, 102935. <https://doi.org/10.1016/j.actpsy.2019.102935>
- Hauch, V., Sporer, S. L., Masip, J., & Blandón-Gitlin, I. (2017). Can credibility criteria be assessed reliably? A meta-analysis of criteria-based content analysis. *Psychological Assessment*, *29*, 819–834. <https://doi.org/10.1037/pas0000426>
- Heuer, F., & Reisberg, D. (1990). Vivid memories of emotional events: The accuracy of remembered minutiae. *Memory & Cognition*, *18*, 496–506. <https://doi.org/10.3758/BF03198482>
- Hocking, J. E., & Leathers, D. G. (1980). Nonverbal indicators of deception: A new theoretical perspective. *Communication Monographs*, *47*(2), 119–131. <https://doi.org/10.1080/03637758009376025>

- Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review*, *103*, 490–517. <https://doi.org/10.1037/0033-295X.103.3.490>
- Leal, S., Vrij, A., Deeb, H., & Jupe, L. (2018). Using the Model Statement to elicit verbal differences between truth tellers and liars: The benefit of examining core and peripheral details. *Journal of Applied Research in Memory and Cognition*, *7*, 610–617. <https://doi.org/10.1016/j.jarmac.2018.07.001>
- 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*(1), 129–146. <https://doi.org/10.1111/lcrp.12017>
- Masip, J., Blandón-Gitlin, I., Martínez, C., Herrero, C., & Ibabe, I. (2016). Strategic interviewing to detect deception: Cues to deception across repeated interviews. *Frontiers in Psychology*, *7*, 1702. <https://doi.org/10.3389/fpsyg.2016.01702>
- Nahari, G., Vrij, A., & Fisher, R. P. (2014). The verifiability approach: Countermeasures facilitate its ability to discriminate between truths and lies. *Applied Cognitive Psychology*, *28*, 122–128. <https://doi.org/10.1002/acp.2974>
- Rogers, S. L., Speelman, C. P., Guidetti, O., & Longmuir, M. (2018). Using dual eye tracking to uncover personal gaze patterns during social interaction. *Scientific Reports*, *8*, Article 4271. <https://doi.org/10.1038/s41598-018-22726-7>
- Sakrisvold, M. L., Granhag, P. A., & Mac Giolla, E. (2017). Partners under pressure: Examining the consistency of true and false alibi statements. *Behavioral Sciences & the Law*, *35*, 75–90. <https://doi.org/10.1002/bsl.2275>
- Shaw, D., Vrij, A., Leal, S., & Mann, S. (2014). ‘We’ll take it from here’: The effect of changing interviewers in information gathering interviews. *Applied Cognitive Psychology*, *28*, 908–916. <https://doi.org/10.1002/acp.3072>
- Shaw, D. J., Vrij, A., Leal, S., Mann, S., Hillman, J., Granhag, P. A., & Fisher, R. P. (2013). Expect the unexpected? Variations in question type elicit cues to deception in joint interviewer contexts. *Applied Cognitive Psychology*, *27*, 336–343. <https://doi.org/10.1002/acp.2911>
- Varao Sousa, T. L., Carriere, J. S. A., & Smilek, D. (2013). The way we encounter reading material influences how frequently we mind wander. *Frontiers in Psychology*, *4*, 892. <https://doi.org/10.3389/fpsyg.2013.00892/full>
- Vrij, A., Leal, S., & Fisher, R. P. (2018). Verbal deception and the Model Statement as a lie detection tool. *Frontiers in Psychiatry*, *9*, Article 492. <https://doi.org/10.3389/fpsyg.2018.00492>
- 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. <https://doi.org/10.1111/lcrp.12126>
- Vrij, A., Leal, S., Mann, S., Dalton, G., Jo, E., Shaboltas, A., . . . Houston, K. (2017). Using the model statement to elicit information and cues to deceit in interpreter-based interviews. *Acta Psychologica*, *177*, 44–53. <https://doi.org/10.1016/j.actpsy.2017.04.011>
- Vrij, A., Mann, S., Leal, S., & Fisher, R. (2010). “Look Into My Eyes”: Can an instruction to maintain eye contact facilitate lie detection? *Psychology, Crime, & Law*, *16*, 327–348. <https://doi.org/10.1080/10683160902740633>

Received 4 March 2021

Appendix A:

The 17 valuable details as derived from Shaw et al. (2014).

- 1-Two people are waiting in the meeting room
- 2-Mr. Black enters and explains the purpose of the meeting.

3-Mr. Black then mentions the time pressure, and what will happen if they run out of time.

4-The two others examine the device and hand it back to Mr Black.

5-Mr. Black then details the physical appearance the device.

6-10-Mr. Black then describes the five features of the device.

11-Mr. Black then introduces the first location and where the device would be planted.

12-Mr. Black then explains why the first location would be suitable and its downside.

13-Mr. Black then introduces the second location and where the device would be located.

14-Mr. Black then explains why the second location would be suitable and its downside.

15-Mr Black then introduces the third location before being saying time is nearly up.

16-All members then cast their votes by a show of hands for each location.

17-Mr. Black then declares the result of the vote and the location where the device will be planted.

Appendix B:

The valuable details 17-point scale analysis

Two 2 (Veracity) \times 5 (Model Statement) ANOVAs were conducted with pre-MS valuable details (17-point scale) and post-MS new valuable details (17-point scale) as dependent variables.

The pre-MS valuable details analysis resulted in a significant Veracity main effect, $F(1, 220) = 28.87, p < .001, d = 0.71, 95\% \text{ CI } [0.43, 0.97]$. Truth tellers ($M = 10.75, SD = 3.88, 95\% \text{ CI } [10.04, 11.48]$) reported more valuable details in Phase 1 than lie tellers ($M = 7.99, SD = 3.89, 95\% \text{ CI } [7.28, 8.71]$). The Model Statement main effect, $F(4, 220) = 0.73, p = .573, \eta_p^2 = .01$, and the Veracity \times Model Statement interaction effect, $F(4, 220) = 0.87, p = .484, \eta_p^2 = .003$, were not significant.

The post-MS new valuable details analysis resulted in a significant Veracity main effect, $F(1, 220) = 4.66, p = .032, d = .29, 95\% \text{ CI } [0.02, 0.54]$. Lie tellers ($M = 2.21, SD = 2.21, 95\% \text{ CI } [1.86, 2.57]$) reported more new valuable details in Phase 2 than truth tellers ($M = 1.66, SD = 1.58, 95\% \text{ CI } [1.30, 2.02]$). The Model Statement main effect, $F(4, 220) = 1.49, p = .208, \eta_p^2 = .03$, and the Veracity \times Model Statement interaction effect, $F(4, 220) = 0.18, p = .948, \eta_p^2 = .003$, were not significant. The two nonsignificant interaction effects mean that Hypothesis 4 was not supported.

Appendix C:

Accurate details in the recall of the model statement

An ANOVA was carried out utilizing a 2 (Veracity) \times 4 (Model Statement) design with the number of accurate details reported about the model statement as dependent variable. The no model statement condition was excluded from the analysis. The analysis revealed a significant Model Statement main effect, $F(3, 175) = 6.60, p < .001, \eta_p^2 = .10$. The Veracity main effect $F(1, 175) = 1.16, p = .282, d = 0.16, 95\% \text{ CI } [-0.13, 0.45]$, and the Veracity \times Model Statement interaction effect, $F(3, 175) = 0.20, p = .898, \eta_p^2 = .003$, were not significant. Most accurate details were recalled in the Read Aloud MS condition, see Table A1. It resulted in significantly more accurate details than the Audio MS and Written MS conditions. The Written MS condition resulted in fewer accurate details reported than the Face to Face MS condition. The other comparisons were not significant, all p 's $> .403$.

Table A1. Statistics for the number of accurate details in recalling the model statement as a function of model statement (4 levels)

Variable	Low engagement MS				High engagement MS							
	Audio MS		Written MS		Face to Face MS		Read aloud MS					
	M	SD	M	SD	M	SD	M	SD				
Number of accurate details	9.77 ^{ab}	9.36	7.24, 12.31	7.86 ^a	5.92	5.32, 10.40	12.61 ^{bc}	8.98	10.07, 15.15	15.42 ^c	9.87	12.89, 18.02

Note. Only mean scores with a different superscript differ significantly ($p < .05$) from each other. In the Audio-MS, the interviewee listens to the Model Statement. In the Written MS, the interviewee reads the Model Statement in silence. In the Face to Face MS, the interviewer reads the Model Statement aloud whilst maintaining eye contact with the interviewee. In the Read Aloud MS, the interviewee reads the Model Statement aloud.