Using Unexpected Questions

to Elicit Information and Cues to Deceit in Interpreter-Based Interviews

Aldert Vrij

Sharon Leal

Samantha Mann

Ronald P. Fisher

Gary Dalton

Eunkyung Jo

Alla Shaboltas

Maria Khaleeva

Juliana Granskaya

Kate Houston

Author Note

Aldert Vrij, Sharon Leal, Samantha Mann and Gary Dalton, Department of Psychology, University of Portsmouth; Ronald P. Fisher, Department of Psychology, Florida International University, Eunkyung Jo, Department of Psychology, Hallym University; Allla Shaboltas, Maria Ivanova, Juliana Granskaya, Department of Psychology, St. Petersburg State University; Kate Houston, Department of Public Affairs and Social Research, Texas A&M International University.

This work is funded by the High-Value Detainee Interrogation Group, DJF-15-1299-V-0010271 awarded to the University of Portsmouth (UK). Any opinions, findings, conclusions, or recommendations expressed in this article are those of the authors and do not necessarily reflect the views of the U.S. Government.

Using unexpected questions in Interpreter-based Interviews

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

Using unexpected questions in Interpreter-based Interviews

Abstract

We examined whether speech-related differences between truth tellers and liars are more profound when answering unexpected questions than when answering expected questions. We also examined whether the presence of an interpreter affected these results. In the experiment, 204 participants from the USA (Hispanic participants only), Russia, and the Republic of Korea were interviewed in their native language by a native-speaking interviewer or by a British interviewer through an interpreter. Truth tellers discussed a trip they had made during the last twelve months; liars fabricated a story about such a trip. The key dependent variables were the amount of information provided and the proportion of all statements that were complications. The proportion of complications distinguished truth tellers from liars better when answering unexpected than expected questions, but only in interpreter-absent interviews. The number of details provided did not differ between truth tellers and liars or between interpreter-absent and interpreter-present interviews.

Keywords: interpreter, non-native speakers, expected vs unexpected questions, information-gathering, deception

Using Unexpected Questions

to Elicit Information and Cues to Deceit in Interpreter-Based Interviews

Meta-analyses have shown that truth tellers and liars differ from each other in terms of speech content, but that these differences are usually small (DePaulo et al., 2003; DePaulo & Morris, 2004; Vrij, 2008). Therefore, around a decade ago, researchers started to examine whether verbal differences between truth tellers and liars could be elicited or enhanced through theory-based interview methods (Vrij & Granhag, 2012, 2014). One strand of this research is called cognitive lie detection and involves three key elements: Imposing cognitive load, encouraging interviewees to say more, and asking unexpected questions (Vrij, Fisher, & Blank, 2017; Vrij, Leal, Mann, Vernham, & Brankaert, 2015). Cognitive lie detection appears to elicit cues to deception (Vrij, Fisher, Blank, Leal, & Mann, 2016) and to facilitate lie detection (Vrij, Fisher, & Blank, 2017). In the current experiment, we elaborate on the asking unexpected questions element.

A consistent finding in the deception literature is that liars prepare themselves for interviews by preparing answers to questions they expect to be asked (Hartwig, Granhag, & Strömwall, 2007; see Tedeschini (2012) for a description of a real-world case). This strategy of preparing answers for possible questions makes sense. Planning makes lying easier and planned lies typically contain fewer cues to deceit than spontaneous lies (DePaulo et al., 2003; Sporer & Schwandt, 2006). However, preparing for answers will be fruitful only if liars correctly anticipate which questions will be asked. Investigators can thus thwart liars' preparations by asking questions that liars do not anticipate. Liars should have more difficulty than truth tellers in answering the unexpected questions as they have to produce spontaneous answers to these questions. As a result, liars' answers may be less detailed than truth tellers'

answers when answering unexpected questions. Liars may have an easier task, and hence provide more detailed answers, if they are asked questions that they are prepared for. In support of the unexpected-questions approach, it has been found that more pronounced differences between truth tellers and liars appeared in answers to unexpected questions than in answers to expected questions (Roos af Hjelmsäter, Ohman, Granhag, & Vrij, 2014; Sooniste, Granhag, Knieps, & Vrij, 2013; Vrij et al., 2009; Warmelink, Vrij, Mann, Jundi, & Granhag, 2012).

Most of the studies examining the unexpected questions approach used a generic measure: Total amount of information. In the present experiment, we break down this generic measure (total amount of information) into components that we believe are more sensitive to the different verbal strategies used by truth tellers and liars. In brief, we expected truth tellers to provide stories that include non-essential details that make the story more complex (complications). By comparison, we expected liars to provide details that are based on common knowledge, or to justify why they cannot provide certain types of information (self-handicapping strategies).

A complication is a reported activity or event that was not expected or planned ("The event was cancelled due to bad weather") (Steller & Köhnken, 1989; Vrij, Leal, et al., 2017). Complications are more likely to occur in truthful statements than in deceptive statements (Amado et al., 2015; Vrij et al., 2017). Making up complications requires imagination and liars may not have adequate imagination to make up facts (Vrij, 2008). In addition, liars prefer to keep their stories simple (Hartwig, Granhag, & Strömwall, 2007) but adding complications makes the story more complex.

Common knowledge details refer to strongly invoked stereotypical information about events ("We visited the Louvre museum where was saw the Mona Lisa") (Vrij, Leal, et al.,

2017). Liars are more likely to include common knowledge details in their statements than truth tellers (Köhnken, 2004; Sporer, 2016; Volbert & Steller, 2014; Vrij, Leal, et al., 2017). Truth tellers, who have personal experiences of an event, are likely to report such unique experiences, and when they do so the statement is no longer scripted. If liars do not have personal experiences of the event they report, they then will draw upon general knowledge to construe the event. Even if they have personal experiences of the event, they may not report them due to their desire to keep their stories simple.

Self-handicapping strategies relate to an admission of not being able to provide information followed by a justification ("I can't remember, it was a while ago when this happened", "Nothing unexpected happened, I am a very organised person") (Vrij, Leal, et al., 2017). Liars are more likely to include self-handicapping strategies in their statements than truth tellers (Vrij, Leal, et al., 2017). For liars, who are inclined to keep stories simple, not having to provide information is an attractive strategy. However, liars are also concerned about their credibility and believe that admitting lack of knowledge and/or memory appears suspicious (Ruby & Brigham, 1998). A potential solution is to provide a justification for the inability to provide information. Note that the justification is crucial for a self- handicapping strategy. The statement 'I can't remember' itself is not a self-handicapping strategy, it is called admitting lack of memory and part of the Criteria-Based Content Analysis tool (Amado et al., 2015).

Examining complications, common knowledge details and self-handicapping strategies has two advantages compared to examining total details. First, the three separate measures should be more effective than the one, combined measure (total details) in discriminating truth tellers from liars. Both truth tellers and liars must provide some details in interviews, with truth tellers perhaps providing somewhat more total details than liars.

However, the combined measure is composed of some details that are more likely to be reported by truth tellers (complications), but other details that are more likely to be reported by liars (common knowledge details and self-handicapping strategies). Combining the three separate measures into one overall score should therefore dilute the diagnostic value of each measure taken separately. Second, we could examine *the proportion of cues to truthfulness*, i.e. the proportion of complications (complications / (complications + common knowledge details + self-handicapping strategies), which is a within-subjects measure. Within-subjects measures are preferred by practitioners (Vrij, 2016) and scholars (Nahari & Pazuelo, 2015; Nahari & Vrij, 2014; Nahari, in press), amongst other reasons, because it creates the opportunity to design cut-off scores rules (e.g., "A statement is considered to be truthful when it has more complications than common knowledge details and self-handicapping strategies combined").

This difference in effectiveness between the total detail variable and the proportion of complications variable may occur particularly in answering unexpected questions. Liars have no option other than to report spontaneous answers to the unexpected questions and spontaneous questions in particular may lack complications and may include common knowledge details and self-handicapping strategies.

The use of an Interpreter

Interpreters are frequently used in investigative interviews but experimental research examining their effect on eliciting information and cues to deceit is scarce. We are aware of only six experimental studies published in this area (Ewens et al., 2016a, b, c, 2017; Houston, Russano, & Ricks, 2017; Vrij, Leal et al., 2017, in press). A consistent finding in these experiments is that when interviewees speak in their own language through an interpreter, fewer details are provided than when interviewees speak in their own language without an

interpreter (Ewens et al., 2016a, b, c, 2017; Vrij, Leal, et al., 2017). Ewens et al. (2016b) found evidence for two explanations. First, interpreters did not translate every detail the interviewee gave. The interpreters implemented a long consecutive style of interpretation in which the interpreter translated segments of talk, a style frequently used in interviews (Viezzi, 2012), including in intelligence interviews (Department of the Army, 2006). Remembering all the details an interviewee conveys is difficult in such a situation and the interpreters failed to translate about 10% of the details given by the interviewees. Second, interviewees actually reporteded less with an interpreter present (Ewens et al., 2016b). Perhaps the interpreter's interruptions disrupted the interviewee's train of thought and made memory retrieval more difficult (Vrij, Hope, & Fisher, 2014; Nelson & Goodmon, 2003) or perhaps the interviewees decided to be concise when an interpreter was present given the extra time it takes to communicate through an interpreter (similar to people being more concise when talking to a hard-of-hearing person).

We explored the effect of an interpreter on complications, common knowledge details and self-handicapping strategies. Vrij, Leal, et al. (2017) found no effect of an interpreter on these variables, which suggest that examining them could be equally effective for lie detection purposes in interpreter-absent and interpreter-present interviews. However, we do not recommend practitioners and policy makers to rely on the findings of a single study, also because we cannot rule out at this stage that some of these variables could be affected by the presence of an interviewer. That is, perhaps interviewees are inclined to be more concise in the presence of an interpreter, which could result in fewer complications and more common knowledge details. Or perhaps interpreters leave out complications and replace them with common knowledge details in their translations. Or perhaps interpreter-present interviews

give interviewees more time to think when the interpreter translates a response, which could result in fewer common knowledge details.

Hypotheses

Truth tellers will report more details, more complications, fewer common knowledge details and fewer self-handicapping strategies than liars (Hypothesis 1a). From Hypothesis 1a follows that the proportion of cues to truthfulness (e.g., proportion of complications) will be higher in truth tellers than in liars (Hypothesis 1b). The differences between truth tellers and liars predicted in Hypothesis 1a, b will be more pronounced when answering unexpected questions than when answering expected questions (Hypothesis 2). The proportion of complications will be more effective in discriminating truth tellers from liars than total details, particularly when answering unexpected questions (Hypothesis 3). Less information will be provided in interviewer-absent than in interviewer-present interviews (Hypothesis 4).

Method

Participants

A total of 204 participants (52 males, 147 females and five not indicated) took part in the study. All participants were university students and their age ranged from 18-39 years with an average age of M = 22.00 years (SD = 3.44). Participation took place in three different universities in the Republic of Korea, Russia and USA, and the background of the participants was Korean (n = 80), Russian (n = 79) and Hispanic (n = 45), respectively.

Procedure

The study consisted of two parts. The first part related to the unexpected – expected questions, the topic of this paper. The second part related to the effect of sketching when narrating on eliciting information and cues to deceit (Vrij et al., in press). The two parts are independent from each other (i.e., the sketching manipulation was introduced after the

unexpected – expected questions part of the study) and the findings are therefore reported in two different papers.

Procedure

Recruitment, pre-condition selection form, preparation and pre-interview questionnaire.

We used the same procedure as Vrij, Leal, et al. (2017). Participants were recruited via an advert on the university intranets and advertisement leaflets. The advert explained that the experiment would require participants to tell the truth or lie about a trip away that they may (or may not) have taken within the last year. We decided upon "last year" so that truth tellers would still remember many details about their trip and liars could not easily say 'I can't remember' when answering the questions. On arrival to the corresponding university, participants received a participant information sheet and signed an informed consent form. They then completed a pre-condition selection form that contained six cities that the researchers thought the participants may have visited during the past year. (Different cities were used for the three different countries.) The participants were also asked to write down the names of two other cities they had visited during the past year.

For each city the participants indicated (a) whether they had been there during the last twelve months, (b) when they had been there during the last twelve months, (c) for how long they stayed there, and (d) whether they have lived there. For truth tellers the experimenter selected a city where the participant had stayed during the last twelve months for at least two nights but had never lived there. Truth tellers were informed that they would be interviewed about this selected city (city X) and asked to answer the questions truthfully. For liars, the experimenter selected a city where the participant had never been in his/her life before. Liars were informed that they would be interviewed about city X (taken from the list of cities truth

tellers were interviewed about) and that they had to pretend to have stayed there for at least two nights during a trip made during the last twelve months. Across all 102 truth tellers, more than twenty cities were used. Liars were also interviewed about these cities.

Truth tellers and liars were then left with a computer with internet access and told they had twenty minutes to prepare for their interview, or to inform the experimenter if they were ready before that time. The participants were told that they were allowed to make notes while doing their research. Truth tellers and liars were told that it was important to be convincing because, if they did not appear convincing, they would be asked to write a statement about what they told the interviewer in the interview. In a pre-interview questionnaire (which was translated in the native language of the participant) the truth tellers and liars rated their thoroughness of preparation via three items: (1) shallow to (7) thorough; (1) insufficient to (7) sufficient; and (1) poor to (7) good. The answers to the three questions were averaged (Cronbach's alpha = .81) and the variable is called 'preparation thoroughness'. They were also asked whether they thought they were given enough time to prepare themselves with the following question: 'Do you think the amount of time you were given to prepare was: (1) insufficient to (7) sufficient.

Experimental conditions.

Participants were allocated randomly to one of the four experimental cells. A total of 102 participants were allocated to the truth condition and 102 to the lie condition; and 101 to the Interpreter-absent condition and 103 to the Interpreter-present condition.

In total, three interpreters were used in the study, one in each country. These were the same interpreters as used in Vrij, Leal, et al. (2017). The Korean and Hispanic interpreters were professional interpreters; the Russian interpreter spoke fluent English and had a Masters degree that included English language. The interpreters used a long consecutive interpreter

style: They interpreted chunks of information uttered by the participant rather than interpreting sentence by sentence, and took notes during the interview.

In the interpreter condition, one British interviewer was used, whereas in the non-interpreter condition one Russian, one Korean and one Hispanic interviewer were used. The British, Russian and Korean interviewers were the same as in Vrij, Leal, et al. (2017). Prior to the experiment the British interviewer (who is a very experienced interviewer and has interviewed in many experiments before) instructed the other interviewers how to conduct the interview. They were instructed to be friendly and not to interrupt the interviewee. Several practice sessions took place until the British interviewer was satisfied with the interview style of the interviewer. That is, she was satisfied with the demeanour of the interviewers (appeared friendly) and the opportunities they gave to the interviewees to talk (no interruptions). To assess consistency in interview style between the interpreter and non-interpreter conditions, participants were asked to assess in a post-questionnaire the rapport they experienced with the interviewer.

The interview.

Prior to the interview the experimenter told the interviewer about which city to interview the participant. The interviewer was unaware of the veracity status of the participant. To make the interviewee feel comfortable and to avoid floor effects in establishing rapport interviewees were offered a glass of water from the interviewer, as offering something helps rapport building (reciprocation principle, Cialdini, 2007).

The interviewer started by saying: “I understand from my colleague that you have visited _____ I would now like to ask you some questions about this visit.” The interview contained five questions. Two questions about the actual trip we thought would be expected: “Could you please tell me in as much detail as possible everything you did when

you were at _____ from the moment you arrived to the moment you left?" (Question 1); "Could you please tell me in as much detail as possible about your accommodation where you stayed, including the location and address if you can remember, and what (tourist) attractions were nearby?" (Question 2). Three questions were asked about the planning of the trip which we thought to be unexpected: "Could you please tell me in as much detail as possible everything you did to plan this trip E.g., organising transport, accommodation, where to visit and so on.?" (Question 1), "Could you please talk us through the order of that planning? What did you plan first, second etc. and what was most important for you to organise correctly?" (Question 2) and "Could you please tell me in as much detail as possible if there was anything unexpected that happened or perhaps something that didn't go to plan when organising this trip?" (Question 3). The questions about the actual trip and planning were counterbalanced, so that 100 participants were asked the two actual trip questions first and 104 participants were asked the three planning questions first. We examined the effect of Order on the main dependent variables (total details, complications, common knowledge details, self-handicapping strategies and proportion of complications) in a 2 (Veracity) x 2 (Interpreter) x 2 (Order) MANOVA. At a multivariate level, none of the effects that included Order were significant, all $F's < 1.183$, all $p's > .305$. We therefore did not include Order as a factor in the subsequent analyses.

The interviews were video (interviewees only) and audio recorded and the English speech in the audiotapes was subsequently transcribed. In other words, in the interpreter conditions the speech from the interpreter was transcribed and analysed. We did this because it is this speech that interviewers will understand in real life interviews with interpreters. In a study in which in the Interpreter-present interviews both the interviewee's and interpreter's speech were both transcribed, coded and analysed, virtual identical findings in the two data

sets emerged in terms of (i) eliciting information and (ii) the difference in reported details between truth tellers and liars (Ewens et al., 2016b).

Post-interview questionnaire.

The post-interview questionnaire was translated in the native language of the participant. Rapport was measured via the nine items Interaction Questionnaire (Vallano & Schreiber Compo, 2011). Participants rated the interviewer on 7-point scales ranging from [1] not at all to [7] extremely on nine characteristics such as *smooth, bored, engrossed,* and *involved* (Cronbach's alpha = .81). The post-questionnaire also measured participant's motivation to perform well during the interview (with a single item).

After questions about rapport and motivation, the two actual trip questions and three planning questions were listed. Participants reported for each of these five questions the extent to which they had expected these questions (on 7-point Likert scales ranging from 1 = not at all to 7 = totally). They were also asked to indicate the extent to which they told the truth in answering these questions on 11-point Likert scales ranging from 0% to 100%.

Coding

All coders were blind to the hypotheses and Veracity condition.

Detail.

The coders were taught the coding scheme by the first author who has more than twenty years of experience in coding details. A coder first read the transcripts and coded each detail in the interview. To give an example, the following answer contains ten details: "I have an aunt there, ah – living in a house close to the lake. We went in my truck, ah, five of us, also my sister's boyfriend and my boyfriend". Each detail in the actual trip section and each detail in the planning of the trip section was coded only once; thus repetitions were not coded. The same detail reported in the actual trip section and in the planning section would

not be considered a repetition and would be coded twice, but this situation never occurred. A second coder coded a random sample of 40 transcripts. Inter-rater reliability between the two coders, using the two-way random effects model measuring consistency, was high (Single Measures ICC = .90).

Two coders coded independently from each other complications, common knowledge details and self-handicapping strategies. A complication is an activity or event that someone describes which was not expected or planned (Steller & Köhnken, 1989; Vrij, Leal, et al., 2017). Three examples of complications are: i) “We couldn't stay at his place because he lived in dormitory”; ii) “Two weeks before the competition, I broke my jaw - I do figure skating- and they couldn't find any another person to go to the competition instead of me. So I had to go to this competition with my injury”, and iii) “The first night, we stayed at that guesthouse because my acquaintance worked there, and for the second night, we went to that hotel because we wanted to go there”.

Common knowledge details refer to strongly invoked stereotypical knowledge about events (Vrij, Leal, et al., 2017). Three examples of common knowledge details are: i) “And we just erm – er walked er around and saw this er medieval architecture which is like quite common for, er, different European cities”; ii) “And we also visited like a standard places like er the cathedral in Prague”, and iii) “I looked for information on the internet, ah where – and read some mm places to go”.

Self-handicapping strategies relate to an admission of not being able to provide information followed by a justification (Vrij, Leal, et al., 2017). Three examples of self-handicapping strategies are: i) “I don't remember the, er, exact address, because I came there by er tourist minibus, so I – I didn't quite remember the city, because we weren't walking much”, ii) “We er travelled um together with my mum, er – and um, so the – the room wasn't

quite memorable for me, maybe also because there was like no er minibar there as it is in Russia" and iii) "For the guesthouse I didn't look for any information because (name city) is very famous, so I – I guess if I go there, there must be some places for me".

Inter-rater reliability between the two coders, using the two-way random effects model measuring consistency, was high for complications (Average Measures, Intraclass correlation coefficient, ICC = .88) and self-handicapping strategies (Average Measures ICC = .81) and satisfactory for common knowledge details (Average Measures ICC = .66). Disagreements were resolved between the two coders. All disagreements occurred because one coder failed to spot a particular complication, common knowledge detail or self-handicapping strategy. To calculate the proportion of complications a total score was computed (number of complications + number of common knowledge details + number of self-handicapping strategies) and the number of complications was divided by this total score. Scores > .50 indicate that the participants reported more complications than common knowledge details or self-handicapping strategies, whereas scores < .50 indicate that the participants reported more common knowledge details or self-handicapping strategies than complications.

Number of days spent away.

One coder also counted the number of days spent away the interviewees described in their interviews. A second coder coded a random sample of 40 transcripts. Inter-rater reliability between the two coders, using the two-way random effects model measuring consistency, was good (Single Measures ICC = .79).

Results

Preparation thoroughness and preparation time

Two oneway ANOVAs with Veracity as factor and preparation thoroughness and preparation time as dependent variable revealed that truth tellers (*M* = 4.92, *SD* = 1.07, *95% CI* [4.71, 5.1]) rated their preparation as more thorough than liars (*M* = 4.17, *SD* = 0.99, *95%* CI [3.97, 4.37]), $F(1, 202) = 26.82$, $p < .001$, $d = 0.73$. Truth tellers (*M* = 5.93, *SD* = 1.49, *95% CI* [5.61, 6.26]) also believed more than liars (*M* = 4.93, *SD* = 1.79, *95% CI* [4.61, 5.25]) that they were given sufficient time to prepare themselves for the interview, $F(1, 202) = 18.83$, $p < .001$, $d = 0.75$. We introduced preparation thoroughness and preparation time as covariates in the hypotheses testing analyses.

Manipulation Checks

Rapport.

An ANOVA with Interpreter as factor and rapport with the interviewer as dependent variable did not show a difference in rapport between the two conditions, $F(1, 202) = 2.77$, $p = .097$, $d = 0.23$. The grand mean revealed that the interviewees perceived the rapport with the interviewer as very good (*M* = 5.60, *SD* = .82 on a 7-point scale).

Expectedness of the questions.

A 2 (Veracity: truth vs lie) X 2 (Interpreter: present vs absent) X 2 (Type of Question: planning vs actual trip) with Veracity and Interpreter as between-subjects factors and Type of Question as within-subjects factor revealed a main effect for Type of Question, $F(1, 200) = 180.90$, $p < .001$, $d = 1.04$ (0.82,1.24). The questions about the actual trip were more expected (*M* = 5.02, *SD* = 1.34, *95% CI* [4.83, 5.20]) than the questions about planning (*M* = 3.57, *SD* = 1.43, *95% CI* [3.37, 3.78]). All other effects were not significant, all *F's* < 1.75, all *p's* > .225. From now on we refer to expected questions when referring to the actual trip questions and to unexpected questions when referring to the planning questions.

Motivation, number of days spent away discussed, percentage truth telling in the interview, and time since the trip was made.

Analyses of variance were carried out utilising a 2 (Veracity) X 2 (Interpreter) design with motivation, number of days spent away discussed, and percentage truth telling in the interview as dependent variables. The grand mean showed that the participants were motivated to perform well during the interview (M = 4.01, SD = .72 on a 5-point scale). The ANOVA revealed no significant main or interaction effects for motivation, all F's < 3.37, all p's > .067. Truth tellers discussed on average more days spent away in their interview (M = 4.08, SD = 2.92, *95% CI* [3.65, 4.52]) than liars (M = 3.03, SD = 1.21, *95% CI* [2.59,3.46]), $F(1, 200) = 11.38$, $p = .001$, $d = 0.46$ (0.18,0.74). All the other effects were not significant, all F's < 1.25, all p's > .067. We included 'the number of days spent away discussed' as a covariate in the hypotheses testing analyses. We did not predict any effects for nationality of the participants. In addition, not only nationality but also interviewer and interpreter varied along with nationality. We wanted to control for the possible effects of this cluster of factors, and therefore included *site* as a covariate in the hypotheses testing analyses.

Truth tellers told the truth significantly more (M = 97.45, SD = 8.67, *95% CI* [94.21,100.63]) than liars (M = 14.20, SD = 21.73, *95% CI* [10.86,17.28]) about the expected questions, $F(1, 200) = 1311.18$, $p < .001$, d = 4.96 (4.40,4.51). Truth tellers also told the truth significantly more (M = 98.42, SD = 4.23, *95% CI* [95.11,101.64]) than liars (M = 18.98, SD = 23.48, *95% CI* [15.58,22.12]) about the unexpected questions, $F(1, 200) = 1140.68$, $p <$ .001, d = 4.64 (4.11,5.16). All the other effects were not significant, all F's < 2.78, all p's > .096.

On average truth tellers made their trip made M = 5. 48 months prior to the interview (SD = 2.96). This variable was not correlated with any of the main dependent variables in the

study (total details, complications, common knowledge details, self-handicapping strategies or proportion of complications), all *r's* < .07, all *p's* > .50).

Hypothesis Testing

A 2 (Veracity) X 2 (Interpreter) MANCOVA was conducted with the ten variables listed in Table 1 as dependent variables and preparation thoroughness, preparation time, site, and number of days spent away discussed during the interview as covariates.[1] The analysis revealed significant main effects for Veracity, $F(10, 187) = 5.597$, $p < .001$, *partial eta*^{2} = .23) and Interpreter, $F(10, 187) = 4.468$, $p < .001$, *partial eta*^{2} = .19, and a significant Veracity x Interpreter interaction effect, $F(10, 187) = 1.914$, $p = .046$, *partial eta*^{2} = .10. The univariate Veracity main effects are presented in Table 1.

Tables 1 about here

Compared to liars, truth tellers reported significantly more complications and fewer common knowledge details and self-handicapping strategies[2] when answering both the unexpected and expected questions. However, truth tellers and liars did not differ from each other in the total number of details they reported. Thus, Hypothesis 1a is supported, except for the variable total details. In support of Hypothesis 1b, the proportion of complications was significantly higher for truth tellers than for liars in answering both the unexpected and expected questions. The difference in proportion scores between truth tellers and liars was similar in answering the unexpected questions ($d = 0.67$) and in answering the expected questions ($d = 0.52$), with a substantial overlap in the two 95% confidence intervals, rejecting Hypothesis 2 (but see below for differential effects for Interpreter absent and Interpreter-present interviews).

Regarding the Interpreter main effect, significant univariate effects emerged for common knowledge details in answering both the unexpected, $F(1, 196) = 9.124$, $p = .003$, *d*

= 0.36 (0.08,0.64) and expected questions, $F(1, 196) = 10.271$, $p = .002$, $d = 0.40$ (0.12,0.68). In answering the unexpected questions, more common knowledge details were reported in Interpreter-absent (*M* = 1.14, *SD* = 0.91, *95% CI* [0.99,1.31]) than in Interpreter-present interviews (*M* = 0.81, *SD* = 0.89, *95% CI* [0.65,0.96]). The same pattern of results emerged in the answering the expected questions with more common knowledge details being reported in Interpreter-absent (*M* = 6.25, *SD* = 2.85, *95% CI* [5.75,6.85]) than in Interpreter-present interviews (*M* = 5.08, *SD* = 2.88, *95% CI* [4.49,5.58]). In Hypothesis 4 it was predicted that fewer details would be reported in Interpreter-present than in Interpreter-absent interviews. Since there was no difference in the number of details reported in Interpreter-present and Interpreter-absent interviews, Hypothesis 4 is rejected.

Regarding the Veracity x Interpreter interaction effect, three significant univariate effects emerged. First, a significant effect for common knowledge details in answering unexpected questions emerged, $F(1, 196) = 8.508$, $p = .004$, *partial eta*^{2} = 0.04. In Interpreter-absent interviews, liars reported more common knowledge details in answering the unexpected questions (*M* = 1.65, *SD* = 0.80, *95% CI* [1.46,1.93]) than truth tellers (*M* = 0.65, *SD* = 0.71, *95% CI* [0.39,0.84]), $F(1, 95) = 38.135$, $p < .001$, $d = 1.30$ (1.00,1.60). In Interpreter-present interviews, the difference between liars (*M* = 0.98, *SD* = 0.97, *95% CI* [0.74,1.21]) and truth tellers (*M* = 0.62, *SD* = 0.75, *95% CI* [0.38,0.87]) in reporting common knowledge details just failed to reach significance, $F(1, 97) = 3.808$, $p = .054$, $d = .41$ (0.13,0.69). Note that this differential pattern of Veracity findings is caused by liars: They reported more common knowledge details in Interpreter-absent (*M* = 1.65) than in Interpreter-present interviews (M = 0.98); truth tellers reported the same number of common knowledge details in Interpreter-absent (*M* = 0.65) and Interpreter-present interviews (*M* = 0.62).

Second, a significant Veracity x Interpreter interaction effect for proportion of complications in answering unexpected questions emerged, $F(1, 196) = 4.955$, $p = .027$, *partial eta*$^2 = 0.032$. In Interpreter-absent interviews, truth tellers ($M = 0.74$, $SD = 0.28$, *95% CI* [0.65,0.80]) provided a higher proportion of complications in answering the unexpected questions than liars ($M = 0.46$, $SD = 0.22$, *95% CI* [0.40,0.55]), $F(1, 95) = 19.689$, $p < .001$, $d = 1.11$ (0.80,1.39). In Interpreter-present interviews, truth tellers obtained a similar proportion of complications score in answering the unexpected questions ($M = 0.69$, $SD = 0.31$, *95% CI* [0.59,0.77]) as liars ($M = 0.59$, $SD = 0.29$, *95% CI* [0.51,0.68]), $F(1, 97) = 1.907$, $p = .170$, $d = .33$ (0.05,0.60). Note that in Interpreter-absent interviews the Veracity *d*-score for proportion of complications for unexpected questions is substantially higher ($d = 1.11$ [0.80,1.39]) than the earlier reported Veracity *d*-score for the proportion of complications for expected questions ($d = 0.52$ [0.23,0.79]). This supports in Interpreter-absent interviews Hypothesis 2, that the proportion of complications variable would be most diagnostic in answering the unexpected questions.

The combination of findings that (1) the total details variable did not discriminate between truth tellers and liars, and (2) the proportion of complications variable was most diagnostic in answering the unexpected questions (in interpreter-absent interviews), supports in interpreter-absent interviews Hypothesis 3 in which it was predicted that the proportion of complications variable would be more effective in discriminating between truth tellers and liars than the total details variable, particularly in answers to the unexpected questions..

Third, a significant Veracity x Interpreter interaction effect for total details in answering the expected questions emerged, $F(1, 196) = 6.037$, $p = .015$, *partial eta*$^2 = 0.03$. Neither of the two simple effects tests revealed Veracity differences, but the effects were in opposite directions, causing the interaction effect to be significant. In Interpreter-absent

interviews, no difference emerged in the number of details reported in answering the expected questions between truth tellers (*M* = 77.78, *SD* = 46.24, *95% CI* [63.52,84.45]) and liars (*M* = 60.47, *SD* = 24.60, *95% CI* [53.78,75.43]), $F(1, 95) = 1.331$, $p = .252$, $d = 0.49$. Also in Interpreter-present interviews, no difference emerged in the number of details reported in answering the expected questions between truth tellers (*M* = 74.10, *SD* = 36.92, *95% CI* [63.55,85.13]) and liars (*M* = 81.21, *SD* = 36.95, *95% CI* [70.52,91.44]), $F(1, 97) = 0.772$, $p = .398$, $d = .27$.

Discriminant Analyses

Table 2 about here

We conducted a series of discriminant analyses to distinguish between truthful and deceptive interviewees in the unexpected and expected questions. In all cases, the objective group belonging (truthful versus deceptive) was the classifying variable, with the number of reported details or the proportion of complications as predictors. We present the cross-validated 'leave-one-out' results. Cross-validation assesses the accuracy of a statistical model across different samples, an important step in generalisation (Field, 2009). Since we found differences between Interpreter-absent and Interpreter-present interviews in the analyses just described, we ran these analyses for the two interpreter conditions separately.

Interpreter-absent interviews.

Regarding Interpreter-absent interviews (Table 2a), significant discriminant functions emerged for the proportion of complications predictors in both the unexpected questions and expected questions analyses. In addition, the truth/lie classifications based on the proportion of complications predictor were more accurate when answering the unexpected questions (71.13% total accuracy) than when answering the expected questions (64.40% accuracy).

Total details emerged as a predictor in the expected questions analysis only, which was not predicted. However, further analyses showed that this finding was an artefact. When we examined the difference in reporting total details between truth tellers and liars in Interpreter-absent interviews with an ANCOVA with Veracity as factor and preparation thoroughness, preparation time, site, and number of days spent away discussed during the interview as covariates, no significant effect emerged, $F(1, 95) = 1.331$, $p = .252$, $d = 0.47$ (0.06,0.85). Since the ANCOVA takes the covariates into account whereas the discriminant analysis does not, we can conclude that the significant effect was caused by the covariates. Indeed, the effect for the covariate number of days spent away discussed on reporting detail was significant, $F(1, 95) = 15.506$, $p < .001$, *partial eta*$^2 = .14$. Thus, the proportion of complications was more effective in discriminating truth tellers from liars than total details, particularly when answering unexpected questions. This supports Hypothesis 3.

We also calculated several within-subjects measures because they allow for decision rules to be formulated. We designed three decision rules and tested their ability to discriminate between truth tellers and liars. All three decision rules relate to a cut-off score of .50. We believe that this is a logical turning point as scores > .50 indicate that the interviewee reported *more cues of truthfulness* (complications) than cues to deceit (common knowledge details and self-handicapping strategies) whereas scores < .50 indicate that the interviewee reported *more cues to deceit* than cues to truthfulness.

The first decision rule we tested was for the answers to the unexpected questions. We tested the decision rule "An interviewee is classified as a truth teller if his/her proportion of complications score > .50 and classified as a liar if his/her proportion of complications score < .50 ". This means that interviewees with a proportion of complications score of .50 were treated as inconclusives. This decision rule resulted for the unexpected questions in a 73.9%

correct classifications rate, with a high percentage of truth tellers (87.2%) classified correctly (13 inconclusives, 13% of the sample). When we used the same decision rule for the answers to the expected questions (9 inconclusives, 9% of the sample), we obtained a 70.7% accuracy rate, with a high percentage of liars (93.5%) classified correctly. Note that the high accuracy rates for truth tellers in the unexpected questions analysis means that many truth tellers obtained a *high* proportion of complications when answering *unexpected* questions. Similarly, the high accuracy rates for liars in the expected questions analysis means that many liars obtained a *low* proportion of complications when answering the *expected* questions.

Finally, we designed a decision rule that combined the unexpected and expected questions. Interviewees who obtained >.50 proportion-of-complications scores on both the unexpected and expected questions indices were classified as truth tellers, whereas interviewees who obtained <.50 proportion-of-complications scores on both the unexpected and expected questions indices were classified as liars. This resulted in a high total accuracy rate (88.2%), with high accuracy rates for both truth tellers (84.0%) and liars (92.3%). However, n = 50 (50% of the sample) of the interviewees were treated as inconclusives.

Interpreter-present interviews.

Regarding interpreter-present interviews (Table 2b), no significant discriminant function emerged when either total details or the proportion of complications were included as predictors. However, significant discriminant functions emerged in two of the three analyses based on decision rules. The decision rule “An interviewee is classified as a truth teller if his/her proportion of complications score > .50 and classified as a liar if his/her proportion of complications score < .50 ” resulted for the unexpected questions in a 63.2% correct classifications rate, with a relatively high percentage of truth tellers (78.7%) classified

correctly (16 inconclusives, 16% of the sample). When we used the same decision rule for the answers to the expected questions no significant discriminant function emerged.

Finally, we ran a discriminant analysis based on the decision rule that combined the unexpected and expected questions: Interviewees who obtained >.50 proportion of complications scores on both the unexpected and expected questions indices were classified as truth tellers, whereas interviewees who obtained <.50 proportion of complications scores on both the unexpected and expected questions indices were classified as liars. This resulted in a 70.2% total accuracy rate. However, $n = 56$ (54% of the sample) interviewees were treated as inconclusives.

Discussion

We found that unexpected questions were more effective for lie detection purposes than expected questions. Compared to expected questions, the unexpected questions resulted in more pronounced differences in speech content between truth tellers and liars and in a better ability to discriminate between them. These findings provide support for asking unexpected questions.

In our experiment the expected and unexpected questions referred to different aspects of a trip the participants allegedly made. The unexpected questions approach can also be used in different ways. For example, it can be used as a two-steps approach, with asking expected questions first followed by unexpected questions about the same aspect of the event. In that case, someone can examine the amount of additional information provided in the unexpected questions. Colwell et al., who used this approach, found that truth tellers provided more additional details than liars (Colwell, Hiscock-Anisman, & Fede, 2013).

In the present experiment we also found that, compared to liars, truth tellers reported more complications, fewer common knowledge details and fewer self-handicapping

strategies. As a result, the proportion of complications was higher for truth tellers than liars. This replicates Vrij, Leal, et al.'s (2017) new within-subjects measures approach of examining speech content and thus contributes to the robustness of this approach. However, we also found that these findings were affected by the absence or presence of an interpreter. Common knowledge details and the proportion of complications differentiated truth tellers from liars in Interpreter-absent interviews but not Interpreter-present interviews. The difference in findings was due to liars providing more common knowledge details in Interpreter-absent interviews than in Interpreter-present interviews. Interviewees have more opportunity to think during an Interpreter-present than during an Interpreter-absent interview because they can think when the interpreter translates their responses. Perhaps during this enhanced thinking time they fabricated details that made their replies less scripted.

The benefit of using within-subjects measures is that decision rules of how to classify truth tellers and liars can be developed. In the interpreter-absent interviews, we achieved high accuracy rates in correctly classifying truth tellers and liars based on decision rules, particularly when both unexpected and expected questions were taken into account (88.2% total accuracy rates), but this decision rule also led to many inconclusive results (50 out of 101 interviewees were judged as inconclusive). Although practitioners in all likelihood would be disappointed with so many inconclusive results, we believe that treating a case as inconclusive is better than just guessing whether that person is telling the truth or lying. The benefit of introducing an inconclusive category is that it distinguishes between those inconclusive cases and the remaining cases. Our findings showed that investigators can be accurate in assessing the veracity status in these remaining cases.

We reasoned that examining complications, common knowledge details and self-handicapping strategies would be more effective for lie detection purposes than examining

total details. We found exactly that. However, we did not expect that the variable total details could not discriminate truth tellers from liars at all, which is somewhat unusual in deception research (Vrij, 2008). In addition, we found no difference in providing detail between the Interpreter-absent and Interpreter-present interviews. This is also an unusual finding as in all experimental studies examining the effect of interpreters so far, it has been found that Interpreter-present interviews result in less information than Interpreter-absent interviews. One obvious explanation for the lack of findings for the total details variable in the current experiment is that our coding system was not sensitive enough to pick up differences in detail between experimental cells. For example, we did not count action details when they could be deciphered from the context of the sentence. For example, in the sentence 'We went by car to Amsterdam' we would not count the word 'went'. We are aware of coding schemes were such action details are counted (e.g., Colwell, Hiscock, & Memon, 2002). We do not think that our coding scheme was not sensitive enough. We have used the same coding scheme in studies before and often found differences in detail between experimental cells (e.g., Ewens et al., 2017; Vrij, Leal et al., 2017; Vrij, Nahari, Isitt, & Leal, 2016). Also, in the drawings part of the current research project (Vrij et al., in press), the variable total details did distinguish truth tellers from liars, as truth tellers reported more new details (details not reported in an earlier part of the interview) than liars.

Although Interpreter-present interviews did not result in a loss of information (as was found in the previous interpreter based experiments), interpreter-present interviews resulted in fewer cues to deceit and made lie detection more difficult. These findings demonstrate the importance of examining the effect of having an interpreter present on eliciting information and cues to deceit, because findings obtained in interpreter-absent interviews cannot always be generalized to interpreter-present interviews.

Two methodological issue merit attention. First, in the third unexpected question in the interview, participants were asked to describe anything unexpected that happened or did not go according to plan when organising the trip. It could be argued that some complications reported in answer to this question were not reported spontaneously but were directly queried. In another deception study, unexpected events (which is similar to complications) that were mentioned spontaneously and that were mentioned after being prompted were examined separately (Jundi, Vrij, Hope, Mann, & Hillman, 2013). Truth tellers reported more spontaneous and prompted unexpected events than liars. This suggests that complications differentiate truth tellers from liars both when they are reported spontaneously or after being prompted.

Second, in the experiment truth tellers were given the opportunity to prepare themselves. We gave then this opportunity to avoid creating a confound between veracity and preparation, as we also gave liars the opportunity to prepare themselves. Although asking truth tellers to prepare themselves increased the robustness of the experimental design someone may argue that this is not realistic as in real life truth tellers are unlikely to prepare themselves. We believe that the advantage of the experimental robustness outweighs this disadvantage, also because good interviewing involves giving interviewees plenty of opportunity to think about the event they are interviewed about (Fisher, 2010).

References

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

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

Colwell, K., Hiscock-Anisman, C. K., & Fede, J. (2013). Assessment Criteria Indicative of Deception: An example of the new paradigm of differential recall enhancement. In B. S. Cooper, D. Griesel, & M. Ternes (Eds*.) Applied issues in investigative interviewing, eyewitness memory, and credibility assessment (pp. 259-292).* New York: Springer

Colwell, K., Hiscock, C. K., & Memon, A. (2002). Interview techniques and the assessment of statement credibility. *Applied Cognitive Psychology*, *16*, 287-300. DOI: 10.1002/acp.788

Department of the Army. (2006). *Human Intelligence Collector Operations (Field Manual 2-22.3).* Washington, DC: Headquarters, Department of the Army (United States of America).

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

DePaulo, B. M., & Morris, W. L. (2004). Discerning lies from truths: Behavioural cues to deception and the indirect pathway of intuition. In P. A. Granhag & L. A. Strömwall (Eds.), *Deception detection in forensic contexts* (pp. 15-40). Cambridge, England: Cambridge University Press.

Du Prel, J-B., Hommel, G., Röhrig, B., & Blettner, M. (2009). Confidence interval or p-value? *Deutzches Arzteblatt International*, *106*, 335-339. Doi:10.3238/arztebl.2009.0335.

Ewens, S., Vrij, A., Leal, S., Mann, S., Jo, E., & Fisher, R. P. (2016a). The effect of interpreters on eliciting information, cues to deceit and rapport. *Legal and Criminological Psychology, 21*, 286-304 DOI: 10.1111/lcrp.12067

Ewens, S., Vrij, A., Leal, S., Mann, S., Jo, E., & Houston, K. (2017). The effect of the presence and seating position of an interpreter on eliciting information and cues to deceit. *Psychology, Crime, & Law*, *23*, 180-200. doi:10.1080/1068316X.2016.1239100

Ewens, S., Vrij, A., Leal, S., Mann, S., Jo, E., Shaboltas, A., Ivanova, M., Granskaya, J., & Houston, K. (2016b). Using the model statement to elicit information and cues to deceit from native speakers, non-native speakers and those talking through an interpreter. *Applied Cognitive Psychology*, *30*, 854-862. DOI: 10.1002/acp.3270

Ewens, S., Vrij, A., Mann, S., & Leal, S. (2016c). Using the reverse order technique with non-native speakers or through an interpreter. *Applied Cognitive Psychology, 30*, 242-249. DOI: 10.1002/acp.3196

Field, A. (2009). *Discovering statistics using SPSS (third edition).* Los Angeles, CA: Sage.

Fisher, R. P. (2010). Interviewing cooperative witnesses. *Legal and Criminological Psychology, 15*, 25-38. Doi: 10.1348/135532509X441891

Fritz, C. O., Morris, P. E., & Richler, J. J. (2012). Effect size estimates: Current use, calculations and interpretation. *Journal of Experimental Psychology: General, 141*, 2-18, doi: 10.1037/a0024338

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

Houston, K. A., Russano, M. B., & Ricks, E. P. (2017). 'Any friend of yours is a friend of mine': Investigating the utilization of an interpreter in an investigative interview. Psychology, Crime, & Law, 23, 413-426. Doi: 10.1080/1068316X.2017.1290091

Jundi, S., Vrij, A., Hope, L., Mann, S., & Hillman, J. (2013). Establishing evidence through undercover and collective intelligence interviewing. *Psychology, Public Policy, & Law, 19*, 297-306. DOI: org/10.1037/a0033571

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

Nahari, G. (in press). Advances in lie detection: Limitations and potential for investigating allegations of abuse. In R. Burnett (Ed.), *Wrongful Allegations of Sexual and Child Abuse*. Oxford: Oxford University Press.

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.

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

Nelson, D. L., & Goodmon, L. B. (2003). Disrupting attention: The need for retrieval cues in working memory theories. *Memory and Cognition, 31*, 65-76. Doi: 10.3758/BF03196083

Roos af Hjelmsäter, E., Öhman, L., Granhag, P. A., & Vrij, A. (2014). Mapping deception in adolescents: Eliciting cues to deceit through an unanticipated spatial drawing task. *Legal and Criminological Psychology,* 19, 179-188. Doi: 10.1111/j.2044-8333.2012.02068.x

Ruby, C. L., & Brigham, J. C. (1998). Can Criteria-Based Content Analysis distinguish between true and false statements of African-American speakers? *Law and Human Behavior*, *22*, 369-388. Doi.org/10.1023/A:1025766825429

Sooniste, T., Granhag, P.A., Knieps, M., & Vrij, A. (2013). True and false intentions: Asking about the past to detect lies about the future. *Psychology, Crime, & Law*, *19*, 673-685. DOI 10.1080/1068316x.2013.793333

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

Steller, M., & Köhnken, G. (1989). Criteria-Based Content Analysis. In D. C. Raskin (Ed.), *Psychological methods in criminal investigation and evidence* (pp. 217-245). New York: Springer-Verlag.

Tedeschini, J. (2012). Overcoming roadblocks to reform. *Journal of Applied Research in Memory and Cognition*, *1*, 134–135. Doi: 10.1016/j. jarmac.2012.04.008

Viezzi, M. (2012). Simultaneous and consecutive interpreting (non-conference settings). In C. Millan & F. Bartrina, (Ed.) *The Routledge handbook of translation studies*, (pp. 377-388). Abingdon: Routledge.

Volbert, R., & Steller, M. (2014). Is this testimony truthful, fabricated, or based on false memory? Credibility assessment 25 years after Steller and Köhnken (1989). *European Psychologist,* 19, 207-220. doi:10.1027/1016-9040/a000200.

Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities, second edition*. Chichester: John Wiley and Sons.

Vrij, A. (2016). Baselining as a lie detection method. *Applied Cognitive Psychology, 30,* 1112-1119. Doi: 10.1002/acp.3288

Vrij, A., & Fisher, R. P. (2016). Which lie detection tools are ready for use in the criminal justice system? *Journal of Applied Research in Memory and Cognition, 5,* 302-307. Doi: 10.1016/j.jarmac.2016.06.014

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

Vrij, A., & Granhag, P. A. (2012). Eliciting cues to deception and truth: What matters are the questions asked. *Journal of Applied Research in Memory and Cognition, 1,* 110-117. doi.org/10.1016/j.jarmac.2012.02.

Vrij, A., & Granhag, P. A. (2014). Eliciting information and detecting lies in intelligence interviewing: An overview of recent research. *Applied Cognitive Psychology, 28,* 936-944. DOI: 10.1002/acp.3071.

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

Vrij, A., Leal, S., Fisher, R. P., Mann, S., Dalton, G. Jo, E., Shaboltas, A., Khaleeva, M., Granskaya, J., & Houston, K. (in press). Sketching as a technique to elicit information and cues to deceit in interpreter-based interviews. *Journal of Applied Research in Memory and Cognition*.

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

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

Vrij, A., Leal, S., Mann, S., 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., Nahari, G., Isitt, R., & Leal, S. (2016). Using the verifiability lie detection approach in an insurance claim setting. *Journal of Investigative Psychology and Offender Profiling*, *13*, 183-197. Doi: 10.1002/jip.1458.

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

Table 1.
Total detail, complications, common knowledge details, self-handicapping strategies and proportion of complications as a function of Veracity.

	Truth	**Lie**	***F***	***p* (two-tailed)**	**Cohen's *d***
	***M* (*SD*) 95% CI**	***M* (*SD*) 95% CI**			**[95% CI]**
Unexpected (Planning of the trip)					
Detail	31.68 (20.87) (28.00,35.76)	32.22 (16.72) (28.11,35.88)	0.001	.971	0.03 (-.30,.24)
Complications	3.48 (2.90) (2.90,3.98)	2.39 (2.29) (1.89,2.97)	6.101	.014	0.42 (.13,.69)
Common knowledge details	0.64 (0.72) (0.46,0.80)	1.30 (0.95) (1.16,1.49)	29.883	< .001	0.79 (.49,1.05)
Self-handicapping strategies	0.32 (0.55) (0.22,0.49)	0.63 (0.74) (0.46,0.73)	5.771	.017	0.48 (.19,.75)
Proportion of complications	0.72 (0.30) (0.65,0.76)	0.53 (0.27) (0.48,0.59)	15.025	<.001	0.67 (.37,.94)
Expected (Actual trip)					
Detail	76.03 (41.77) (67.23,82.24)	71.24 (33.15) (64.59,79.62)	0.216	.642	0.13 (-.14,.40)
Complications	6.51 (7.38) (5.23,7.51)	3.78 (2.99) (2.75,5.03)	8.329	.004	0.48 (.20,.76)
Common knowledge details	5.05 (2.80) (4.35,5.50)	6.26 (2.91) (5.83,6.99)	11.735	.001	0.42 (.14,.69)
Self-handicapping strategies	0.17 (0.40) (0.08,0.30)	0.43 (0.65) (0.30,0.53)	7.226	.008	0.48 (.20,.75)
Proportion of complications	0.46 (0.26) (0.41,0.51)	0.34 (0.20) (0.29,0.39)	10.793	.001	0.52 (.23,.79)

Table 2a
Hit Rates for total detail and proportions of complications as a function of Veracity. Interpreter-absent interviews.

	Hit Rate			χ^2	Wilks' λ	*p*-value	Canonical correlation
INTERPRETER-ABSENT	Truths (%)	Lies (%)	Total (%)				
Unexpected (Planning of the trip)							
Total detail	----	----	----	00.165	.998	.684	.041
Proportion of complications	75.0	67.3	71.3	26.472	.764	<.001	.485
Expected (Actual trip)							
Total detail	51.9	79.6	65.3	05.312	.948	.021	.229
Proportion of complications	61.5	67.3	64.4	14.980	.859	<.001	.376
Decision rules	(n=47)	(n=41)					
Proportion complications **unexpected** > .50 is truth teller, < .50 is liar and .50 is inconclusive	87.2	58.5	73.9	22.570	.768	<.001	.482
	(n=46)	(n=46)					
Proportion complications **expected** > .50 is truth teller, < .50 is liar and .50 is inconclusive	47.8	93.5	70.7	21.725	.784	<.001	.464
	(n=25)	(n=26)					
Proportion complications **unexpected** and **expected** combined (both process and outcome > .50 is truth teller, both process and outcome < .50 is liar; others are inconclusive)	84.0	92.3	88.2	42.973	.412	<.001	.767

Table 2b
Hit Rates for total detail and proportions of complications as a function of Veracity. Interpreter-present interviews.

	Hit Rate			χ^2	Wilks' λ	*p*-value	Canonical correlation
INTERPRETER-PRESENT	Truths (%)	Lies (%)	Total (%)				
Unexpected (Planning of the trip)							
Total detail	----	----	----	00.009	1.00	.925	.009
Proportion of complications	----	----	----	02.911	.971	.088	.169
Expected (Actual trip)							
Total detail	----	----	----	00.943	.991	.331	.097
Proportion of complications	----	----	----	02.817	.972	.093	.166
Decision rules	(n=47)	(n=40)					
Proportion complications **unexpected** > .50 is truth teller, < .50 is liar and .50 is inconclusive	78.7	45.0	63.2	05.593	.936	.018	.253
	(n=45)	(n=48)					
Proportion complications **expected** > .50 is truth teller, < .50 is liar and .50 is inconclusive	----	----	----	02.998	.967	.083	.181
	(n=27)	(n=20)					
Proportion complications **unexpected** and **expected** combined (both process and outcome > .50 is truth teller, both process and outcome < .50 is liar; others are inconclusive)	74.1	65.0	70.2	07.373	.847	.007	.391

[1] We prefer this analysis above an analysis that included Type of Question as a within-subjects factor. Such a design would include a Type of Question main effect in which we were not interested. In addition, the Veracity main effect would not refer to Veracity effects in the unexpected and expected questions separately (the interesting analyses), but in Veracity effects overall, which are less specific and less interesting analyses. Thus, we would then also have to report the Veracity effects we report now. Finally, we believe that comparing the *d*-values for Veracity in the expected and unexpected questions to be a better way to test Hypotheses 2 and 3 than interpreting a significant interaction effect. A *p*-value provides information about the statistical relevance but not about the practical importance of an effect (Fritz, Morris, & Richler, 2012; du Prel, Hommel, Röhrig, & Blettner, 2009). In this article, we are interested in the practical relevance of the effect and *d*-values are indicators of practical relevance (Fritz et al., 2012).

[2] The SHS results refer to the number of self-handicapping strategies made. When we analysed how many truth tellers and liars reported self-handicapping strategies we found that more liars (49%) than truth tellers (28.4%) reported self-handicapping strategies while answering the

process questions, $X^2(1, n = 204) = 9.195$, $p < .001$, *phi* = .211. Also when answering outcome questions, more liars (35.3%) than truth tellers (15.7%) reported self-handicapping strategies, $X^2(1, n = 204) = 10.324$, $p < .001$, *phi* = .225.