

The Effect of Interpreters on Eliciting Information, Cues to Deceit and Rapport

Sarah Ewens, Aldert Vrij, Sharon Leal and Samantha Mann

University of Portsmouth

Eunkyung Jo

Hallym University

Ronald P. Fisher

Florida International University

Author Note

Sarah Ewens, Aldert Vrij, Sharon Leal and Samantha Mann, Department of Psychology, University of Portsmouth; Eunkyung Jo, Department of Psychology, Hallym University; Ronald P. Fisher, Department of Psychology, Florida International University.

This work is funded by the High-Value Detainee Interrogation Group, J-FBI-12-194 awarded to the University of Portsmouth (UK). Any opinions, findings, and conclusions or recommendations expressed in this article are those of the authors and do not necessarily reflect the views of the U.S. Government.

Correspondence concerning this article should be addressed to Sarah Ewens, Department of Psychology, University of Portsmouth, King Henry Building, King Henry 1 Street, PO1 2DY, Hants, United Kingdom. Email sarah.ewens@port.ac.uk

Abstract

Background. The present experiment examined how the presence of an interpreter during investigative interviews affects eliciting information, cues to deceit and rapport.

Method. A total of 60 native English speakers were interviewed in English and 183 non-native English speakers were interviewed in English (a foreign language) or through an interpreter who interpreted their answers sentence by sentence (short consecutive interpretation) or summarised their answers (long consecutive interpretation). Interviewees discussed the job they had (truth tellers) or pretended to have (liars).

Results. Interviewees who spoke through an interpreter provided less detail than interviewees who spoke in their first language and a foreign language (English) without an interpreter. Additionally, cues to deceit occurred more frequently when interviewees spoke without an interpreter. The presence of an interpreter had no effect on rapport.

Conclusion. The findings suggest that at present there are no benefits to using an interpreter with regards to eliciting information. Future research should investigate how best to utilise an interpreter to gain maximum detail from an interview.

Keywords: interpreter, rapport, information gathering, deception

The Effect of Interpreters on Eliciting Information, Cues to Deceit and Rapport.

In today's society, with widespread travel, it is often the case that investigators and interviewees do not share the same first language. As a result, investigators may have little to no understanding of the interviewee's first language and vice versa. The investigator's inability to speak the interviewee's first language may result in one of two practices. Firstly, investigators may conduct the interview in their first language with interviewees responding in that language which is, for them, a foreign language. Alternatively, the interview may be conducted with the aid of an interpreter in the interviewee's first language.

Interpreting can be simultaneous, when the interpreter speaks at the same time as the individual they are interpreting, or consecutive/alternate, when the interpreter interprets what has been said after the individual has finished talking (Department of the Army, 2006; Viezzi, 2012). A further distinction can be made between two types of consecutive interpreting. Short consecutive interpretation, when the interpreter translates all turns of talk sentence by sentence, and long consecutive interpretation, when the interpreter translates segments of talk which may vary considerably in length (Viezzi, 2012). The US Navy field manual states that the interpreter listens to the entire phrase, sentence, or paragraph before translating (Department of the Army, 2006). In other words, it states that both types of consecutive interpretation are allowed but it does not give a preference for either method. To date, little is known about the effect of short and long consecutive interpretation on, eliciting the maximum amount of information, cues to deceit, and rapport with non-native speaking interviewees.

Information-gathering

It is reasonable to suggest that interviewees speaking in their first language are more talkative and provide more detail than interviewees who are less proficient in that language. Interviewees who speak in their first language have a larger vocabulary and can better

express themselves compared to those who speak in a foreign language (Ullman, 2001). In addition, those speaking in a foreign language may opt to leave out information simply because they do not know how to express these details in that language (Huang, 2010): a strategy known as ‘message reduction’ (Dornyei & Scott, 1995). Speaking in a foreign language is also cognitively demanding (Evans, Michael, Meissner, & Brandon, 2013). Hence, to lower this demand, interviewees may choose to provide a shorter statement which includes less detail. Finally, those speaking in a foreign language say less because they need to actively inhibit neural control mechanisms that would otherwise automatically make them respond in their first language (Wang, Xue, Chen, Xue, & Dong, 2007).

Having an interpreter present allows interviewees to speak in their first language. However, the introduction of an interpreter disrupts the flow of conversation and it is likely that those speaking through an interpreter will provide fewer details than interviewees speaking in their first language. Research has shown that interruptions lead to annoyance and anxiety (Bailey & Konstan, 2006), and interviewees who are annoyed may volunteer less information (Bull, 2010; Fisher, 2010). In addition, interruptions may make memory retrieval more difficult, which would result in less information being reported (Nelson & Goodmon, 2003). Finally, interviews with interpreters can take longer and the flow of information exchange is slow. This may make an interviewee decide to be as concise as possible and only discuss the core issues without elaboration. Indeed, the second author of this paper experiences such issues when giving presentations through interpreters. In fact, physicians who communicated with patients through an interpreter were less likely to engage in small talk and, in those conversations, the patients asked fewer questions (Aranguri, Davidson, & Ramirez, 2006).

How short and long consecutive interpreting relates to conveying detail is difficult to predict. Short consecutive interpretation will result in a more complete and accurate

translation of the interviewee's speech. However, short consecutive interpretation will also take longer and produces more disruptions to the flow of the conversation. Thus interviewees may become more reluctant to volunteer details in a short consecutive interpretation interview than in a long consecutive interpretation interview.

Verbal cues to deceit

When interviewees says more, the likelihood of verbal cues to deceit occurring will increase. We argued earlier that the interviews whereby interviewer and interviewee share the same first language are expected to elicit most detail. These interviews are, therefore, also most likely to elicit verbal cues to deceit.

Interviewees who are interviewed without an interpreter in a foreign language are likely to experience cognitive difficulty when communicating in that language (Evans et al., 2013). This additional mental load may further elicit cues to deceit. Lying is often more mentally taxing than truth telling, because lying involves more tasks, e.g., fabricating and maintaining a lie, creating a convincing impression, and scrutinizing the interviewer to check if they are believed (Vrij et al., 2008). Consequently, liars have fewer cognitive resources left over to cope when cognitive demand is further raised in an interview. Cognitive demand is further raised by requesting interviewees to communicate in a foreign language (Akca & Elkilic, 2011; Evans et al., 2013). Such a request should thus affect liars more than truth tellers, with verbal cues to deceit likely to occur.

With an interpreter present, the interview becomes considerably easier for interviewees. First, it allows them to speak in their first language, which is cognitively easier. Second, the presence of an interpreter gives interviewees plenty of opportunity to think during the interview. Each time the interpreter or interviewer speaks the interviewee has time to contemplate what to say next. The opportunity to think combined with the possibility that

limited detail will be conveyed in interviews with interpreters makes it less likely that cues to deceit will occur.

We examined two verbal cues to deceit, detail and plausibility. Deception research has demonstrated that truth tellers typically give more detail than liars (DePaulo et al., 2003; Masip, Sporer, Garrido, & Herrero, 2005; Vrij, 2008). Liars may lack imagination needed to convey the amount and type of detail that truth tellers convey. Liars may also be reluctant to provide much detail as they fear this detail may provide leads to investigators to check. Deception research further demonstrated that liars' statements sound less plausible than truth tellers' statements, suggesting that if liars manage to include fabricated detail in their statement they sometimes struggle to do so in a convincing way (DePaulo et al., 2003; Leal, Vrij, Warmelink, & Fisher, 2012; Vrij, Leal, Mann, & Fisher, 2012; Vrij, Mann, Leal, & Fisher, 2012). With regards to how an interpreter will affect plausibility, whilst saying more does not necessarily mean that what is said sounds more plausible, saying little or nothing when communicating through an interpreter would sound less plausible.

Rapport

Rapport is defined as a harmonious, positive and productive relationship between an interviewer and interviewee (Evans, Houston, & Meissner, 2012; Walsh & Bull, 2012). It is the most critical element of investigative interviewing, according to a US Intelligence Science Board report on gathering information (Fein, 2006). This conclusion is echoed by others. For example, the FBI argues that the most effective way to obtain accurate information from interviews is to use rapport-building techniques (Driskell, Blickensderfer, & Salas, 2013). Establishing rapport is important as it facilitates talking and cooperation (Bull & Soukara, 2010; Drolet & Morris, 2000; Macintosh, 2009; Valley, Thompson, Gibbons, & Bazerman, 2002), more accurate recall (Collins, Lincoln, & Frank, 2002; Vallano & Schreiber Compo, 2011), helps investigators gain interviewees' trust and, in turn,

facilitates relationship building between interviewers and interviewees resulting in a more productive interpersonal experience (Abbe & Brandon, 2012).

The question arises about how the presence of a third person (another interviewer or an interpreter) affects rapport. Dyad compared with triad interactions are seen as fundamentally different in terms of intimacy with closeness being more revealed within dyad interactions (Simmel, 1964). Indeed, intelligence investigators in the field have mentioned that interpreters have a negative effect on rapport (Soufan, 2011). Furthermore, the US Department of Defence field manual on intelligence collection cautions that a third person may negatively impact the establishment of rapport (Driskell et al., 2013). Driskell et al., (2013) examined how the introduction of a third party affected rapport in police interviews. In contrast, to what those in the field report, no difference was found in rapport when they compared interviews conducted by one or two interviewers.

The role of an interpreter is fundamentally different from that of an interviewer. The interpreter's role is not to question or interrogate interviewees but to aid communication by bridging the barrier between two people who do not share the same mother tongue. A study which focussed on the effect of interpreters on rapport building found that trust or rapport was not affected when physicians interacted with patients through an interpreter. However, physicians reported difficulty in eliciting symptoms and discussing treatment plans through an interpreter (Karlner, Perez-Stable, & Gildengorin, 2004). In sum, it has been argued that the presence of an interpreter will hamper rapport, however, further research is needed to investigate the effects of an interpreter.

Although it is difficult to predict how the presence of an interpreter will affect rapport, the effect that lying or truth telling will have seems more straightforward to predict. Liars can feel guilty about lying or can be afraid of having their lies exposed (Ekman, 1985), and subsequently liars can express more negative affect than truth tellers (DePaulo et al.,

2003). When someone experiences negative affect, s/he may perceive the environment (i.e., the interview or interviewer) in a negative frame of mind (Jundi, Vrij, Hope, Mann, & Hillman, 2013; Mann et al., 2012). Those in a negative mind set may become uncooperative and unresponsive to any attempt to build rapport. Currently, no research has investigated the effect that lying has on rapport. The current paper will examine this.

Hypotheses

We predicted that the greatest amount of detail would be provided in the interview whereby the interviewer and interviewee shared their first language (Hypothesis 1). We further explored how the short and long consecutive interpreter groups and the interviewees speaking in a foreign language compare to each other in terms of providing detail. Finally, we predicted that verbal cues to deceit (lack of detail and lack of plausibility) are more likely to occur in interviews where interviewer and interviewee shared the same first language and in interviews where the interviewee speaks in a foreign language compared to interviews where an interpreter is present (Hypothesis 2). We predicted that interviewees would experience less rapport with the interviewer when an interpreter was present than when an interpreter was absent (Hypothesis 3). Additionally, due to negative affect, we predicted that liars would experience less rapport with the interviewer than truth tellers (Hypothesis 4).

Method

Participants

A total of 243 participants (145 Females and 98 Males) took part in the study. They were of British ($n = 60$), Chinese ($n = 45$), Arabic ($n = 19$), Korean ($n = 68$), Hispanic ($n = 48$) and Urdu ($n = 3$) background. Ages ranged from 16-75 years with an average age of 26.44 years ($SD = 10.91$). Participation took place in three different universities located in the United Kingdom, USA and Republic of Korea (South Korea). Analyses revealed a similar gender distribution across all four conditions $\chi^2(3, 243) = 4.12, p = .25, phi = .13$. Age

differed between conditions $F(3, 238) = 12.43, p < .001, \eta^2 = .14$, with the participants in the two interpreter conditions being older ($M = 30.83, SD = 12.16; M = 30.52, SD = 15.71$) than the participants in the two non-interpreter conditions ($M = 22.02, SD = 5.30; M = 23.33, SD = 5.24$). However, when age was used as a covariate in all proceeding analyses it did not change the findings reported in the Results section regarding Subjective Detail and Plausibility (the effect of age was not significant for Subjective Detail, $F(1, 233) = .52, p = .473, ns, \eta^2 = .00$ and Plausibility, $F(1, 233) = .87, p = .353, ns, \eta^2 = .00$). The results for Rapport changed as the Veracity main effect was no longer significant, $F(1, 233) = 1.20, p = .27, \eta^2 = .01$ (the effect of age was significant for Rapport, $F(1, 233) = 9.96, p = .002, \eta^2 = .04$).

Procedure

An ‘occupation scenario’ similar to Mann et al. (2012) was used. Participants were recruited via advertising posters and internet announcements and asked to take part in a study about ‘Improving cross-cultural communications in interviews’. Participants were emailed a ‘selection briefing form’, which contained a list of 17 different jobs, and were asked to indicate how much they knew about each job (1 = very little to 7 = a lot). They were further asked which job, if any, they currently had.

Participants were allocated randomly to the truth telling ($n = 128$) or lying ($n = 115$) condition. Truth tellers were informed that they would be interviewed about their current job. Liars were told that their task would be to convince the interviewer they had a job that they did not have. The job chosen was one participant’s knew less about (i.e., an indicated score of 2 or 3 on the selection form). This score ensured the participants were in fact lying but it was not a completely impossible lie and thus reflects a real life situation.

Interviews were scheduled for a minimum of three days following the veracity allocation and participants were not restricted on how much they could prepare. In order to

motivate participants to be convincing, we informed them that they would receive a £5 (or equivalent) reward if the interviewer believed them to be telling the truth. Participants were further told that if they were not believed they would have to write a statement detailing why they thought this was the case. For ethical reasons all participants were told that the interviewer believed them and received their reward.

On arrival to the corresponding university the participants were greeted by members of the research team. All participants completed a pre-interview questionnaire in which they were asked to what extent they were motivated to perform well in the interview on a 5 point scale (1 = not at all motivated to 5 = very motivated). All forms were translated and completed in the first language of each participant; any answers were translated into English. Before being interviewed, truth tellers were reminded to answer the questions truthfully about their current job and liars were reminded which job they needed to convince the interviewer they had. All participants said that they had understood the instructions and all liars said they had remembered the jobs they had been allocated previously. A check of the transcripts revealed that all truth tellers discussed their current job and all liars discussed their allocated job. The roles of the interviewer and interpreter were explained to each participant, ensuring that they understood the interpreter was not a fellow interviewer but an impartial person bridging the communication gap.

Participants were then brought to the interview room and introduced to the interviewer and, if present, the interpreter. Both the interviewer and the interpreter were blind to the veracity of the participant. Two female interviewers were used for all interviews. Both interviewers are British and spoke English during the interviews. The interviewers were instructed to keep an open posture but to avoid displaying any expressiveness, as being supportive or sceptical influences participant's responses during an interview (Mann et al., 2012). In total, twelve interpreters were used in the study; Chinese ($n = 1$), Arabic ($n = 2$),

Urdu ($n = 1$), Korean ($n = 2$) and Hispanic ($n = 6$). Of these twelve, five had previous interpreting experience. The interpreters were requested to speak in the first person. They were situated next to the interviewer both of whom faced the interviewee, thus forming a triangle. The interpreters either interpreted the interviewee's answer short consecutive (sentence by sentence, $n = 64$) or long consecutive (gave a complete rendition of the interviewee's response [to the best of their ability] after the interviewee had finished answering each question, $n = 50$). Participants were randomly assigned to one of the two interpreter conditions. In both interpreter conditions the interpreters were instructed to give a complete rendition of the interviewee's response (rather than a summary).

The study included two further conditions which did not have an interpreter present. One condition (native-English) consisted of native English speaking participants ($n = 60$) who were interviewed in English. In the other condition (non-native English), Chinese, Arabic, Hispanic and Korean native speakers ($n = 69$) were interviewed in English (and answered in English). The non-native languages were equally distributed across the non-native English and interpreter conditions, ensuring that language did not affect the non-native conditions. Inclusion criteria were used for the two conditions which did not involve an interpreter. The native English condition consisted of participants whose first language was English. These participants were recruited at the university in the UK. In the non-native English condition, all participants had an intermediate level of English, ensuring that they would be able to get by in the interview. These participants were recruited at all three universities. All participants, regardless of condition, contained a sample of university administrative/maintenance staff, students, and the general public.

The interview commenced with three questions: 'What is your job and how many hours a week do you work?', 'How long have you been in your job?' and 'Where do you work?' To make the interviewee feel comfortable and to avoid floor effects in establishing

rapport (i.e., no rapport in any of the experimental conditions), these questions were followed by self-disclosing information from the interviewer. Following this, the remaining five detail-eliciting questions were asked. They were open rather than closed questions and required long answers: (1) 'Please describe your place of work in as much detail as you can.' (2) 'There must be one single experience in your job that must stand out – what is that? What happened?' (3) 'Can you describe a typical day/shift at work, hour by hour?' (4) 'Can you tell me about a recent interaction or event that you were involved in within the last week that occurred in your workplace? Just something that springs to mind, but doesn't have to be out of the ordinary, but please do describe it in detail' (5) 'If you were training me to do your job for a day, what things would I need to know about it?' The questions were derived from Mann et al. (2012) and Vrij et al. (2012).

After the interview, participants completed a post-interview questionnaire which measured motivation, likelihood of receiving the £5, likelihood of writing a statement and rapport with the interviewer. To measure motivation participants were asked to what extent they were motivated to perform well on a 5 point scale (1 = not at all motivated to 5 = very motivated). Likelihood of receiving the £5 or writing a statement was measured on 7 point scales (1 = not at all to 7 = totally). 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'. The post-interview questionnaire further asked the non-native English participants whether they would request an interpreter if they were arrested in an English speaking country via a yes/no response.

The interviews were video and audio recorded and the English speech in the audiotapes was subsequently transcribed.

Coding

All coders read the transcripts and were blind to the hypotheses and experimental conditions of the study.

Subjective detail. A coder rated each of the five detail-eliciting questions on a 7-point scale (1 = not detailed to 7 = very detailed) and the average of their scores formed the subjective coding score. Their coding showed good inter-rater reliability with a second coder's ratings of subjective detail in a sample of 60 transcripts (i.e., 25%) (Intra-class Correlation Coefficient [ICC] = .91)¹.

Plausibility. Five coders rated the plausibility of the responses to each of the five detail-eliciting questions on a 7-point scale (1= not plausible to 7= very plausible). Plausibility is defined as a seemingly or apparently valid, likely, or acceptable response to the questions asked. A total plausibility score (used in this article) was calculated by averaging each coder's five plausibility scores. Agreement between the five coders was satisfactory (Cronbach's alpha = .66).

Grasp of English. Three coders rated English proficiency by listening to the interviews and using a scale from Embassy English, an English language training scheme. The scale consists of five categories: [1] Beginner (those who know a few English words), [2] Elementary (those who can communicate in a basic way/can make simple sentences), [3] Pre-Intermediate (those with a good basic ability to communicate and understand), [4] Intermediate (those who have the grammar to talk about a wide number of subjects), and [5] Upper-Intermediate (those who can talk fluently and almost completely accurately). The scale is available from: http://www.embassyces.com/about/should_know.aspx. A reliability analysis revealed that the agreement between coders was excellent (Cronbach's $\alpha = .93$). When there was a disagreement between the three coders, two coders gave the same ratings and that a third coder was an outlier. In such situations, the classification made by the two coders who agreed was used. The interviewees were classified as Beginner 6%, Elementary

34%, Pre-Intermediate 36%, Intermediate 20% and Upper-Intermediate 4%. Additionally the non-native English participants were asked if they would request an interpreter in a police interview in an English-speaking country. Of these, 77% would have requested an interpreter in an interview situation.

Interpretation Checks. A MANOVA with Veracity (truth versus lie) X Interpreter (short versus long consecutive) X Interpreter Experience (no versus yes) as factors and ‘rapport’, ‘subjective detail’, and ‘plausibility’ as dependent variables, revealed a significant Interpreter Experience main effect, $F(3, 104) = 2.78, p = .045, \eta^2 = .07$. The interaction effects that involved Interpreter Experience were not significant, all F 's < 2.36 , all p 's $> .075$. Regarding the Interpreter Experience main effect, at a univariate level one significant effect emerged. Rapport with the interviewer was better when the interpreter was inexperienced ($M = 5.78, SD = .92$) than when the interpreter was experienced ($M = 5.32, SD = .89$), $F(1, 106) = 6.43, p = .013, \eta^2 = .057$. The univariate effects regarding subjective detail and plausibility were not significant, both F 's < 2.03 , both p 's $> .15$. We, therefore, did not take Interpreter Experience into account in the subsequent analyses regarding subjective detail and plausibility.

To check that the interpreters interpreted correctly, the number of segments of talk they gave was recorded for the five detail-eliciting questions. Segments refer to the interpreters' renditions of the interviewee's answers. Both interpreter conditions were adhered to and, in the short consecutive condition ($M = 39.08, SD = 22.30, 95\% \text{ CI } [34.92, 43.24]$), more segments of talk were given than in the long consecutive condition ($M = 7.38, SD = 2.39, 95\% \text{ CI } [2.67, 12.09]$), $F(1, 112) = 99.90, p < .001, d = 2.00$.² The interviewer said more in the long consecutive condition ($M = 547.48, SD = 226.86$, ranging from 232 to 1316 words) than in the short consecutive condition ($M = 463.76, SD = 205.39$, ranging from 193 to 1037 words), $F(1, 112) = 4.15, p = .044, d = 0.51$. In addition, we transcribed the

speech of 41 interviewees who spoke through an interpreter, translated these statements into English and coded these statements for detail. We then compared the number of details originally recalled by the interviewee with the number of details translated by the interpreter. In the short consecutive condition the interpreters translated more detail ($M = 77.81$, $SD = 23.23$, 95% CI [67.24, 88.38]) than given by the interviewee ($M = 72.90$, $SD = 26.28$, 95% CI [60.94, 84.87]), $F(1, 20) = 3.98$, $p = .060$, $d = 0.20$. Similarly, also in the long consecutive condition the interpreters translated more detail ($M = 75.95$, $SD = 25.87$, 95% CI [63.84, 88.06]) than given by the interviewee ($M = 67.75$, $SD = 23.89$, 95% CI [56.57, 78.93]), $F(1, 19) = 7.62$, $p = .012$, $d = 0.33$. A comparison of the mean scores between the originally given and translated detail indicate that the interpreters in both conditions translated about 10% more detail than the interviewees gave. Separate analyses for different type of detail revealed that interpreters ($M = 45.20$, $SD = 13.70$, 95% CI [40.87, 49.52]) appeared to give more visual details than interviewees ($M = 40.05$, $SD = 14.60$, 95% CI [35.44, 44.66]), $F(1, 40) = 20.65$, $p < .001$, $d = 0.36$. One participant used a clarification to explain their situation. The interpreter also interpreted this same clarification. However, the interviewee then says ‘I took them there’ whilst the interpreter says ‘I took them to the hospital’ adding a further visual, ‘hospital’. It would appear that interpreters were not adding detail as such but using what the interviewee has previously said. Another example is that gestures made by the interviewee were replaced with words by the interpreter. A look at the video revealed that when describing a hall the interviewee indicated with their hands that it was big. Although they did not say it was big the interpreter interpreted this action and indicated ‘there is a big hall’.

Analyses revealed that, in the short consecutive condition, the five experienced interpreters ($M = 32.51$, $SD = 15.41$, 95% CI [28.53, 36.50]) made fewer segments of talk than the seven inexperienced interpreters ($M = 48.07$, $SD = 27.04$, 95% CI [40.59, 55.56]), $F(1, 62) = 8.50$, $p < .001$, $d = 0.71$. In the long consecutive condition experienced and

inexperienced interpreters did not differ in segments of talk, $F(1, 48) = 3.86, p = .055, d = 0.56$.

Perhaps more important is the total number of details that the interpreters conveyed. In the short consecutive condition the experienced interpreters conveyed a similar number of details as the inexperienced interpreters, $F(1, 62) = 1.40, p = .241, d = 0.30$. In the long consecutive condition the difference between experienced interpreters and inexperienced interpreters was not significant either, $F(1, 48) = 3.74, p = .059, d = 0.55$.

Interview length. A 2 (Veracity) X 4 (Interpreter Condition; Native-English, Short Consecutive, Long Consecutive, Non-Native English) analysis was carried out with interview length as dependent variable. The truthful and deceptive interviews were of a similar length, $F(1, 235) = .32, p = .575, ns, d = 0.11$. Interpreter Condition had an effect on the length of interview, $F(3, 235) = 33.12, p < .001, d = 1.87$. The native-English interviews ($M = 584.92$ seconds, $SD = 183.52$, 95% CI [509.14, 660.69]) were significantly shorter than the short consecutive ($M = 1027.22, SD = 278.89$, 95% CI [954.46, 1102.36]), long consecutive ($M = 1058.98, SD = 372.12$, 95% CI [977.76, 1144.98]) and non-native English ($M = 991.93, SD = 329.60$, 95% CI [921.85, 1063.18]) interviews. These latter three conditions did not differ in duration. There was no significant Veracity X Interpreter Condition effect, $F(3, 235) = .60, p = .614, ns, \eta^2 = .00$.

Results

Motivation, Likelihood of Receiving Incentive and Receiving a Penalty

Four 2 (Veracity) X 4 (Interpreter Condition) ANOVAs were conducted on the four manipulation checks. Motivation before being interviewed (measured in the pre-interview questionnaire) revealed a significant Interpreter Condition effect, $F(3, 235) = 6.84, p < .001, \eta^2 = .08$. Tukey post hoc tests revealed that native-English participants ($M = 4.13, SD = .75$, 95% CI [3.95, 4.32]) were significantly more motivated than the short consecutive (M

= 3.73, $SD = .74$, 95% CI [3.54, 3.91]), long consecutive ($M = 3.62$, $SD = .75$, 95% CI [3.41, 3.83]) and non-native English participants ($M = 3.59$, $SD = .71$, 95% CI [3.42, 3.77]). The latter three groups did not differ from each other. The Veracity main effect, $F(1, 235) = 1.25$, $p = .265$, *ns*, $d = 0.14$, and Veracity X Interpreter Condition interaction effect were not significant, $F(3, 235) = .20$, $p = .893$, *ns*, $\eta^2 = .00$. The grand mean score ($M = 3.77$, $SD = .76$ on a 5-point scale) revealed that the participants were motivated to perform well during the interview. When motivation was used as a covariate in all proceeding analyses it did not change the findings reported in the Results section regarding Subjective Detail, Plausibility and Rapport. The effect of motivation was not significant for Subjective Detail, $F(1, 234) = 1.49$, $p = .223$, *ns*, $\eta^2 = .01$ and Plausibility, $F(1, 234) = 2.45$, $p = .119$, *ns*, $\eta^2 = .01$, but was significant for Rapport, $F(1, 234) = 4.68$, $p = .031$, $\eta^2 = .02$. Motivation during the interview (measured in the post-interview questionnaire) showed a significant Veracity effect, $F(1, 235) = 4.68$, $p = .032$, $d = 0.29$, with truth tellers ($M = 3.95$, $SD = .75$, 95% CI [3.82, 4.09]) being more motivated than liars ($M = 3.73$, $SD = .79$, 95% CI [3.60, 3.88]). The Interpreter Condition main effect, $F(3, 235) = 1.96$, $p = .121$, *ns*, $\eta^2 = .02$ and Veracity X Interpreter Condition interaction effect, $F(3, 235) = 1.02$, $p = .385$, *ns*, $\eta^2 = .01$, were not significant. The grand mean score ($M = 3.84$, $SD = .80$ on a 5-point scale) revealed that the participants were motivated to perform well during the interview. When motivation was used as a covariate in all proceeding analyses it did not change the findings reported in the Results section regarding Subjective Detail and Plausibility, but did change the findings regarding Rapport as the Veracity main effect was no longer significant, $F(1, 234) = 3.527$, $p = .062$, $\eta^2 = .02$. The effect of motivation was not significant for Subjective Detail, $F(1, 234) = 1.26$, $p = .262$, *ns*, $\eta^2 = .01$ and Plausibility, $F(1, 234) = .176$, $p = .676$, *ns*, $\eta^2 = .00$, but was significant for Rapport, $F(1, 234) = 12.61$, $p < .001$, $\eta^2 = .051$.

The likelihood of receiving an incentive of £5 (or equivalent) resulted in a main effects for Veracity, $F(1, 235) = 9.32, p = .003, d = 0.37$, with truth tellers ($M = 5.33, SD = .1.55, 95\% \text{ CI } [5.08, 5.61]$) being more convinced that they would receive the incentive than liars ($M = 4.76, SD = 1.51, 95\% \text{ CI } [4.45, 5.02]$). The Interpreter Condition main effect, $F(3, 235) = 1.98, p = .118, ns, \eta^2 = .03$ and Veracity X Interpreter Condition interaction effect were not significant, $F(3, 235) = .23, p = .875, ns, \eta^2 = .00$.

The ANOVA regarding receiving a penalty of writing a statement revealed no effect (all F 's < 2.29 , and all p 's $> .08$). These results suggest that all participants, regardless of Veracity and Interpreter Condition, thought they were equally likely to receive a penalty ($M = 3.89, SD = 1.48$ on a 7-point Likert scale).

Subjective Detail (Hypotheses 1 and 2)

All analyses reported in this article refer to the answers given to these five questions only. A 2 (Veracity) X 4 (Interpreter Condition) ANOVA, with subjective detail as the dependent variable, revealed a significant Veracity main effect, $F(1, 235) = 56.31, p < .001, d = 0.75$, a significant Interpreter Condition main effect, $F(3, 235) = 32.17, p < .001, \eta^2 = .29$, and a significant Veracity X Interpreter Condition interaction effect, $F(3, 235) = 16.18, p < .001, \eta^2 = .17$. Regarding the Veracity effect, truth tellers ($M = 2.63, SD = 1.05, 95\% \text{ CI } [2.53, 2.77]$) gave significantly more subjective detail than liars ($M = 1.99, SD = .60, 95\% \text{ CI } [1.86, 2.11]$). Regarding the Interpreter Condition effect, Tukey post-hoc tests revealed that participants in the native English condition gave significantly more subjective detail ($M = 3.06, SD = 1.11, 95\% \text{ CI } [2.89, 3.23]$) than participants in the short consecutive condition ($M = 2.21, SD = .73, 95\% \text{ CI } [2.02, 2.36]$), participants in the long consecutive condition ($M = 1.93, SD = .58, 95\% \text{ CI } [1.73, 2.11]$), and participants in the non-native English condition ($M = 2.10, SD = .74, 95\% \text{ CI } [1.94, 2.26]$). The three latter groups did not differ significantly from each other.

Regarding the significant interaction effect, post hoc analyses were carried out in which truth tellers and liars were compared in each of the four interpreter conditions. Truth tellers ($M = 3.92$, $SD = .83$, 95% CI [3.67, 4.18]) gave more detail than liars ($M = 2.19$, $SD = 0.52$, 95% CI [1.94, 2.45]) in the native English condition, $F(1, 58) = 93.81$, $p < .001$, $d = 2.50$. The same pattern of results emerged in the non-native English condition, with truth tellers ($M = 2.31$, $SD = .83$, 95% CI [2.07, 2.56]) reporting more detail than liars ($M = 1.89$, $SD = .57$, 95% CI [1.65, 2.13]), $F(1, 67) = 6.06$, $p = .016$, $d = 0.59$. In contrast, in the short consecutive condition, truth tellers ($M = 2.34$, $SD = .72$, 95% CI [2.10, 2.58]) and liars ($M = 2.04$, $SD = .72$, 95% CI [1.77, 2.32]) reported a similar amount of detail, $F(1, 62) = 2.60$, $p = .112$, $d = 0.42$. The same pattern of results occurred in the long consecutive condition with truth tellers ($M = 2.03$, $SD = .63$, 95% CI [1.81, 2.24]) and liars ($M = 1.81$, $SD = .50$, 95% CI [1.56, 2.06]) reporting a similar amount of detail, $F(1, 48) = 1.75$, $p = .192$, $d = 0.39$.

Plausibility (Hypothesis 2)

A 2 (Veracity) X 4 (Interpreter Condition) ANOVA with plausibility as the dependent variable revealed a significant Veracity main effect, $F(1, 235) = 38.86$, $p < .001$, $d = 0.73$, a significant Interpreter Condition main effect, $F(3, 235) = 17.46$, $p < .001$, $\eta^2 = .18$, and a significant Veracity X Interpreter Condition interaction effect, $F(3, 235) = 3.44$, $p = .018$, $\eta^2 = .04$. The interaction effect is the most informative of these three effects, and the only effect we discuss.

Post hoc analyses were carried out in which truth tellers and liars were compared in each of the four interpreter conditions. The same pattern of results emerged in three out of four conditions. Truth tellers ($M = 5.52$, $SD = .36$, 95% CI [5.37, 5.67]) gave significantly more plausible answers than liars ($M = 4.74$, $SD = .47$, 95% CI [4.59, 4.89]) in the native English condition, $F(1, 58) = 52.38$, $p < .001$, $d = 1.86$. Truth tellers ($M = 4.82$, $SD = .51$, 95% CI [4.66, 4.98]) gave also significantly more plausible answers than liars ($M = 4.58$, SD

= .41, 95% CI [4.40, 4.75]) in the short consecutive condition, $F(1, 62) = 4.19, p = .045, d = 0.52$, and truth tellers ($M = 4.70, SD = .70, 95\% \text{ CI } [4.46, 4.93]$) gave significantly more plausible answers than liars ($M = 4.22, SD = .69, 95\% \text{ CI } [3.98, 4.45]$) in the non-native English condition, $F(1, 67) = 8.09, p < .001, d = 0.69$. In the long consecutive condition, the difference between truth tellers ($M = 4.79, SD = .51, 95\% \text{ CI } [4.61, 4.97]$) and liars ($M = 4.56, SD = .42, 95\% \text{ CI } [4.36, 4.76]$) was not significant, $F(1, 48) = 2.86, p = .097, ns, d = 0.49$.

Rapport with the Interviewer (Hypotheses 3 and 4)

A 2 (Veracity) X 4 (Interpreter Condition) ANOVA, with rapport with the interviewer as the dependent variable, revealed a significant Veracity effect, $F(1, 235) = 5.48, p = .020, d = 0.31$. Truth tellers reported significantly higher levels of rapport with the interviewer ($M = 5.65, SD = .88, 95\% \text{ CI } [5.49, 5.80]$) than liars ($M = 5.38, SD = .87, 95\% \text{ CI } [5.22, 5.54]$), although this effect was no longer significant when age or motivation during the interview were introduced as a covariate (see above). The Interpreter Condition main effect, $F(3, 235) = .84, p = .476, ns, \eta^2 = .01$, and the Veracity X Interpreter Condition interaction effect, $F(3, 235) = 1.91, p = .129, ns, \eta^2 = .02$, were not significant.³

Discussion

Eliciting Information

The English participants who were interviewed in English (their first language) provided more detail than the non-native English participants who spoke in English (for them a foreign language) and who were interviewed through an interpreter, supporting Hypothesis 1. In all likelihood the non-native participants who spoke in English lacked the vocabulary to be as detailed as their English counterparts or experienced considerable cognitive load during the interview. Lack of vocabulary became evident in the length of the interviews. Although those who spoke in a foreign language provided less detail ($M = 81.41$ details) than the native

English speakers ($M = 136.93$ details), their interviews lasted considerably longer ($M = 991.93$ versus 584.92 seconds).

Furthermore, it could be that the participants who were interviewed through an interpreter became annoyed because of the interpreter interrupting them and, therefore, said less. It could also be that the interpreter's disruptions hampered memory retrieval. Finally, perhaps the presence of an interpreter made the interviewee decide to be as concise as possible. Of these three explanations, the first explanation is the least likely. If participants became annoyed by the presence of the interpreter, this would have had an effect on rapport, which was not the case (see below). Also participants in both interpreter conditions provided a similar amount of detail as those who spoke in a foreign language. With an interpreter present the interviewees could speak in their first language which gives them the possibility to provide much detail, yet they provided the same amount of detail as those who spoke in a foreign language and less detail than the native English speakers. Thus, it could be that interpreters are not being used effectively to gain the maximum amount of detail possible.

Interestingly, short and long consecutive interpretations resulted in the same amount of information being conveyed. One could argue that compared to a long consecutive interpretation, short consecutive interpretation will result in a more complete and accurate translation of the interviewee's speech and, therefore, in more detail. However, this was not found. For us, the most likely (albeit speculative) explanation is that the many disruptions in the short consecutive interpretation made interviewees more reluctant to volunteer information.

An alternative explanation is that the lack of difference in detail between the two interpreter conditions is due to an inflated account from the interpreters. That is, perhaps interviewees in the long consecutive condition provided fewer detail than interviewees in the short consecutive condition, but that the interpreter in the long consecutive condition

‘corrected’ this by adding more detail. We believe that this is an unlikely explanation. Indeed, when we made a comparison between what the interpreter reported and what the interviewee actually said, we found that interpreters did in fact interpret more information than was reported by the interviewee. However, this information was not additional information. Rather the interpreters sometimes repeated an aspect that had been previously mentioned by the interviewee. Importantly, this occurred in both the short and long consecutive interpretation styles and shows that even when an interpreter is interpreting sentence by sentence they have the ability to make ‘errors’. This is an important aspect which requires further investigation and is something that also occurs in real life, as seen in the trial of Oscar Pistorius who is accused of murdering his then girlfriend Reeva Steenkamp. The first witnesses, Michelle Burger, gave evidence through an interpreter, as her native language was Afrikaans and the trial was being conducted in English. Although the witness could speak and understand English very well (as became clear in the trial), she chose to give her evidence in the language she felt most comfortable with. Surprisingly, throughout her evidence she was correcting the interpreter and she told the judge that the interpreter was not interpreting correctly what she had said. These errors had a big impact on her evidence as the defence picked up on any inconsistency in details, trying to discredit her as a witness. A clip of this is available from: <https://www.youtube.com/watch?v=CcASMGKhkAU>

Cues to Deceit

Detail emerged as a cue to deceit. Liars were less detailed in the native English speaking condition but not in the conditions where an interpreter was present, supporting Hypothesis 2. Moreover, in the non-native English speaking condition liars provided less detail than truth tellers (also supporting Hypothesis 2). We believe that this is the result of the difficulty the interviewees experienced while speaking English. Because the act of lying is more difficult than the act of truth telling the additional request to speak in a foreign language

affects liars more than truth tellers, resulting in cues to deceit. Apart from being less detailed, liars were also less plausible than truth tellers when they spoke in their first-language or in a foreign language. In addition, liars were less plausible than truth tellers in the short consecutive interpretation condition, despite not being less detailed than truth tellers in this condition. Apparently, it was the quality rather than the quantity of detail that gave liars away in that condition. One could argue that during long consecutive interpretation an interpreter has more opportunity to express his or her own 'voice' than during short consecutive interpretation. Thus, the measure of plausibility might be created by the interpreter rather than the truth tellers and liars, which could explain the lack of difference in plausibility between truth tellers and liars in the long consecutive interpreter condition.

Plausibility was measured by reading the transcripts of the interviews and the coders were unaware that an interpreter was present or absent. Plausibility has been reliably coded in this way in deception research before (DePaulo et al., 2003; Leal et al., 2012; Vrij et al., 2012; Vrij et al., 2012). However, in theory plausibility may change if someone was to view video footage where an interpreter is clearly present or not, which would match real life more closely. For example, the interpreter may think that an answer sounds plausible or implausible and may reveal this through his/her demeanour. This demeanour, in turn, may affect the veracity decision made by the observer. Future research should investigate this.

Rapport

The presence of an interpreter had no effect on the interviewee's judgement of rapport with the interviewer, leading us to reject Hypothesis 3 that interviewees would experience less rapport with the interviewer when an interpreter was present. This supports the scarce research in this area (Karliner et al., 2004). However, it goes against the views of criminal and intelligence investigators in the field, who believe that the presence of an interpreter has a negative effect on rapport (Soufan, 2011; Driskell et al., 2013).

There are some noticeable differences between our experiment and real life criminal and intelligence interviews which may explain this discrepancy. Firstly, real-life interviews are considerably longer than the interviews in the present experiment. We cannot rule out that some consequences of having an interpreter present in an interview, for example the interruptions they cause in the flow of conversation will have a negative effect in the longer term. Perhaps people tolerate disrupting factors initially but tolerate them less as time progresses. Secondly, the interviewees in the present experiment experienced high rapport with the interviewers. This rapport could have been higher than typically obtained in the field due to the context of the experiment and reduced stakes involved. In real-life situations we would expect suspects to feel more uncomfortable and potentially reluctant during interviews and this may have a negative effect on establishing rapport. Alternatively, it may be that interpreters have no effect on rapport. Research has shown that when interruptions are relevant they do not lead to emotional arousal or disruption of performance (Morris & Perez, 1972). Thus, as long as the interpreter's disruptions aid the interviewee, rapport may not be affected from their perspective at least. Alternatively, it could be that being interviewed in English without an interpreter could be frustrating for a non-native English speaker due to not being able to express him/herself in the way s/he wants to. This frustration could hamper rapport with the interviewer. Although officers might perceive interpreters as interfering, this may not be an accurate measure of the actual effect. Thus the relationship between the interviewer and interpreter is also an important one, which would benefit from future investigation.

Methodological reasons may also be responsible for the null finding regarding the presence of an interpreter and rapport. A null finding could occur due to lack of sensitivity in the measurement of rapport. However, our measurement was not insensitive as it did reveal that rapport was correlated with age and motivation. We further found that rapport with the

interpreter was better when inexperienced interpreters were present than when experienced interpreters were present. We have no plausible explanation for this finding, but it shows that the effect of interpreter experience on rapport with the interviewee is worth to examine in future research.

Rapport was not influenced by the act of lying, although we hypothesised such an effect in Hypothesis 4. The findings therefore reveal a complex picture for rapport. It was not influenced by 'obvious' factors such as the use of an interpreter and deception, but was correlated with factors such as experience of the interpreter, age of the interviewee and interviewee's motivation to perform well.

Short – Long Consecutive Interpretation Comparison

One of the aims of the experiment was to compare short and long consecutive interpretation. We found little difference between them, including no difference in rapport and the same amount of detail elicited. This detail was significantly less than detail elicited in the first-language English speaking condition and similar to the amount of detail elicited in the foreign language speaking condition. Truth tellers and liars provided similar amount of detail in both interpreter conditions, unlike in the two non-interpreter conditions, where truth tellers reported more detail than liars. Only plausibility revealed a difference between the two interpreter conditions. Truth tellers were more plausible than liars in the short consecutive interpreting condition (as well as in the two non-interpreter conditions) but no difference in plausibility emerged in the long consecutive interpreting condition. However, the plausibility results for the short and long consecutive interpreting condition were almost identical and the results in the short consecutive interpreting condition only just reached significance. With barely a difference emerging between the two interpreter conditions, someone could imagine that interviewers and interviewees would prefer long consecutive interpreting as it leads to

fewer disruptions. However, further research is required to examine whether indeed no difference occurs between short and long consecutive interpreting.

Most of the questions that were asked in our experiment were short (one sentence questions). The result of this was that not much difference existed between the two interpreter conditions in how the questions were presented to the interviewee. In real life longer questions could be asked which may result in larger differences between short and long consecutive interpretation in presenting the questions. However, whether a difference in short or long consecutive interpretation of the questions does have an effect on the response given by the interviewee remains to be seen, and is a question for future research.

Experience of the Interpreter

We used a mixture of experienced and inexperienced interpreters. Although in real life it is more likely that experienced interpreters are used, it is not uncommon for police officers, relatives (including children) or even crime scene witnesses to carry out the interpretation (Berk-Seligson, 2000). The comparison between the inexperienced and experienced interpreters in the present experiment can be summarised as follows. The experienced interpreters made more renditions than the inexperienced interpreters but in terms of detail translated no difference between experienced and inexperienced interpreters emerged. Since detail is the crucial variable it means that the experience of the interpreter had no discernible effect in the present experiment. However, we do not suggest that interpreting is a job that does not require experience. Our findings showed that the difference in the amount of detail between experienced and non-experienced interpreters just failed to reach significance. Thus, if more participants had been recruited a significant medium effect size could have been obtained, suggesting that experience does matter in terms of the amount of detail gained.

Methodological Issues

In a study like this, it is difficult to decide what to use as a control group. We decided upon using native English speakers, speaking in English as a control group. We did so because this is the interesting comparison from an applied perspective. That is, English speaking interviewers are interested in how the English speech delivered by non-native speakers, who either speaks in English or in their native language through an English speaking interpreter, compares to when they interview native English speakers. This, of course, meant that the allocations of participants to the experimental conditions were not entirely random as native English speakers were not allocated to the interpreter conditions and vice versa. However, both the native and non-native English speakers were recruited from similar populations (university students and people working at the university) so we do think that the native and non-native English speakers were comparable on characteristics other than having English as their first language. In cases where we found differences in characteristics between experimental groups, in age and motivation, analyses of covariance revealed that such differences had no effect on eliciting detail and cues to deceit.

A limitation of the study is the lack of ground truth in that all the information from the truth tellers about their current jobs could not be verified. This lack of ground truth is not uncommon in deception research (Vrij, 2008). We asked participants what their current job was before being allocated to a veracity condition and have no reason to believe anyone would be lying at this stage. In fact, participants were asked to rate, on separate scales of 0-100%, the extent to which they had been truthful or lied during the interview. From the truth tellers, 83% reported to have been 100% truthful in the interview, whereas 13% reported to have been 90% truthful. From the liars, 34% reported to have been 100% deceitful in the interview.

This study only measured English proficiency in participants in the non-native English condition and not those in the interpreter conditions. It is possible that the language

proficiency in the interpreter conditions affected the engagement and behaviour of the participants. However, since it is unlikely that the English proficiency of participants differed between the short and long consecutive conditions (the participants were randomly allocated to these two conditions) it is unlikely that the level of English proficiency has affected the findings in these conditions. However, to measure the effect of English proficiency on engagement and behaviour, future research should examine the effect of English proficiency on interviews involving interpreters.

We measured English proficiency with a scale from Embassy English. Of course, alternatives such as IELTS exist. We chose the Embassy English scale because it is easy to apply and resulted in high inter-rater agreement between the different coders.

The decision to call an interpreter for non-English speaking suspects in police interviews in the United Kingdom usually lies jointly with the interviewing and/or arresting officer and the Custody Sergeant (Russell, 2002). Perhaps our findings could advise those who have to make decisions about the use of interpreters in interviews. Our results suggest that it is preferable to interview interviewees in their own language. Thus, rather than introducing an interpreter, it may be preferable to have an interviewer who speaks the interviewee's language. This, of course, will not always be possible.

Conclusion

The presence of an interpreter makes interviewees say less compared to when they speak in their first language without an interpreter. In fact, talking through an interpreter did not result in more detail than speaking in a foreign language. Cues to deceit emerged when interviewers and interviewees shared the same first language or when the interviewee spoke in a foreign language. However, the former requires an interviewer who speaks the interviewee's first language, whereas the latter requires that the interviewee has a certain understanding of the foreign language s/he is requested to use in the interview. The presence

of an interpreter did not affect rapport in the present experiment. We do not rule out, however, that an interpreter will have an effect when the interviews take longer and when the rapport between interviewer and interviewee is more difficult to establish.

Acknowledgments

This work is funded by the High-Value Detainee Interrogation Group, J-FBI-12-194 awarded to the University of Portsmouth (UK). Any opinions, findings, and conclusions or recommendations expressed in this article are those of the authors and do not necessarily reflect the views of the U.S. Government.

References

- Abbe, A., & Brandon, S. E. (2012). The role of rapport in investigative interviewing: A Review. *Journal of Investigative Psychology and Offender Profiling*, *10*, 237-249. doi: 10.1002/jip. 1386
- Akca, C., & Elkilic, G. (2011). The effects of age and production mode on the bilingual performance of Turkish learners of English in a serial recall task. *International Journal of Psychology*, *46*, 9-17. doi: 10.1080/00207594.2010.503762
- Aranguri, C., Davidson, B., & Ramirez, R. (2006). Patterns of communication through interpreters: A detailed sociolinguistic analysis. *Journal of General Internal Medicine*, *21*, 623–629. doi: 10.1111/j.1525-1497.2006.00451.x
- Bailey, B. P., & Konstan, J. A. (2006). On the need for attention-aware systems: Measuring effects of interruption on task performance, error rate, and affective state. *Computers in Human Behavior*, *22*, 685-708. doi: 10.1016/j.chb.2005.12.009
- Berk-Seligson, S. (2000). Interpreting for the police: Issues in pre-trial phases of the judicial process. *International Journal of Speech Language and the Law*, *7*, 212–237.
- Bull, R. (2010). The investigative interviewing of children and other vulnerable witnesses: Psychological research and working/professional practice. *Legal and Criminological Psychology*, *15*, 5–24. doi: 10.1348/014466509X440160
- Bull, R., & Soukara, S. (2010). Four studies of what really happens in police interviews. In G. D. Lassiter, & C. A. Meissner (Eds.), *Police interrogations and false confessions: Current research, practice, and policy recommendations* (pp. 81–95). Washington, DC: American Psychological Association.
- Collins, R., Lincoln, R., & Frank, M. G. (2002). The effect of rapport in forensic interviewing. *Psychiatry, Psychology, and Law*, *9*, 69–78. doi: 10.1375/pplt.2002.9.1.69

- 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
- Dornyei, Z., & Scott, M. L. (1995). Communication strategies in a second language: Definitions and taxonomies. *Language Learning*, *47*, 173–210. doi: 10.1111/0023-8333.51997005
- Driskell, T., Blickensderfer, E. L., & Salas, E. (2013). Is three a crowd? Examining rapport in investigative interviews. *Group Dynamics: Theory, Research and Practice*, *17*, 1-13, doi: 10.1037/a0029686
- Drolet, A. L., & Morris, M. W. (2000). Rapport in conflict resolution: Accounting for how face to face contact fosters mutual cooperation in mixed-motive conflicts. *Journal of Experimental Social Psychology*, *36*, 26–50. doi: 10.1006/jesp.1999.1395
- Ekman, P. (1985). *Telling lies: Clues to deceit in the marketplace, politics and marriage*. New York: W. W. Norton.
- Evans, J. R., Houston, K. A., & Meissner, C. A. (2012). A positive, collaborative, and theoretically-based approach to improving deception detection. *Journal of Applied Research in Memory and Cognition*, *1*, 122–123. doi: 10.1016/j.jarmac.2012.02.004
- Evans, J. R., Meissner, C. A., Micheal, S. W., & Brandon, S. E. (2013). Validating a new assessment method for deception detection: Introducing a psychologically based credibility assessment tool. *Journal of Applied Research in Memory and Cognition*, *2*, 33-41. doi: 10.1016/j.jarmac.2013.02.002
- Fein, R. A. (Ed.). (2006). *Educing information: Interrogation: Science and art: Intelligence*

Science Board Phase I report. Washington, DC: United States Department of Defense, National Defense Intelligence College.

- Fisher, R. P. (2010). Interviewing cooperative witnesses. *Legal and Criminological Psychology, 15*, 25-38. doi: 10.1348/135532509X441891
- Huang, C. P. (2010). Exploring factors affecting the use of oral communication strategies. *Department of Applied Foreign Languages, Lunghwa University of Science and Technology*. Retrieved from:
[http://www.lhu.edu.tw/m/oaa/synthetic/publish/publish/30/8.%E9%BB%83%E7%A7%8B%E8%90%8D-Exploring%20Factors%20Affecting%20the%20Use%20of%20Oral %20Communication%20Strategies.pdf](http://www.lhu.edu.tw/m/oaa/synthetic/publish/publish/30/8.%E9%BB%83%E7%A7%8B%E8%90%8D-Exploring%20Factors%20Affecting%20the%20Use%20of%20Oral%20Communication%20Strategies.pdf)
- 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: 10.1037/a0033571
- Karliner, L. S., Pérez-Stable, E. J., & Gildengorin, G. (2004). The language divide: The importance of training in the use of interpreters for outpatient practice. *Journal of General Internal Medicine, 19*, 175–183. doi: 10.1111/j.1525-1497.2004.30268.x
- Leal, S., Vrij, A., Warmelink, L., & Fisher, R. P. (2012). You can't hide your telephone lies: Priming as an aid to detect deception in insurance telephone calls. *Legal and Criminological Psychology*.
- Macintosh, G. (2009). The role of rapport in professional services: Antecedents and outcomes. *Journal of Services Marketing, 23*, 70–78. doi: 10.1108/08876040910946332
- Mann, S., Vrij, A., Shaw, D. J., Leal, S., Ewens, S., Hillman, J., . . . Fisher, R. P.

- (2012). Two heads are better than one? How to effectively use two interviewers to elicit cues to deception. *Legal and Criminological Psychology, 18*, 324-340.
doi:10.1111/j.20448333.2012.02055
- Masip, J., Sporer, S., Garrido, E., & Herrero, C. (2005). The detection of deception with the reality monitoring approach: A review of the empirical evidence. *Psychology, Crime, & Law, 11*, 99-122. doi: 10.1080/10683160410001726356
- Morris, L. W., & Perez, T. L. (1972). Effects of test-interruption on emotional arousal and performance. *Psychological reports, 31*, 559-564. Retrieved from:
<http://www.amsciepub.com/doi/pdf/10.2466/pr0.1972.31.2.559>
- 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
- Russell, S. (2002). "Three's a crowd: shifting dynamics in the interpreted interview". In J. Cotterill, (Ed.). *Language in the legal process*. (pp. 111-126). New York: Palgrave Macmillan.
- Simmel, G. (1964). *The sociology of Georg Simmel* (K. H. Wolf, Ed. and Trans.). Glencoe, IL: Free Press.
- Soufan, A. H. (2011). *The Black Banners: Inside the hunt for Al Qaeda*. London: Allen Lane.
- Ullman, M. T. (2001). The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism: language and Cognitive, 4*, 105-122.
doi: 10.1017/S1366728901000220
- Vallano, J. P., & Schreiber Compo, N. (2011). A comfortable witness is a good witness: Rapport-building and susceptibility to misinformation in an investigative mock-crime interview. *Applied Cognitive Psychology, 25*, 960-970. doi: 10.1002/acp.1789
- Valley, K., Thompson, L., Gibbons, R., Bazerman, M. H. (2002). "How communication

- improves efficiency in bargaining games". *Games and Economic Behavior*, 38, 127–155. doi: 10.1006/game.2001.0855
- 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.
- Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities*. Chichester: John Wiley and Sons.
- Vrij, A., Leal, S., Mann, S., & Fisher, R. (2012). Imposing cognitive load to elicit cues to deceit: Inducing the reverse order technique naturally. *Psychology, Crime & Law*, 18, 579-594. doi: 10.1080/10683160902776843
- Vrij, A., Mann, S., Fisher, R., Leal, S., Milne, B., & Bull, R. (2008). Increasing cognitive load to facilitate lie detection: The benefit of recalling an event in reverse order. *Law and Human Behavior*, 32, 253-265. doi: 10.1007/s10979-007-9103-y
- Vrij, A., Mann, S., Leal, S., & Fisher, R. (2012). Is anyone there? Drawings as a tool to detect deceit in occupation interviews. *Psychology, Crime & Law*, 18, 377-388. doi: 10.1080/1068316x.2010.498422
- Walsh, D., & Bull, R. (2012). Examining rapport in investigative interviews with suspects: Does its building and maintenance work? *Journal of Police and Criminal Psychology*, 27, 73-84. doi: 10.1007/s11896-011-9087-x
- Wang, Y. P., Xue, G. M., Chen, C. S., Xue, F., & Dong, Q. (2007). Neural bases of asymmetric language switching in second-language learners: An ER-fMRI study. *NeuroImage*, 35, 862-870. doi: 10.1016/j.neuroimage.2006.09.054

¹ Objective coding was also carried out on the data. A coder read the transcripts and rated the number of ‘visual’, ‘spatial’, ‘temporal’, ‘auditory’ and ‘action’ details in the five detail-eliciting questions. A second coder rated a sub-sample of 60 transcripts (25%). The inter-rater reliability between the two coders for the objective detail was very good (Intra-class Correlation Coefficient (ICC) = .87). We correlated objective and subjective coding. The correlation revealed a significant overlap between the objective coding and subjective coding ($r = .81$). This indicates that the subjective ratings gave a good indication of the details actually present in the statements. To avoid repetition we decided to only report the subjective analyses. We opted for presenting the subjective analyses as investigators in real life make such judgements. The results for the objective and subjective details showed a 100% match. That is, all effects that were significant in the objective analyses were also significant in the subjective analyses and vice versa. Readers wishing more information about the coding and statistics of the objective coding can do so by contacting the authors.

² Occasionally during the interviews, interviewees asked for questions to be clarified or the interviewer intervened when the participant’s answer did not match the question. This explains the average seven segments in talk in the long consecutive condition for the five detail eliciting questions.

³ As we reported above, a 2(Veracity) X 2 (Interpreter Condition) X 2 (Interpreter Experience) ANOVA resulted in a main effect for Interpreter Experience only.