Abstract

Background. A common strategy in interviewing is to repeatedly focus on the same topics, for example by asking to recall an event first in chronological order and then in reverse order. We examined the effect of changing interviewers between the two questions or keeping the same interviewers throughout on cues to deception. Truth tellers may be most encouraged to recall again what they have witnessed when confronted with new interviewers, as these new interviewers have not heard their story before. Liars may be most encouraged to recall again their story when confronted with the same interviewers, realising that these interviewers will check for consistency in their answers. The impact of changing interviewers should lead to more pronounced differences between truth tellers and liars in terms of detail and repetition in the 'Changed Interviewers' condition compared to the 'Same Interviewers' condition.

Method. Participants were interviewed by two interviewers about a mock security meeting they attended. In half the interviews the same two interviewers remained throughout, and in the other half two new interviewers took over half-way through.

Results. As predicted, differences between truth tellers and liars in terms of detail and repetition were most pronounced in the 'Changed Interviewers' condition.

Conclusions. Changing interviewers during an interview effectively differentiates liars and truth tellers with respect to detail and repetition. We discuss this finding and its place within investigative interviewing and deception detection literature.

'We'll take it from here': the effect of changing interviewers in information gathering interviews

A common strategy employed by investigators is to repeatedly focus on the same topics throughout an investigative interview (Fisher, 2010). This technique is aimed at giving interviewees the chance to say everything they know, and to ensure that nothing is missed. One way to achieve this is by asking the same question in a different format, for example, by asking an interviewee to recall an event in reverse chronological order (Fisher & Geiselman, 1992). Although the invitation to recall an event in reverse order is mostly used in interviews with collective witnesses (Fisher, 2010), it has also been used when interviewing suspects (Geiselman, 2012). In the present experiment we also used this reverse order recall in interviews with (mock) suspects. We were particularly interested to know whether using different interviewers in different stages of an interview can heighten the verbal differences between truth tellers and liars. We looked at the effect of (i) having the same interviewers ask both a normal chronological order question at Stage One and a reverse order question at Stage Two ('Same Interviewers' condition) and (ii) changing interviewers after Stage One so that the interviewers who ask the normal order question at Stage One are different from the interviewers who ask the reverse order question at Stage Two ('Changed Interviewers' condition).

For truth tellers, the reverse order question at Stage Two of the interview and the introduction of new interviewers at Stage Two may elicit new information. Anderson and Pichert (1978) showed that recalling an event from a different perspective stimulates an interviewee to think again, subsequently evoking new information. Similarly, recalling an event in reverse order in the presence of two new interviewers may encourage an interviewee to think harder about the event in question. For liars, the reverse order question or the introduction of new interviewers is less likely to result in new information. Liars may find it difficult to come up with new information that sounds plausible (Köhnken, 1996, 2004; Leal,

Vrij, Warmelink, Vernham, & Fisher, in press). Equally, liars may be discouraged from being detailed as this increases the risk of that information being falsified (Hartwig, Granhag, & Strömwall, 2007; Masip & Ces, 2011; Nahari, Vrij, & Fisher, 2012, 2013). Therefore, methods that can evoke new information from truth tellers (reverse order recall or the introduction of new interviewers) may not have such an effect on liars. We thus predicted that at Stage Two truth tellers would provide more new details than liars (Hypothesis 1a). We further predicted that truth tellers would provide more new details at Stage Two than liars, particularly when confronted with new interviewers (Hypothesis 1b).

Apart from providing new information at Stage Two, we were interested in the amount of detail provided during the interview and the amount of repetition in answers between Stages One and Two. Having the same or new interviewers at Stage Two may have opposite effects on truth tellers and liars, notably in terms of detail at Stage Two and the amount of repetition between Stages One and Two responses. Truth tellers will realise that the new interviewers have not heard what they said at Stage One. Therefore, they may be encouraged to be detailed at Stage Two and to repeat what they have said previously when new interviewers are present. Of course, a good reason for truth tellers to be detailed at Stage Two and to repeat themselves when the same interviewers are present is to show consistency in their answers. Consistency is widely seen as a sign of honesty (Strömwall, Granhag, & Hartwig, 2004). However, because truth tellers typically take their credibility for granted (DePaulo et al., 2003; Gilovich, Savitsky, & Medvec, 1998) and have no reason to believe that interviewers will doubt them, they are typically not concerned with conveying their honesty (Kassin, 2005; Kassin & Gudjonsson, 2004). Given that the same interviewers have already heard their story, and given that reverse order recall is cognitively demanding, truth tellers may be less motivated to tell them the entire story again.

In contrast, liars may be particularly keen to be detailed at Stage Two and to repeat what they have said before when interviewed by the same interviewers. Unlike truth tellers,

liars do not take their credibility for granted. Also, given that consistency is perceived as a sign of honesty, they may be keen to show consistency in their answers when the same interviewers are present, resulting in them being detailed at Stage Two and repeating what they have said at Stage One. When two new interviewers take over the interview, the idea that their answers will be compared is less obvious (Vrij, Leal, Mann, & Granhag, 2011). This notion, combined with the difficulty of reverse order recall, may result in liars providing a shorter statement that includes less detail when two new interviewers are present.

We therefore predicted that, because liars may find it difficult to fabricate information or may fear that their responses will be falsified, overall truth tellers will be more detailed than liars (Hypothesis 2a). We further predicted that truth tellers will be *more* detailed in the 'Changed Interviewers' condition compared to the 'Same Interviewers' condition, whereas liars will be *less* detailed in the 'Changed Interviewers' condition compared to the 'Same Interviewers' condition. As a result, differences in detail between truth tellers and liars will be more pronounced in the 'Changed Interviewers' condition than in the 'Same Interviewers' condition (Hypothesis 2b). We finally predicted that truth tellers will repeat more at Stage Two what they have said at Stage One in the 'Changed Interviewers' condition compared to the 'Same Interviewers' condition, and conversely, that liars will show less repetition in the 'Changed Interviewers' condition compared to the 'Same Interviewers' condition. As a result, differences in repetition between truth tellers and liars will be more pronounced in the 'Changed Interviewers' condition than in the 'Same Interviewers' condition (Hypothesis 3).

The idea that liars display less repetition than truth tellers when the questions are asked by different interviewers has been noted twice before (Jundi, Vrij, Hope, Mann, & Hillman, in press; Vrij, Leal, Mann, & Granhag, 2011). On both occasions, the authors speculated that the change of interviewer may have been responsible for the reduced overlap in liars' (compared to truth tellers') responses. In this article we put this suggestion to an empirical test.

Method

Design

We used a 2 (Veracity: truth vs lie) X 2 (Interviewer: same interviewers vs changed interviewers) between-subjects design with the following dependent variables: (i) the number of details in Stage One recollections; (ii) the number of details in Stage Two recollections; (iii) the frequency of 'meeting events' (see below) recalled in Stage One recollections; (iv) the frequency of meeting events recalled in Stage Two recollections; and (v) the amount of repetition between Stage One and Stage Two recollections.

Participants

A total of 165 participants (58 males and 107 females) took part in the study. The sample was made up of undergraduate students ($N = 144$), university staff ($N = 18$), and members of the general public ($N = 3$). The average age was $M = 22.56$ years ($SD = 6.64$).

Procedure

Participants were recruited via posters, leaflets, and online advertisements on the University's staff and student portals. An advert was also placed in a local newspaper. Participants were invited to play the role of a secret agent, attending a meeting and then an interview. The advert provided contact details and offered a £5 reward to those who were convincing in the interview.

After arriving at the Department, participants were informed that they were going to play the role of an intelligence officer, attending a secret security meeting. The participants were then directed to a small room where the meeting took place.

The meeting. The purpose of the meeting was to vote on a suitable location to plant a spy device, and included a visual presentation of the following details: the three members

(confederates) present at the meeting; the spy device and its physical and technical features; and the shortlisted locations suitable to host the device, including floor plans and details on suitability. Before the third and final location (a hotel reception) was presented, the meeting was interrupted, triggering a vote on which location should host the device. The outcome of this vote was pre-determined, and the participant's vote could not affect the result. The participant then returned to the room where s/he started the experiment. It was at this stage that all participants were randomly assigned to a veracity condition, either truth tellers ($N = 82$) or liars ($N = 83$).

Prior to being interviewed, the truth tellers were informed that a sister organisation, HMR, was aware of the meeting they had just attended. Truth tellers were told that HMR knew the target and have been pursuing this person for some time. Truth tellers were therefore instructed to have an interview with HMR and volunteer information about its content. They were told that their task is to fully cooperate with the interviewers and to volunteer all the information they ask for. In addition, the experimenter informed truth tellers that they would receive £5 as a reward if they managed to convince the interviewers. Alternatively, if they failed to convince the interviewers, they would have to write a report about the meeting instead. (All participants in this experiment, truth tellers and liars, received the £5 for taking part, and no one had to write a report). Finally, before being led to the interview room, truth tellers completed a 'Pre-Interview Questionnaire' which measured how motivated they were to perform well in the interview. This was measured on a 5-point Likert scale ranging from [1] 'Not at all motivated' to [5] 'Very motivated'.

On returning from the meeting, the liars were first informed that a foreign intelligence agency, 'EFA', was aware of the meeting they had attended. To prevent an investigation into HMI, liars were told that they must now meet with EFA and do their utmost to convince the EFA interviewers that they are telling the truth. The liars' task required them to provide a mixture of truthful and false information. The truthful information, it was argued, would help

convince EFA that they are being cooperative. Therefore, liars were first instructed to be completely honest about the room where the meeting took place, and the location that did not win the vote. Second, they were told that they must be completely dishonest about the location that did win the vote. In all cases, liars were instructed to say that the Hotel Reception was the location selected to host the device, and make up the following details: a floor plan, one reason why it is a suitable location, and one reason why it is not. Finally, liars were told that EFA knew something about the spy device and who attended the meeting, though it's not clear what they knew. As a result, to appear cooperative, liars' third task was to provide a mixture of truthful and false information about the device, and the people present at the meeting. How much truthful and false information provided was at the discretion of the participant. Liars also received the same information as truth tellers regarding the reward for being convincing and the penalty of being unconvincing and completed the same 'Pre-Interview Questionnaire'. Liars were then left alone and given as much time as they needed to think about the details of the Hotel Reception. They were not provided with any writing materials during this time. Also, the time (in seconds) that liars took to consider what they would say about the Hotel Reception was recorded by the experimenter ($M = 264.63$, $SD = 149.01$, ranging between 70 and 900 seconds).

The Interview. Our interview protocol consisted of two interviewers. A two interviewer protocol is frequently adopted by the police (Driskell, Blickensderfer, & Salas, 2013; Sim & Lamb, 2012) and intelligence agencies (Soufan, 2011), and thus reflects real life. We used four interviewers, all female, aged between 31-54. The interviewers had not been formally trained in investigative interviewing, but had extensive experience interviewing participants in previous experiments. All interviewers were also blind to the experimental hypotheses. Before the interview commenced, the speaking interviewer introduced herself and the silent interviewer, stressing that the silent interviewer has been trained to detect deception in interviews. The interviewee was also informed that the

interviewers knew that s/he had attended the meeting with HMI, and that the purpose of the meeting was to select a location to host a spy device.

The interview schedule consisted of two stages. Stage One required participants to recall what happened during the meeting in normal chronological order, and Stage Two required participants to recall the meeting in reverse order. Participants also responded to questions about the meeting room, the device, the locations, and the confederates present at the meeting. While participants' responses to these questions were also transcribed and coded, the results are beyond the scope of the present article, and the data is not included in our analyses.

We implemented our 'Interviewer' manipulation immediately after participants had answered questions in Stage One. In the 'Same Interviewers' condition, a confederate entered the room and passed a note to the speaking interviewer. The confederate then exited the room, and both interviewers read the note without revealing its contents to the interviewee. In the 'Changed Interviewers' condition, two new interviewers (one silent and one speaking) entered the room. Following the exchange of a note between the new and existing speaking interviewers, the existing pair of interviewers exited the room and were replaced by the new pair. Participants were not given an indication as to why the interviewers changed. The speaking interviewer from the new pair introduced herself and the new silent interviewer, explaining that they would be conducting the interview henceforth. Stage Two commenced after the interviewers had changed.

All participants then completed a 'Post Interview Questionnaire', in which they estimated the likelihood that they would receive the £5 (measured on a 7-point Likert scale ranging from [1] 'Not at all likely' to [7] 'Very likely), and the likelihood that they would have to write a statement (measured on a 7-point Likert scale ranging from [1] 'Not at all likely' to [7] 'Very likely).

The questionnaire also examined what the participants could remember about the device, the locations, and the confederates. Knowing participants' actual memory of the meeting allowed us to determine exactly how much accurate information was volunteered by each participant. After completing the questionnaire, the participants were debriefed and received £5 for taking part.

Counterbalancing. We counterbalanced the locations presented in the meeting, the location that was voted for in the meeting, and the interviewers used in the interviews.

Coding

Overview. We coded the amount of detail conveyed in Stage One and Stage Two and we did this both subjectively and objectively. We also looked at the amount of repetition between Stage One and Stage Two responses, and for any new information provided at Stage Two that was not provided at Stage One, again coded subjectively and objectively.

Subjective Detail. The interviews were videotaped, audiotaped and transcribed, and the subjective ratings of detail were derived from the transcripts. A coder blind to the hypotheses and experimental conditions rated participants' responses using a 7-point scale, ranging from 1 (low on valuable detail) to 7 (high on valuable detail). The responses were not coded, however, according to the accuracy of the details provided. The amount of valuable detail was determined by the degree to which the participant went into detail about events or topics they introduced. For example, the sentence 'The meeting started with some introductions, then I saw the device, then we looked at some locations, then we had a vote, and then I left' would be rated as low on valuable detail, as topics are introduced without further description. In contrast, the sentence 'The meeting started with Mr. Black introducing himself, saying that he was the operations manager and that he had worked at the company for 7 years' offers a much more detailed account of individual topics, and therefore would have been rated high on valuable detail.

A second coder also blind to the hypotheses and experimental conditions coded for the same details using a sub sample of 42 transcripts (25%). The inter-rater reliability between the two coders for the Stage One (Intra-class Correlation Coefficient, ICC = .93) and Stage Two details (ICC = .86) was very high.

Subjective Repetition. To determine the amount of repetition between participants' responses, again using the transcripts, Stage Two answers were compared with Stage One answers. A coder blind to the hypotheses and experimental conditions compared participants' responses, rating them using a 7-point scale ranging from 1 (low in repetition) to 7 (high in repetition). A response would be rated as high in repetition if much of the detail mentioned in Stage One recollections were repeated in Stage Two recollections, and no contradictions emerged between the Stage One and Stage Two recollections. It should be noted, however, that contradictions rarely occurred. Only eight contradictions were noted in the 165 transcripts by the objective coder. A similar measure of overlap was used by Vrij et al. (2009). A second coder, also blind to the hypotheses and experimental conditions, also coded for overlap using a sub sample of 42 transcripts (25%). The inter-rater reliability between the two coders (ICC = .60) was satisfactory.

Subjective New Information. To determine the amount of new valuable detail in participants' Stage Two responses, Stage Two answers were compared with Stage One answers. A coder blind to the hypotheses and experimental conditions compared participants' responses, rating them using a 7-point scale, ranging from 1 (low in new information) to 7 (high in new information). A response would be rated high in new information if the participant introduced new information, and subsequently went into detail about the new event of topic. Therefore, the coding of new information was identical to the coding of detail, except that it only applied to new information. A second coder also blind to the hypotheses and experimental conditions coded for the same details using a sub sample of 42 transcripts

(25%). The inter-rater reliability between the two coders for new information (Intra-class Correlation Coefficient, ICC = .92) was very high.

Objective Detail. The verbal coding was also derived from the transcripts. A coder blind to the hypotheses and experimental conditions read each answer carefully and marked every detail the interviewee gave. A second coder also blind to the hypotheses and experimental conditions coded for details using a sub sample of 42 transcripts (25%). The inter-rater reliability between the two coders for the Stage One (Intra-class Correlation Coefficient, ICC = .98) and Stage Two details (ICC = .97) was very high.

Meeting Events. We further compiled a checklist of, in our view, 18 key events of the meeting. This checklist was used to create the objective overlap variable, discussed below. A coder blind to the hypotheses and experimental conditions scored participants' responses using this checklist. Participants could score a maximum of 18 points if they mentioned all events on the checklist, and they scored 0 points if none of the details they mentioned were on the checklist. To demonstrate how the checklist works, item 16 on the checklist was: 'All members casted their votes by a show of hands for each location'. In order to score a point for each item, the participant must clearly make a reference to that event. If a participant's response did not have a clear meaning, then s/he did not score a point for that particular item. The complete checklist of key events coded is included in Appendix 1.

A second coder also blind to the hypotheses and experimental conditions coded participants' responses using a sub sample of 42 transcripts (25%). The inter-rater reliability between the two coders for the Stage One (ICC = .99) and Stage Two (ICC= .98) checklist scores was high. The 'total detail' and 'checklist' variables were significantly correlated with each other in both Stage One, $r(165) = .69$, $p < .001$, and Stage Two, $r(165) = .60$, $p < .001$.

Objective Repetition. To calculate the overlap score, a coder blind to the hypotheses and experimental conditions compared the checklist scores from the Stage One and Stage

Two responses. Each event that the participant mentioned in both responses counted towards their overall repetition score. The overlap score could range from 0 to 18. From that score deducted were the few contradictions that occurred in the transcripts.

A second coder, also blind to the hypotheses and experimental conditions, coded participants' responses using a sub sample of 42 transcripts (25%). The inter-rater reliability between the two coders for repetition (ICC = .92) was high.

Objective New Information. To calculate the new information score, a coder blind to the hypotheses and experimental conditions compared the checklist scores from the Stage One and Stage Two responses. Each event that the participant mentioned in Stage Two but not in Stage One counted towards their new information score. The new information score could range from 0 to 18.

A second coder also blind to the hypotheses and experimental conditions coded participants' responses using a sub sample of 42 transcripts (25%). The inter-rater reliability between the two coders for new information (ICC = .96) was high.

Correlations between Subjective and Objective Codings. The correlation between subjective and objective detail coding was very satisfactory, $r = .67$ for Stage One detail and $r = .67$ for Stage Two detail. The correlation between subjective new information and objective new information was also satisfactory, $r = .55$. The correlation between subjective and objective repetition coding was lower, $r = .44$, but it should be noted that the subjective and objective repetition coding measured slightly different things. Subjective repetition measures the repetition between the statements as a whole, whereas objective repetition measures repetition between mentioning key events of the meeting only.

Results

Motivation, Incentive and Penalty. Participants were motivated to do well in the experiment ([$M = 4.33$, $SD = .69$] on a 5-point Likert scale), with 43% reporting that they were 'quite motivated' (score of 4), and 44% 'very motivated' (score of 5). A 2 (Veracity) X 2 (Interviewers) ANOVA with motivation as the dependent variable revealed no significant main or interaction effects (all Fs <.77, all ps >.380) indicating that participants' motivation level was similar amongst the experimental conditions.

A 2 (Veracity) X 2 (Interviewers) ANOVA regarding the likelihood of receiving the £5 reward resulted in a main effect for Veracity, F (1, 161) = 41.85, p <.001, $\eta^2 = .21$, $d =$ 1.01. Truth tellers ($M = 4.83$, $SD = 1.33$) more than liars ($M = 3.40$, $SD = 1.51$) were inclined to think that they would receive a £5 incentive. The Interviewers main effect and the Veracity X Interviewers interaction effect were not significant, both Fs <.3.71, both ps >.056.

A 2 (Veracity) X 2 (Interviewers) ANOVA regarding the likelihood of having to write a report resulted in a main effect for Veracity, F (1, 161) = 42.94, p <.001, $\eta^2 = .21$, $d = 1.04$. Liars ($M = 4.52$, $SD = 1.62$) more than truth tellers ($M = 3.03$, $SD = 1.23$) were inclined to think that they would be requested to write an essay. The Interviewer main effect and the Veracity X Interviewers interaction effect were not significant, both Fs <2.15, both ps >.144. Taken together, the above analyses suggest that participants were motivated to be convincing and that the incentive and penalty appeared realistic.

Meeting Recollections. Three 2 (Veracity) X 2 (Interviewers) ANOVAs examining participants' post interview recollections of the device, the locations, and the confederates resulted in no significant main or interaction effects (all Fs <1.51, all ps >.221), indicating that participants' memory of the meeting was similar amongst the experimental conditions. The lack of an 'Interviewers' effect (or interaction) suggests that any differences in detail reported during the interview reflect the strategies employed by truth tellers and liars. Participants correctly recalled 84.75% of the device characteristics, 86.50% of the locations

characteristics and 67.44% of the confederate's characteristics. This represents a satisfactory memory of the meeting.

Hypothesis Testing

New Information. A 2 (Veracity) X 2 (Interviewers) ANCOVA with participants' additions at Stage Two (new information) as the dependent variable (*subjective* coding) and the Stage One subjective details score as covariate[1] revealed a significant main effect for Veracity, $F (1, 160) = 12.62$, $p = .001$, $\eta^2 = .07$, $d = .42$. Truth tellers ($M = 2.38$, $SD = 1.30$) gave more new information than liars ($M = 1.90$, $SD = 1.00$), supporting Hypothesis 1a. The Interviewers main effect, $F (1, 160) = 1.45$, $p = .23$, $\eta^2 = .009$, and the Veracity X Interviewers interaction effect, $F (1, 160) = 2.93$, $p = .089$, $\eta^2 = .02$, were not significant. The absence of an interaction-effect means that Hypothesis 1b was not supported.

A 2 (Veracity) X 2 (Interviewer) ANCOVA with participants' *objective* new information at Stage Two as the dependent variable and the Stage One objective details score as covariate revealed a significant main effect for Veracity, $F (1, 160) = 4.06$, $p = .045$, $\eta^2 = .025$, $d = .08$. Truth tellers ($M = 2.46$, SD 2.28) had more new information than liars (M = 2.28, SD = 2.01), supporting Hypothesis 1a. The Interviewer main effect, $F (1, 160) = .20$, $p = .660$, $\eta^2 = .001$, and the Veracity X Interviewer interaction effect, $F (1, 160) = 0.01$, $p = .926$, $\eta^2 < .001$, were not significant. The absence of an interaction-effect means that Hypothesis 1b was not supported.

Detail. A 2 (Veracity) X 2 (Interviewers) ANCOVA with the Stage Two *subjective* details as dependent variable was carried out. Since the Stage Two subjective detail are linked with the Stage One subjective detail (we compare Stage One and Stage Two detail within the same interviewee) we introduced subjective detail at Stage One as covariate (which was significant, $F (1, 160) = 35.87$, $p <.001$, $\eta^2 = .18$).[2] The analysis revealed a main effect for Veracity, $F (1, 160) = 13.90$, $p <.001$, $\eta^2 = .08$, $d = .68$. Truth tellers ($M = 4.45$, SD

= 1.28) provided significantly more detailed responses than liars ($M = 3.63$, $SD = 1.13$), supporting Hypothesis 2a. The Interviewers main effect was not significant, $F (1, 160) = .51$, $p = .475$, $\eta^2 = <.01$, but the Veracity X Interviewers interaction effect was, $F (1, 160) = 7.83$, $p = .006$, $\eta^2 = .05$.

Insert Figure 1 about here

Figure 1 shows that truth tellers gave more detail in the 'Changed Interviewers' condition than in the 'Same Interviewers' condition, whereas liars gave less detail in the 'Changed Interviewers' condition than in the 'Same Interviewers' condition. In the 'Same Interviewers' condition, truth tellers ($M = 4.17$, $SD = 1.10$) and liars were equally detailed ($M = 3.84$, $SD = 1.21$), $F (1, 82) = .37$, $p = .548$, $\eta^2 <.01$, $d = .29$ In the 'Changed Interviewers' condition, truth tellers ($M = 4.75$, $SD = 1.39$) were more detailed than liars ($M = 3.40$, $SD = 1.01$), $F (1, 77) = 19.27$, $p <.001$, $\eta^2 = .20$, $d = 1.11$. This supports Hypothesis 2b.

A 2 (Veracity) X 2 (Interviewer) ANCOVA was conducted with the *objective* details variable from Stage Two as dependent variable and the Stage One objective details as covariate. The analysis revealed a main effect for Veracity, $F (1, 160) = 21.13$, $p <.001$, $\eta^2 = .12$, $d = .97$. Truth tellers ($M = 28.91$, $SD = 12.06$) provided significantly more detailed responses that liars ($M = 19.00$, $SD = 7.87$). The Interviewer main effect was not significant, $F (1, 160) = .68$, $p = .410$, $\eta^2 = .004$, but the Veracity X Interviewer interaction effect was, $F (1, 160) = 5.49$, $p = .020$, $\eta^2 = .03$. In the 'Same Interviewers' condition truth tellers ($M = 27.29$, $SD = 10.85$) and liars ($M = 19.63$, $SD = 8.09$) gave a similar amount of detail, $F (1, 82) = 2.66$, $p = .107$, $\eta^2 = .03$, $d = .80$. In the 'Changed Interviewers' condition, truth tellers ($M = 30.63$, $SD = 13.13$) gave significantly more details than liars ($M = 18.33$, $SD = 7.67$), $F (1, 77) = 20.69$, $p <.001$, $\eta^2 = .21$, $d = 1.14$. This supports Hypothesis 2b.

Repetition. Next we analysed the repeated information data. A 2 (Veracity) X 2 (Interviewers) ANCOVA with participants' *subjective* repetition scores as the dependent

variable and the Stage One subjective details score as covariate (which was significant, F (1, 160) = 117.56, p <.001, η^2 = .42)[3]. The analysis revealed a significant main effect for Veracity, F (1, 160) = 6.41, p = .012, η^2 = .04, d = .55. Truth tellers (M = 3.76, SD 1.39) showed more repetition than liars (M = 3.06, SD = 1.13). The Interviewers main effect, F (1, 160) = .01, p = .926, η^2 < .001 was not significant, but the Veracity X Interviewers interaction effect was, F (1, 160) = 6.33, p = .013, η^2 = .04.

Insert Figure 2 about here

Figure 2 shows that truth tellers showed more repetition in the 'Changed Interviewers' condition than in the 'Same Interviewers' condition, whereas liars showed less repetition in the 'Changed Interviewers' condition than in the 'Same Interviewers' condition. In the 'Same Interviewers' condition, truth tellers' (M = 3.57, SD = 1.25) and liars' (M = 3.30, SD = 1.28) repetition did not differ from each other, F (1, 82) = .01, p = .920, η^2 < .01, d =.21. In the 'Changed Interviewers' condition, truth tellers repeated themselves more (M = 3.95, SD = 1.52) than liars (M = 2.80, SD = .88), F (1, 77) = 13.62, p <.001, η^2 = .15, d =.93. This supports Hypothesis 3.

A 2 (Veracity) X 2 (Interviewer) ANCOVA with participants' *objective* repetition scores as dependent variable and the Stage One objective detail as covariate revealed a main effect for Veracity, F (1, 160) = 10.79, p =.001, η^2 = .06, d = 0.80. Truth tellers repeated themselves more (M = 8.26, SD = 3.52) than liars (M = 5.81, SD = 2.62). The main effect for Interviewer was not significant, F (1, 160) = .01, p = .910, η^2 <.001, but the Veracity X Interviewer interaction effect was, F (1, 160) = 8.61, p = .004, η^2 = .05. In the 'Same Interviewers' condition, truth tellers (M = 7.86, SD = 3.54) and liars (M = 6.35, SD = 2.84) did not differ in terms of repetition, F (1, 82) = .009, p = .924, η^2 < .001. In the 'Changed Interviewers', truth tellers repeated themselves more (M = 8.78, SD = 3.43) than liars (M = 5.33, SD = 2.32), F (1, 77) = 21.88, p <.001, η^2 = .22, d = 1.17. This supports Hypothesis 3.

Discussion

In the present experiment truth tellers and liars were first asked to recall an event in chronological order (Stage One) and then in reverse order (Stage Two). We examined the effect of changing interviewers between the two stages or keeping the same interviewers throughout on cues to deceit (level of detail, new information, and repetition in answers).

The 'Changed interviewers' condition was more successful in eliciting cues to deceit than the 'Same Interviewers' condition. When the same interviewers were present throughout the interview, no differences emerged between truth tellers and liars in terms of the amount of detail in Stage Two, and the amount of repetition between Stages One and Two. In contrast, when confronted with new interviewers at Stage Two truth tellers were more detailed in Stage Two than liars, and their answers in Stages One and Two showed more repetition than the answers given by liars.

Truth tellers were more detailed in recalling again what they have witnessed when confronted with new interviewers than when confronted with the same interviewers. The new interviewers had not heard their story before, which may have encouraged the truth tellers to say more. In contrast, liars were less detailed in telling their story again when confronted with new interviewers than when confronted with the same interviewers. To convey honesty, liars may wish to produce consistent responses, but the need to do this may be less when confronted with new interviewers who have not heard their previous answer than when confronted with the same interviewers who have heard their previous answers before.

The findings that truth tellers and liars showed similar repetition in the 'Same Interviewers' condition and that liars showed less repetition than truth tellers in the 'Changed Interviewers' condition, sheds new light on the ongoing debate about whether liars are more or less consistent than truth tellers. The idea that liars are less consistent than truth tellers is a popular view amongst practitioners (Strömwall, Granhag, & Hartwig, 2004) and promoted in

police manuals (Vrij & Granhag, 2007), but research findings show that liars are not always less consistent than truth tellers (Fisher, Vrij, & Leins, 2013). In this respect, Granhag and colleagues introduced the 'reconstruct – versus repeat' hypothesis (Granhag & Strömwall, 1999, 2001; Granhag, Strömwall, & Jonsson, 2003), which states that when asked to report information for a second time, truth tellers will search their memory for the original event and will reconstruct their story again based on these memoires. In contrast, liars will think about what they have said the first time and will try to repeat this information. Repetition may lead to the same level or even to more consistency than reconstruction. The present results suggests that when confronted with two new interviewers, truth tellers' tendency to 'reconstruct' and liars' tendency to 'repeat' is weakened. The end result is that truth tellers repeat themselves more, and liars repeat themselves less. The notion that liars are likely to be less consistent than truth tellers only after a change of format is introduced, was also found by Leins, Vrij, & Fisher (2012). In their experiment, truthful participants had visited a room whereas deceptive participants had not. However, in the interview all participants claimed to have visited the room. Participants were asked to verbally recall the layout of the room twice, to sketch it twice, or to verbally recall it once and to sketch it once. Liars contradicted themselves more than truth tellers, but only in the 'verbal recall – drawing' condition. In other words, to find differences in consistency between truth tellers and liars, it may be necessary to introduce a change in strategy when asking for the same information, either by using different interviewers (the present experiment) or by asking the same question in different formats (Leins et al., 2012).

Our findings suggest that 'external factors' (such as changing versus maintaining the same interviewers) can have a profound effect on consistency in truth tellers' and liars' responses. It may well be that truth tellers and liars have different interview strategies which are highlighted by external factors. Future research could attempt to shed light on the external factors that highlight such strategies.

One of our hypotheses was not supported. Although truth tellers provided, as predicted, more new information at Stage Two than liars did, having the same or new interviewers at Stage Two did not affect the amount of new information truth tellers gave. We had predicted that truth tellers would give more new information with new interviewers, because we expected that this change in interviewers would stimulate them to search their memory more for detail. We can only speculate why the predicted effect did not occur. The request to report the event in reverse order already triggered truth tellers to provide new information and perhaps this task overshadowed the effect that new interviewers may have on eliciting new information.

We are aware that the 'Same Interviewers' condition is the standard procedure in the Cognitive Interview (CI, Fisher & Geiselman, 1992). We do not suggest that this procedure should be changed due to our findings. The reverse order question in the CI is designed to elicit *new* information. The Reverse Order question did elicit new information and this was not influenced by the interviewer condition. In the present experiment we were also interested in the *repetition of information previously reported*, something the CI is less concerned about. It is in this elicitation of *old* information where the interviewer condition differentiated truth tellers and liars. In other words, if the interviewer has no reason to believe that the interviewee is lying, which is often the case in the Cognitive Interview as it is mainly used when interviewing cooperative witnesses (Fisher, 2010), only the elicitation of new information is relevant and there is therefore no need to change interviewers during the interview.

The present experiment fits well in the new wave of 'interviewing to detect deception' research aimed at eliciting cues to deceit through specific interventions (Vrij & Granhag, 2012). Other strategies that have proved effective at eliciting cues to deceit are imposing cognitive load on interviewees (Evans, Meissner, Michael, & Brandon, 2013), encouraging truth tellers to say more (Leal, Vrij, Warmelink, Vernham, & Fisher, in press), asking

unexpected questions (Vrij et al., 2009) and introducing evidence in interviews in a strategic manner (Granhag, Strömwall, Willén, & Hartwig, 2013; Hartwig, Granhag, Strömwall, & Kronkvist, 2006).

Since truth tellers said more than liars, particularly in the 'Changed Interviewers' condition, we believe that this manipulation fits particularly well in the 'encouraging truth tellers to say more' research domain. Encouraging truth tellers to say more has several benefits. It addresses the core of investigative interviewing, which is to obtain as much information as possible from interviewees (Fisher, 2010). In addition, if truth tellers provide lots of information they are more likely to be believed, because the richer an account is perceived to be, the more likely it is to be believed (Bell & Loftus, 1989; Johnson, 2006; Johnson, Foley, Suengas, & Raye, 1988). Finally, the additional information truth tellers provide could provide leads for investigators to pursue.

Methods that encourage truth tellers to say more are unlikely to have the same effect on liars. First, liars may find it too cognitively demanding to add as many details as truth tellers do. Moreover, if liars do add a sufficient amount of detail, the additional information may be of lesser quality or may sound less plausible (Leal et al., in press). Finally, liars may be reluctant to add more information out of fear that it will provide leads to investigators and, consequently, give their lies away (Nahari, Vrij, & Fisher, in press).

In conclusion, the present experiment demonstrated the beneficial effect of changing interviewers half-way through an interview, notably when interviewers discuss the same topics twice over the course of an interview. In their second response, truth tellers provided more detail and repeated themselves more compared to liars, but only if the interviewers had changed half-way through the interview. Our findings challenge previous findings (Granhag & Strömwall, 1999, 2001; Granhag, Strömwall, & Jonsson, 2003) in that truth tellers appeared more consistent that liars. The present findings highlight the complex relationship

between consistency and deception, suggesting that consistency may depend on both internal (retrieval strategies) and external factors (interview setting).

Acknowledgements

References

Bell, B. E., & Loftus, E. F. (1989). Trivial persuasion in the courtroom: The power of (a few) minor details. *Journal of Personality and Social Psychology*, *56*, 669-679.

Burgoon, J. K. (1991). Relational message interpretations of touch, conversation distance, and posture. *Journal of Nonverbal Behaviour*, *15* (4), 233-259.

Burgoon, J. K., Birk, & T., Pfau, M. (1990). Nonverbal behaviours, persuasion, and credibility. *Human Communication Research*, *17* (1), 140-169.

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.

Evans, J. R., Meissner, C. A., Michael, 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.

Fisher, R. P. (2010). Interviewing cooperative witnesses. *Legal and Criminological Psychology*, *15*, 25-38.

Fisher, R. P., & Geiselman, R. E. (1992). *Memory enhancing techniques for investigative interviewing: The cognitive interview*. Springfield, Il: Charles C. Thomas.

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

Geiselman, R. E. (2012). The Cognitive Interview for suspects. *American Journal of Forensic Psychology*, *30*, 1-12.

Geiselman, R. E & Callot, R. (1990). Reverse and forward order recall of script based text. *Applied Cognitive Psychology, 4*, 141-144.

Gilbert, J. A. E., & Fisher, R. P. (2006). The effects of varied retrieval cues on reminiscence in eyewitness memory. *Applied Cognitive Psychology*, *20*, 723-739.

Gilovich, T., Savitsky, K., & Medvec, V. H. (1998). The illusion of transparency: Biased assessments of others' ability to read one's emotional states. *Journal of Personality and Social Psychology*, *75*, 332-346.

Granhag, P. A., & Strömwall, L. A. (1999). Repeated interrogations: Stretching the deception detection paradigm. *Expert Evidence: The International Journal of Behavioural Sciences in Legal Contexts*, *7*, 163-174.

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

Granhag, P. A., Strömwall, L. A., & Jonsson, A. C. (2003). Partners in crime: How liars in collusion betray themselves. *Journal of Applied Social Psychology*, *33*, 848-868.

Granhag, P.A., Strömwall, L.A., Willén, R., & Hartwig, M. (2013). Eliciting cues to deception by tactical disclosure of evidence: The first test of the Evidence Framing Matrix. *Legal and Criminological Psychology, 18,* 341-355. DOI. 10.1111/j.2044-8333.2012.02047.x

Hartwig, M., Granhag, P. A., Strömwall, L., & Kronkvist, O. (2006). Strategic use of evidence during police interrogations: When training to detect deception works. *Law and Human Behavior*, *30*, 603-619.

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

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

Johnson, M. K., Foley, M. A., Suengas, A. G., & Raye, C. L. (1988). Phenomenal characteristics of memories for perceived and imagined autobiographical events. *Journal of Experimental Psychology: General*, *117*, 371-376.

Jundi, S., Vrij, A., Hope, L., Mann, S., & Hillman, J. (in press). Establishing evidence through undercover and collective intelligence interviewing. *Psychology, Public Policy, & Law*.

Kahana, M. J. (1996). Associate retrieval processes in free recall. *Memory & Cognition. 24*, 103-109.

Kassin, S. M. (2005). On the psychology of confessions: Does innocence put innocents at risk? *American Psychologist*, *60*, 215-228.

Kassin, S. M., & Gudjonsson, G. H. (2004). The psychology of confessions: A review of the literature and issues. *Psychological Science in the Public Interest*, *5*, 33-67.

Köhnken, G. (1996). Social psychology and the law. In G. R. Semin, & K. Fiedler (Eds.), *Applied Social Psychology* (pp. 257-282). London, Great Britain: Sage Publications.

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.

Leal, S., Vrij, A., Warmelink, L., Vernham, Z., & Fisher, R. (in press). You can't hide your telephone lies: Providing a model statement as an aid to detect deception in insurance telephone calls. *Legal and Criminological Psychology*.

Leins, D., Fisher, R., & Vrij, A. (2012). Drawing on liars' lack of cognitive flexibility: Detecting deception through varying report modes. *Applied Cognitive Psychology, 26, 601-607*. DOI 10.1002/acp.2837.

Mann, S., Vrij, A., Shaw, D., Leal, S., Ewans, S., Hillman, J., Granhag, P. A., & Fisher, R. P. (in press). Two heads are better than one? How to effectively use two interviewers to elicit cues to deception. *Legal and Criminological Psychology*.

Münte, T. F., Schiltz, K., & Kutas, M. (1998). When temporal terms belie conceptual order. *Nature*, *395*, 71-73.

Nahari, G., Vrij, A., & Fisher, R. P. (2012a). Does the truth come out in the writing? SCAN as a lie detection tool. *Law & Human Behavior*, 36, 68-76.

Nahari, G., Vrij, A., & Fisher, R. P. (in press). Exploiting liars' verbal strategies by examining the verifiability of details. *Legal and Criminological Psychology*.

Seay, T. A., & Alterkrise, M. K. (1979). Verbal and nonverbal behaviour in judgments of facilitative conditions. *Journal of Counselling Psychology*, *26*, 108-119.

Shaw, D. J, Vrij, A., Mann, S., Leal, S., & Hillman, J. (2013). Expect the unexpected? Variations in question type elicit cues to deception in joint interviewer contexts. *Applied Cognitive Psychology.*

Sim, M. P. Y., & Lamb, M. E. (2012). Police perceptions of interviews with juvenile suspects. Paper presented at the Annual Conference of the American Psychology-Law Society, Puerto Rico, 14-17 March 2012.

Smith-Hanen, S. S. (1977). Effects of nonverbal behaviour on judged levels of counsellor warmth and empathy. *Journal of Counselling Psychology*, *24*, 87-91.   Soufan, A.

H. (2011). *The black banners: The inside story of 9/11 and the war against al-Qaeda*. New York: W. W. Norton & Company.

Soufan, A. H. (2011). *The black banners: The inside story of 9/11 and the war against al-Qaeda*. New York: W. W. Norton & Company.

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

Vrij, A., & Granhag, P. A. (2007). Interviewing to detect deception. In S. A. Christianson (Ed.), *Offenders' memories of violent crimes* (pp. 279-304). Chichester, England: John Wiley & Sons, Ltd.

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

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., & Granhag, P. A. (2011). A comparison between lying about intentions and past activities: Verbal cues and detection accuracy. *Applied Cognitive Psychology*, *25*, 212-218. DOI 10.1002/acp.1665.
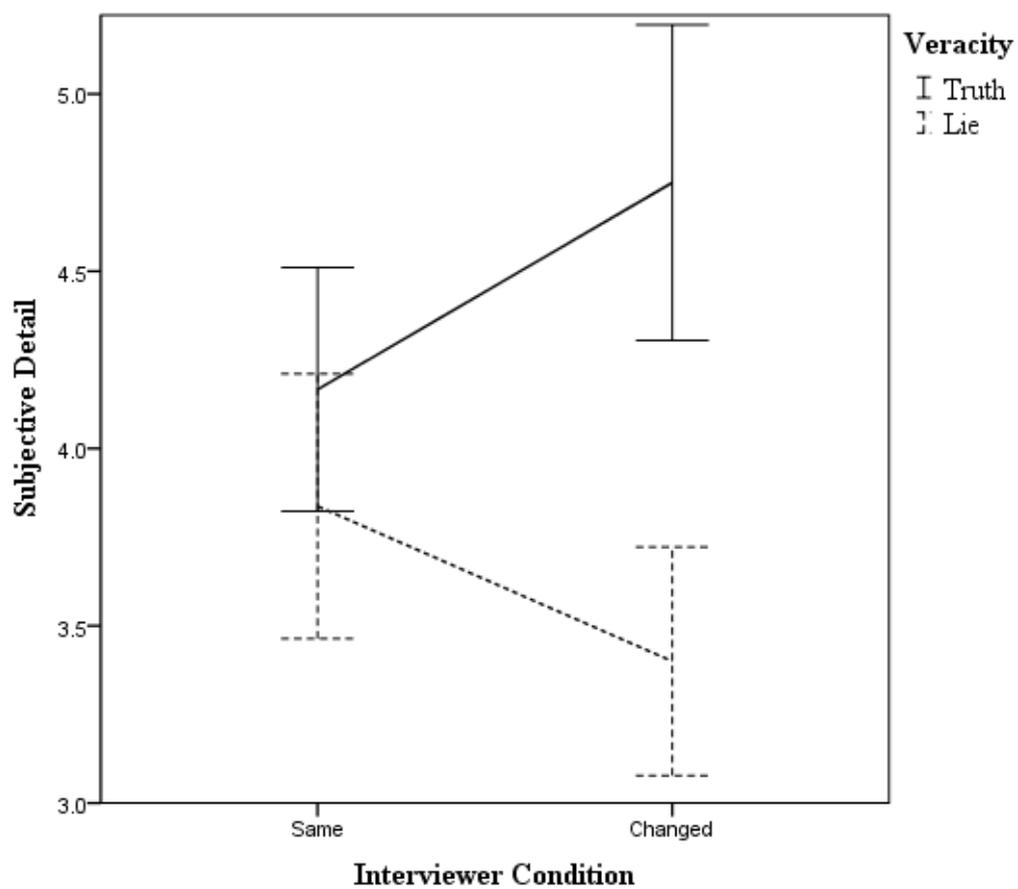
Figure 1. Subjective detail at Stage Two as a Function of Veracity and Interviewer. Error bars
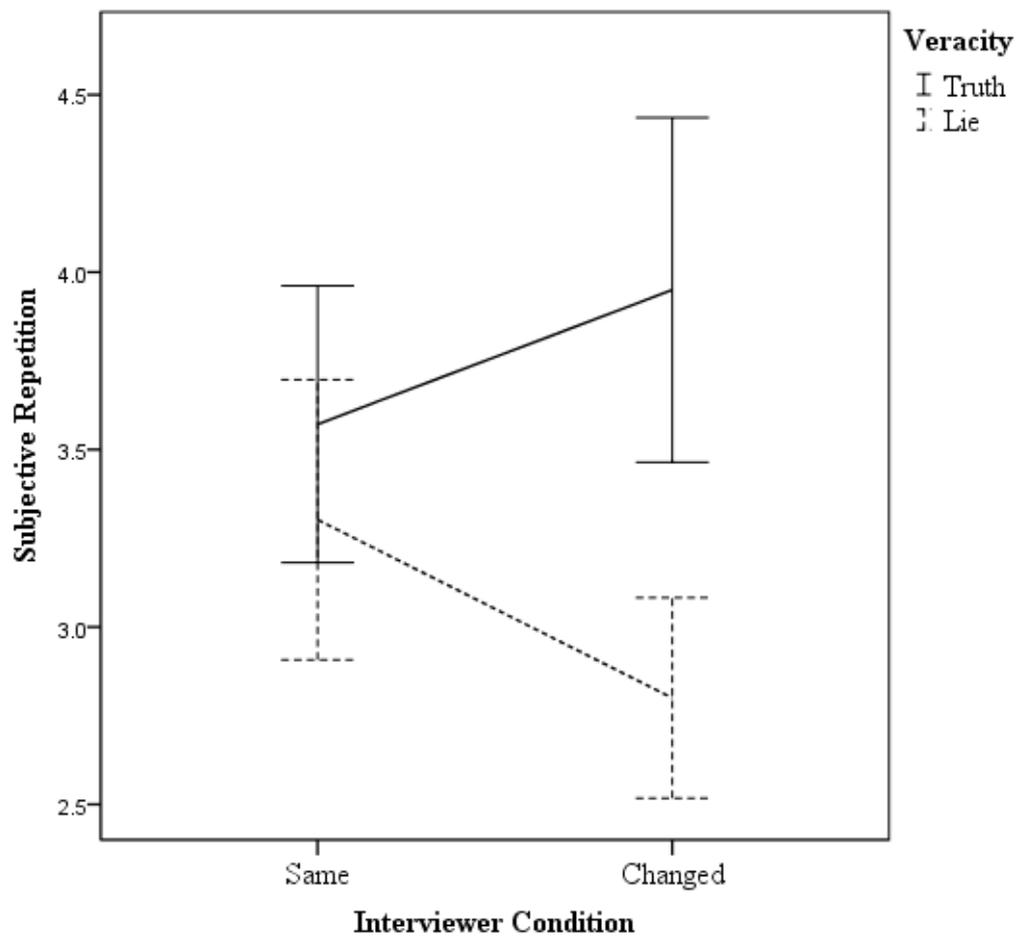
represent confident intervals (95%).

Figure 2. Subjective repetition between Stages One and Two as a Function of Veracity and Interviewer. Error bars represent confidence intervals (95%).

[1] The new information at Stage Two results could have been affected by the Stage One subjective details results. For example, if someone was very detailed at Stage One, new information at Stage Two is less likely to occur. We therefore introduced subjective details as Stage One as a covariate.

[2] An alternative analysis would be a 2 (Veracity) X 2 (Interview) X 2 (Time: Stage One versus Stage Two) analysis. The disadvantage of this analysis is that at Stage One the 'Same/Changed Interviewers' manipulation was not yet introduced, whereas the analysis treats the Stage One data as if this factor was introduced at Stage One.

[3] The subjective detail at Stage Two results could have been affected by the Stage One subjective details results. For example, those who reported many details at Stage One have a greater likelihood to obtain a higher repetition score on Stage Two. We therefore introduced subjective details as Stage One as a covariate.