



The European Journal of Psychology Applied to Legal Context

<https://journals.copmadrid.org/ejpalc>



The Effect of the Interviewer's Cognitive Load on the Quality of the Forensic Interview

Dora Giorgianni, Aldert Vrij, Sharon Leal, and Haneen Deeb

University of Portsmouth, UK

ARTICLE INFO

Article history:

Received 8 November 2024
Accepted 6 June 2025

Keywords:

Interview presentation modality
Appropriate questions
Inappropriate questions
Lie detection
Truth/lie bias

ABSTRACT

Background/Aim: The aim of this experiment was to investigate the impact of the interviewer's cognitive load resulting from the interview presentation modality on the interviewer's performance. **Method:** A total of 120 participants took on the role of the interviewer and were either exposed to a video with audio of an interview with a mock suspect (high cognitive load condition) or just the audio clip of that interview (low cognitive load condition). The suspects were either truthful or deceptive. The mock interviewers were asked to recall what the suspect said during the interview, propose follow-up questions, and determine whether the suspect was truthful or deceptive. **Results:** The interviewer's cognitive load did not affect recall, but participants in the interviewer's low cognitive load condition asked more questions and more high-quality questions and demonstrated better accuracy in determining whether the interviewees were truthful or deceptive than participants in the interviewer's high cognitive load condition. **Conclusions:** The results suggest that access to both vocal and visual cues (interviewer's high cognitive load) hampers the quality of the interview and veracity assessments.

El efecto de la carga cognitiva del entrevistador en la calidad de la entrevista forense

RESUMEN

Palabras clave:

Tipo de presentación de la entrevista
Preguntas adecuadas
Preguntas inadecuadas
Detección de mentiras
Sesgo verdad/mentira

Antecedentes/Objetivo: El objetivo del presente estudio fue investigar el impacto de la carga cognitiva del entrevistador producto del tipo de presentación de la entrevista en el desempeño del mismo. **Método:** Un total de 120 participantes actuaron como entrevistadores, que visionaron bien un video con audio de una entrevista con un falso sospechoso (condición de gran carga cognitiva) o, bien un mero audioclip de dicha entrevista (condición de carga cognitiva baja). Los sospechosos podían ser verdaderos o falsos. Se pidió a los entrevistadores simulados que recordaran lo que había dicho el sospechoso durante la entrevista, propusieran preguntas de seguimiento y y determinaran si el sospechoso estaba diciendo la verdad o mintiendo. **Resultados:** La carga cognitiva del entrevistador no influyó en el recuerdo, si bien los participantes de la condición de baja carga cognitiva del entrevistador hicieron más preguntas y de gran calidad y mostraron una mayor precisión a la hora de decidir si los entrevistados estaban diciendo la verdad o mintiendo que los participantes de la condición de carga cognitiva alta. **Conclusiones:** Los resultados indican que el acceso tanto a los indicadores vocales como a los visuales (carga cognitiva del entrevistador elevada) disminuye la calidad de la entrevista y la evaluación de la veracidad.

Research on the effect of increased cognitive load in interviews has mainly focused on the interviewee (Gombos, 2006; Vrij et al., 2011; Walczyk et al., 2013). However, cognitive load also affects the interviewer. Conducting an interview is mentally demanding, as the interviewer carries out multiple tasks simultaneously (Vrij & Fisher, 2019, 2020). The interviewer must listen to the interviewee's responses while thinking about the next question, and, at the same time, often watches the nonverbal behaviour of the interviewee to determine veracity. In our experiment, we sought to investigate the influence of the interviewers' cognitive load resulting from the presentation modality (audiovisual vs. audio) on their performance, particularly with respect to their recall of the interviewee's

statements, the quantity and quality of follow-up questions posed, and their ability to determine veracity.

Working Memory and Cognitive Load

Working memory refers to a brain system that provides temporary storage and manipulation of the information necessary to perform complex cognitive tasks. Working memory is thought to be divided into subcomponents: (1) the central executive, which is an attentional-controlling system, (2) the visuospatial sketch pad, which manipulates visual images, (3) the phonological loop, which

Cite this article as: Giorgianni, D., Vrij, A., Leal, S., & Deeb, H. (2025). The effect of the interviewer's cognitive load on the quality of the interview. *European Journal of Psychology Applied to Legal Context*, 7(2), 101-110. <https://doi.org/10.5093/ejpalc2025a9>

Correspondence: dora.giorgianni@port.ac.uk (D. Giorgianni).

ISSN: 1889-1861/© 2025 Colegio Oficial de la Psicología de Madrid. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

stores and rehearses speech-based information, and (4) the episodic buffer, a limited-capacity system for temporarily storing multimodal information (Baddeley, 1992; Baddley, 2000; Baddeley & Hitch, 1974). The existence of two separate systems for auditory and visual material implies that both components are activated when audiovisual information is perceived. Indeed, auditory stimuli significantly interfere with visual working memory and vice versa (He et al., 2022). Consequently, audiovisual presentations generally impose more cognitive load on individuals than single-modality presentations do (He et al., 2022).

Working memory supports the active maintenance of task-relevant information during the performance of a cognitive task. It can be overloaded by focusing attention on multiple simultaneous tasks, which causes a decrease in performance (Chabris & Simmons, 2011). An overload of cognitive capacity may undermine the reliability of decision making at all phases of the legal process, including during interrogations (Kleider-Offutt et al., 2016). As audiovisual presentations generally impose more cognitive load than auditory-only presentations, the cognitive load during investigative interviewing is higher if the interviewer listens to the interviewee and simultaneously watches their nonverbal behaviour (audiovisual presentation) than if they only listen to the speech (audio presentation). In other words, analysing the nonverbal behaviour while listening to the interviewee increases the cognitive load of the interviewer (Vrij & Fisher, 2019). Increased cognitive load can negatively impact the performance of the interviewer. For instance, research has shown that the ability to interpret nonverbal emotional expressions is negatively affected by working memory load (Phillips et al., 2008).

The Interviewer's Recall of the Interview

When presented with new information, working memory capacity is needed to process and transfer information to long-term memory effectively. If working memory is overloaded, less information is transferred to long-term memory, and recall is negatively affected (Camos & Portrat, 2015). We expect this also to happen in investigative interviews. Hanway et al. (2021) assessed the impact of the interviewers' cognitive load on their recall. All participants were instructed to watch and listen to a recorded interview. Participants in the high cognitive load condition were required to carefully comprehend what was said by the interviewee and think about follow-up questions; in the middle cognitive load condition, participants were only required to comprehend what was said; and in the control condition, no instructions were given. The results demonstrated that participants in the high and middle cognitive load conditions were less able to accurately recall what the suspect said than those in the control condition. In our experiment we manipulated the interviewer's cognitive load via the presentation modality (audiovisual vs. audio). The effect of modality presentation on recall was not examined before in an investigative interview experiment focusing on lie detection.

The Interviewer's Questions

We are not aware of research examining the impact of the interviewers' cognitive load on their ability to formulate questions during interviews (in Hanway et al., 2021 the follow-up questions the participants asked were not further analysed). Since working memory capacity is needed for complex tasks, it should also affect the ability to pose questions. Posing enough high-quality questions in interviews is important to gather as much relevant information as possible (Bull, 2010; Fisher, 2010; Walsh et al., 2025). Research has shown that different types of questions lead to different results: Appropriate questions are more effective than inappropriate questions in obtaining relevant accurate information (Oxburgh et al., 2012; Vrij et al., 2015). Appropriate questions are open-ended

questions framed in such a way that the interviewee can give an unrestricted answer (e.g., questions starting with "tell" or "describe") and probing questions (questions beginning with one of the WH words – "what?", "why?", "when?", "where?", "who?", "which?" – or beginning with "how"). Inappropriate questions include closed questions (yes/no questions), forced-choice questions (questions that only offer the interviewee a choice of possible responses), leading questions (questions that suggest an answer to the interviewee), multiple-at-once questions (questions that comprise a number of sub-questions asked all at once), opinion/statement questions (where an interviewer simply reads a statement or provides their own opinion and expects the interviewee to answer), and echo questions (closed questions that repeat what the interviewee has just said asking them to confirm) (Griffiths & Milne, 2006; Oxburgh et al., 2010; Vrij et al., 2015).

The Interviewer's Ability to Assess Veracity

Previous studies have compared truth/lie detection accuracy in various presentation modalities, such as visual, audio, audiovisual, and text-only. Bond and DePaulo's (2006) meta-analysis synthesised research results from 206 documents and 24,483 judges to analyse the accuracy of deception judgements. Overall, people achieved an average of 54% correct lie-truth judgements, correctly classifying 47% of lies as deceptive and 61% of truths as non-deceptive. A distinction in different modalities showed that the accuracy in distinguishing between truths and lies was lower in the visual (50%) than in the audio (54%) and audiovisual (54%) modalities. In other words, Bond and DePaulo (2006) found no differences between the audio and audiovisual conditions, the two modalities compared in the present experiment.

Differences between the audio and audiovisual modalities may occur in terms of bias rather than accuracy. People typically show a "truth bias", that is, they are more likely to judge a message as truthful than deceptive (Levine, 2014; Levine & Street, 2024; Street & Masip, 2015). However, the existence of a truth bias could depend on the presentation format. Nonverbal cues are generally thought to be cues to deceit, that is, cues that lie tellers are supposed to display more than truth tellers, such as gaze aversion and movements (Vrij et al., 2023). Access to visual information gives individuals the opportunity to focus on such cues to deceit. In contrast, verbal cues thought to be related to veracity are typically cues to truthfulness, that is, cues that truth tellers are supposed to display more than lie tellers, such as details and complications (Vrij et al., 2023). Access to auditory information gives individuals the opportunity to focus on such cues to truthfulness. Therefore, a "truth bias" is less likely to occur when judging visual messages. Indeed, Bond and DePaulo (2006) found that 63% of audio messages were judged as truthful, compared to 56% for audiovisual messages and 52% for video messages. In other words, messages were perceived as less truthful in the audiovisual condition than in the audio only condition.

Hypotheses

The aim of this experiment was to investigate the impact of cognitive load resulting from the presentation modality on the performance of the interviewer. The experiment was pre-registered (https://osf.io/x3nz4/?view_only=27df9614ca22426f8b3b8df5f-32badcc). We tested four hypotheses: Hypotheses 1, 2, and 3 were pre-registered and Hypothesis 4 is an exploratory hypothesis.

Pre-registered Hypotheses

Hypothesis 1: In the low cognitive load condition, participants will recall more correct details, and the number of questions, number of

appropriate questions and proportion of appropriate questions asked will be higher than in the high cognitive load condition¹.

Hypothesis 2: Participants will make more lie judgements in the high cognitive load condition than in the low cognitive load condition.

Hypothesis 3: In the high cognitive load condition, the lie detection accuracy rate will be higher, and the truth detection accuracy rate will be lower than in the low cognitive load condition.

Non-registered Exploratory Hypothesis

Hypothesis 4: In the high cognitive load condition, truth/lie detection accuracy will be higher when participants focus more on verbal cues and less on nonverbal cues. (This hypothesis only refers to the high cognitive load condition because participants in the low cognitive load condition had no access to visual cues.)

Method

Participants

A power analysis was conducted using G*Power to determine the necessary sample size. The analysis revealed that a sample size of approximately 120 participants was needed for 95% power, $\alpha = .05$ and a large effect size ($f = .40$), consistent with prior studies on the effect of cognitive load on the interviewer during investigative interviews (Hanway et al., 2021).

A total of 120 participants were recruited from the University of Portsmouth. The sample consisted of students and staff members. The sample included 36 males, 82 females, one non-binary person, and one person who preferred not to say. The average age of the participants was $M = 28.47$ years ($SD = 10.37$). A total of 36 participants had no university degree, 38 had a Bachelor's degree, 42 had a Master's degree, and four had a PhD. Recruiting students and staff members at the University ensured that, regardless of nationality, all participants were highly proficient English speakers. Each participant received £8 for their participation in the experiment.

Informed consent was obtained from all individual participants included in the study.

All procedures performed in this study involving human participants were in accordance with the ethical standards of the institution [University of Portsmouth Faculty of Science and Health Ethics Committee SHFEC 2021-131] and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Design

The experiment employed a 2 (Veracity: truth vs. lie) $\times$ 2 (Cognitive Load: high vs. low) between-subjects design. The 120 participants were randomly assigned to a Veracity and a Cognitive Load condition, with 30 participants in each cell. The dependent variables were the number of details recalled by the participants, the number of questions posed, the number of appropriate questions posed, the proportion of appropriate questions posed, truth/lie judgements and truth/lie detection accuracy.

Procedure

The Stimulus Materials

Before running the experiment, the required stimulus materials were prepared. Six volunteers were sent on a secret mock mission (the procedure was adapted from Vrij et al., 2021). They were told they had to imagine that they were secret agents for the government and that they had to carry out an important mission in the fight

against corrupted police officers. They were instructed to go to a specific shop and buy a specific item, and then go to a café to meet a fellow agent who could be identified through another specific item. They then showed the agent the purchased item and exchanged certain verbal codes to ensure that they knew that they were dealing with the correct person. Subsequently, the fellow agent gave the volunteers a package that they had to hide in a specific location making sure nobody was following them or looking at them while they were doing it.

The volunteers were interviewed after completing the mission. Three volunteers were instructed to tell the truth as follows: "You are going to be interviewed by a security official regarding the mission you have just been on. Please tell the interviewer everything you can remember about your mission in every possible detail. You need to convince the interviewer that you are telling the truth." The other three volunteers were instructed to mislead the interviewer about the mission, but at the same time to convince the interviewer that they were telling the truth: "You are going to be interviewed by a security official regarding the mission you have just been on. Unfortunately, we believe that the person who will interview you cannot be trusted as they have links to some corrupt officials. Therefore, you need to mislead the interviewer about the exchange. You need to report that you went on a mission, but to lie about 1) the location where you met the secret agent, 2) the identity and appearance of the agent (you must describe them differently from what they look like), 3) the content of the package you received, and 4) the place where you hid the package. However, you need to convince the interviewer that you are telling the truth. It is extremely important that the interviewer does not suspect that you are lying." The volunteers were then asked to report all they could remember in a single open-ended question (after their free recall no follow-up questions were asked).

The interviews were recorded and constituted the stimulus material for the experiment. Each of the six interviews was listened to/seen by twenty participants (half in the low cognitive load "listen" condition and half in the high cognitive load "watch with sound" condition). The six interviews consisted of three truth tellers and three lie tellers. The average length of the three lie tellers' accounts was 2 minutes and 16 seconds and the average length of the three truth tellers' accounts was 4 minutes and 23 seconds.

Recruitment

The participants, students and staff members from the University of Portsmouth, were recruited through an email invitation sent to a list of people interested in upcoming studies. Participants were informed that the study would be carried out remotely, that it would last around 30 minutes, and that it would imply playing the role of an interviewer.

Pre-interview questionnaire

The experiment was conducted online via Zoom. First, participants completed a pre-interview questionnaire. After reporting details about their age, gender and highest level of education, participants rated their level of motivation to perform well during the experiment on a 7-point scale from (1) *not motivated at all* to (7) *very motivated*; their level of stress on a 7-point scale from (1) *not at all stressed* to (7) *very stressed*; and their level of mental fatigue on a 7-point scale from (1) *exhausted* to (7) *very fresh*.

Audio Clip or Video

After completing the pre-interview questionnaire, participants either watched a video with sound (high cognitive load condition) or listened to an audio clip (low cognitive load condition) of an

interview with one of the six volunteers (see the stimulus materials section above).

Participants in the low cognitive load condition received the following instructions: "You will now listen to an audio clip, in which a person is interviewed about carrying out a mission. The person may be telling the truth or lying. Truth tellers report the entire mission truthfully whereas lie tellers do not. Four elements of the mission they discuss are of particular importance: 1) the location where they received a package, 2) the agent from whom they received the package, 3) the content of the package, and 4) where they dropped the package. You have to imagine that you are a police officer and that the interviewee is talking to you. After listening to the audio clip, you will be asked to tell everything you remember, propose some questions you would ask the interviewee and decide whether the interviewee is lying or telling the truth."

Participants in the high cognitive load condition received the following instructions: "You will now watch a video in which a person is interviewed about carrying out a mission. The person may be telling the truth or lying. Truth tellers report the entire mission truthfully whereas lie tellers do not. Four elements of the mission they discuss are of particular importance: 1) the location where they received a package, 2) the agent from whom they received the package, 3) the content of the package, and 4) where they dropped the package. You have to imagine that you are a police officer and that the interviewee is talking to you. Your task is to listen to the speech and observe the interviewee's behaviour. After watching the video, you will be asked to tell everything you remember, propose some questions you would ask the interviewee and decide whether the interviewee is lying or telling the truth."

The Interview

The participants then watched the video with sound or listened to the audio clip. They were not allowed to take any notes. After watching the video or listening to the audio clip participants were asked the following five questions which were written on slides presented to them to avoid any interviewer effect:

1) "Please recall the information provided in the audio clip/video as accurately and as completely as possible. Please recall every detail that comes to your mind."

2) "Now imagine that the interviewee is in front of you, and you have just listened to their statement. Please list any questions you would ask the interviewee. You can ask any questions that you think could help you gather the information needed to have a clearer picture and be able to decide whether the person was lying."

3) "To what extent do you think the interviewee was telling the truth or lying? (On a 7-point scale ranging from [1] *truth* to [7] *lie*)."

4) "If I now would force you to make a decision, do you think that the interviewee was telling the truth or lying? (On a dichotomous truth/lie scale.) The results were labelled truth detection accuracy (percentage of participants in the truth condition who classified the truth tellers correctly), lie detection accuracy (percentage of participants in the lie condition who classified the lie tellers correctly) and total accuracy (percentage of participants in the total sample who made a correct veracity judgement)."

5) "Finally, please tell me why you think the interviewee was telling the truth or lying."

Post-Interview Questionnaire

After the interview, participants were asked to complete a post-interview questionnaire. They were asked about their level of agreement on a 7-point scale from (1) *totally disagree* to (7) *totally agree* with the following items measuring cognitive load (the items were derived from [Vrij and Mann, 2006](#)): "I felt I was under pressure

during the experiment", "I felt that the experiment required a lot of thinking", "I felt that the experiment was mentally difficult", "I had to concentrate a lot during the experiment", and "I feel mentally tired now". We clustered these five variables (Cronbach's $\alpha = .81$) and called this clustered variable "experienced cognitive load".

Coding

All interviews were transcribed and coded.

Number of Details Recalled by the Participants

The participants' recall of the interviewee's statement was coded into PLATO [person, location, action, temporal, and object] details. PLATO coding scheme has recently been introduced in lie detection research ([Deeb et al., 2022](#)) with truth tellers typically providing more PLATO details than lie tellers ([Deeb et al., 2021](#)). Both correct and incorrect details were coded for each PLATO category. Any detail reported by the interviewee was coded only once, and repetitions were ignored. Person details involve mentions (e.g., pronouns and names) and physical descriptions of persons. For example, "The agent was wearing a blue T-shirt" includes three person details. Location details refer to directions (e.g., left, towards), positions (e.g., in front, upstairs), static places (e.g., building, room), and their descriptions (e.g., old building, dark room). For example, "The agent went left, reached the main square and then entered the bar" includes four location details. Action details include action verbs, such as walked, entered, and turned. For instance, "The agent went to the main square and then turned left and entered the bar" includes three action details. Temporal details denote any information that is time-relevant, such as then, afterwards, and before. For example, "In the morning, precisely at 9 the agent went on the mission" includes two temporal details. Object details refer to non-static objects, such as cars, chairs, computers, and their descriptions. For instance, "Outside the bar there were red chairs and black tables" includes four object details. In this experiment, we were interested in the number of total correct details recalled by the mock interviewer; therefore, we considered the sum of correct PLATO details. The number of reported incorrect details was low ($M = 2.79$, $SD = 3.59$) and did not differ between the experimental conditions. (See [Appendix](#) for the analyses).

The first author coded all the transcripts and the fourth author coded 60 (50%) transcripts. Inter-rater reliability was measured using the intra-class correlation (ICC) coefficient. The consensus is that inter-rater reliability is poor for ICC values less than .40, fair for values between .40 and .59, good for values between .60 and .74, and excellent for values between .75 and 1 ([Hallgren, 2012](#)). Inter-rater reliability between the coders, using the two-way random effects model measuring consistency, was excellent for all PLATO categories (correct total details single measures ICC = .96; correct person details single measures ICC = .96, correct location details single measures ICC = .84, correct action details single measures ICC = .90, correct temporal details single measures ICC = .84, and correct object details single measures ICC = .89).

Questions

The questions asked by the participants (in response to question 2 in the interview) were classified into appropriate and inappropriate questions. Appropriate questions were (i) open-ended (e.g., "Describe the scene"), (ii) probing (e.g., "What time did you meet the agent?"), and (iii) asking for verifiable details questions (e.g., "Can you show me the receipt?"). Inappropriate questions were (i) closed (e.g., "Were you alone?"), (ii) forced choice (e.g., "Were you walking or driving?"), (iii) leading (e.g., "Do you think your behaviour looks suspicious if you keep checking if

somebody is following you?”), and (iv) accusation questions (e.g., “How are we supposed to trust you?”). The first author coded all the questions asked by participants and the fourth author coded questions asked by 60 (50%) participants. Inter-rater reliability was excellent for both total questions (single measures ICC = .90) and appropriate questions (single measures ICC = .92).

Cues

The cues participants said they used to decide whether the person was lying or not (in response to question 5 in the interview), were clustered into separate categories. The initial classification resulted in 18 categories, but as very few participants observed some of them, we grouped them into larger categories. Finally, we had ten categories, of which six were visual, three were verbal, and one was vocal. The six visual cues were: body language (e.g., the agent was moving a lot), demeanour (e.g., the agent was relaxed), facial expressions (e.g., the agent was looking to the side), laughing (e.g., the agent was laughing), hands/arms movements (e.g., the agent was moving their hands while talking), and leg movements (e.g., the agent was moving their legs). The three verbal cues were style of presentation (including the presence of corrections, confusion/clarity, structured/unstructured production, and contradictions), details (i.e., the number of details), and plausibility (i.e., whether the story is plausible and makes sense). The vocal cue - labelled as “speech disturbances” - included pauses, silences, non-fluencies like stuttering, filler words like um, uh, etc.

The first author coded all the responses and determined the ten final categories. The fourth author coded the responses of 60 (50%) participants using the ten categories given by the first author. Inter-rater reliability was measured using the Cohen's kappa coefficient. The consensus is that values below 0 indicate no agreement, values between .00 and .20 indicate slight agreement, values between .21 and .40 indicate fair agreement, values between .41 and .60 indicate moderate agreement, values between .61 and .80 indicate substantial agreement, and values between .81 and 1.00 indicate almost perfect agreement (Landis & Koch, 1977). Agreement between coders was almost perfect for leg movements ($\kappa = 1$), body language ($\kappa = .92$), laughing ($\kappa = .91$), facial expressions ($\kappa = .90$), speech disturbances ($\kappa = .89$), hands/arms movements ($\kappa = .88$), and style of presentation ($\kappa = .87$); there was substantial agreement for demeanour ($\kappa = .76$), details ($\kappa = .79$), and plausibility ($\kappa = .63$).

Results

Questionnaire Variables Results

Pre-Interview Questionnaire

Three ANOVAs with a 2 (Cognitive Load) $\times$ 2 (Veracity) factorial design were conducted, with Cognitive Load and Veracity as between-subjects factors and the level of motivation, stress, and

mental fatigue as dependent variables. There was no statistically significant main effect of Cognitive Load (all F s < 1.75, all p s > 0.188) or Veracity (all F s < 0.53, all p s > .472). The interaction effect was not significant either (all F s < 0.38, all p s > .541). These results show that the mean scores of these three variables were similar among the experimental conditions. The participants' average level of motivation to perform well during the experiment was quite high ($M = 5.96$, $SD = 1.07$, 95% CI [5.76, 6.15]) and their level of stress was quite low ($M = 2.57$, $SD = 1.52$, 95% CI [2.29, 2.84]). Participants experienced average mental fatigue ($M = 4.88$, $SD = 1.33$, 95% CI [4.63, 5.12]).

Post-Interview Questionnaire

A 2 (Cognitive Load) $\times$ 2 (Veracity) ANOVA with experienced cognitive load as the dependent variable revealed no statistically significant main effects of Veracity, $F(1, 116) = 0.39$, $p = .535$, $d = 0.11$, 95% CI [-0.24, 0.47] nor of Cognitive Load, $F(1, 116) = 0.35$, $p = .555$, $d = -0.11$, 95% CI [-0.47, 0.25] and no statistically significant Cognitive Load $\times$ Veracity interaction effect, $F(1, 116) = 0.43$, $p = .515$, $\eta_p^2 = .004$. The participants reported to have experienced an average amount of cognitive load ($M = 3.77$, $SD = 1.16$, 95% CI [3.56, 3.98]).

Hypothesis Testing: Pre-registered Hypothesis 1

Number of Correct Details Recalled

A 2 (Cognitive Load) $\times$ 2 (Veracity) ANOVA with the number of correct details recalled as the dependent variable revealed a non-significant Cognitive Load $\times$ Veracity interaction effect, $F(1, 116) = 0.55$, $p = .460$, $\eta_p^2 = .005$. The main effect of Cognitive Load was not significant, see Table 1 for the Cognitive Load main effect results. The Veracity main effect was significant, $F(1, 116) = 18.53$, $p < .001$, $d = -0.79$, 95% CI [-1.17, -0.40]. Participants in the truth condition recalled more correct details ($M = 48.08$, $SD = 24.66$, 95% CI [43.04, 53.13]) than participants in the lie condition ($M = 32.58$, $SD = 12.71$, 95% CI [27.54, 37.63]). The significant Veracity effect was due to the length of the accounts, because the lie tellers' accounts (on average 291 words) were shorter than the truth tellers' accounts (on average 586 words). When the analysis was conducted with the percentage of correct details (number of correct details recalled divided by the total number of details included in the account) as the dependent variable, the main effect of Veracity was no longer significant, $F(1, 116) = 2.55$, $p = .113$, $d = 0.29$, 95% CI [-0.07, 0.65]. The Cognitive Load main effect (see Table 1) and the Cognitive Load $\times$ Veracity interaction effect, $F(1, 116) = 1.56$, $p = 0.214$, $\eta_p^2 = .013$, were also non-significant. Hypothesis 1, that predicted a Cognitive Load main effect, was therefore not supported for the number of correct details recalled.

Table 1. Univariate Main Effects of Cognitive Load

Dependent Variable	Low Cognitive Load		High Cognitive Load		F	p	d [95% CI]
	M (SD)	95% CI	M (SD)	95% CI			
Correct Details	41.25 (20.48)	36.21, 46.29	39.42 (21.70)	34.37, 44.46	0.26	.612	0.09 [-0.27, 0.44]
Percentage of Correct Details	40.45 (16.01)	36.32, 44.58	38.41 (16.59)	34.28, 42.54	0.48	.491	0.13 [-0.23, 0.48]
Number of Questions	6.33 (3.81)	5.37, 7.30	4.82 (3.71)	3.85, 5.78	4.82	.030	0.40 [0.04, 0.77]
Number of Appropriate Questions	4.93 (3.23)	4.14, 5.73	3.77 (2.94)	2.97, 4.56	4.24	.042	0.38 [0.01, 0.74]
Number of Inappropriate Questions	1.40 (1.66)	0.97, 1.83	1.05 (1.55)	0.65, 1.45	1.42	.235	0.22 [-0.14, 0.58]
Percentage of Appropriate Questions	77.22 (25.08)	70.72, 83.71	80.39 (25.32)	73.90, 86.89	0.47	.495	-0.13 [-0.48, 0.23]
Truth/Lie Rated Judgements	3.90 (1.59)	3.52, 4.28	4.60 (1.43)	4.22, 4.98	6.53	.012	-0.46 [-0.83, 0.09]

Table 2. Lie Judgements on a Truth/Lie Dichotomous Scale as a Function of Cognitive Load

Lie judgements	Low Cognitive Load	High Cognitive Load	<i>n</i>	χ^2	<i>p</i>	OR [95% CI]	ϕ
Total	58.33%	65.00%	120	0.56	.453	1.33 [0.63, 2.78]	.07
Truth condition	46.67%	80.00%	60	7.18	.007	4.57 [1.45, 14.39]	.35
Lie condition	70.00%	50.00%	60	2.50	.114	0.43 [0.15, 1.24]	.20

Number of Questions Asked

A 2 (Cognitive Load) $\times$ 2 (Veracity) ANOVA with the number of questions as the dependent variable revealed no significant main effect of Veracity, $F(1, 116) = 0.17$, $p = .682$, $d = -0.07$, 95% CI [-0.43, 0.28], but the Cognitive Load main effect was significant (see Table 1). Participants in the low cognitive load condition asked more questions than participants in the high cognitive load condition, supporting Hypothesis 1 for the number of questions asked. There was no significant Cognitive Load $\times$ Veracity interaction effect, $F(1, 116) = 0.26$, $p = .613$, $\eta_p^2 = .002$.

Number of Appropriate Questions Asked

A 2 (Cognitive Load) $\times$ 2 (Veracity) ANOVA with the number of appropriate questions as the dependent variable revealed a significant main effect of Cognitive Load. In support of Hypothesis 1, participants in the low cognitive load condition asked more appropriate questions than participants in the high cognitive load condition, see Table 1. The Veracity main effect, $F(1, 118) = 0.50$, $p = .482$, $d = -0.13$, 95% CI [-0.49, 0.23], and the Cognitive Load $\times$ Veracity interaction effect were not significant, $F(1, 116) = 0.00$, $p = 1.000$, $\eta_p^2 = .000$.

Proportion of Appropriate Questions Asked

A 2 (Cognitive Load) $\times$ 2 (Veracity) ANOVA with the proportion of appropriate questions asked (number of appropriate questions divided by the total number of questions) as the dependent variable revealed no significant main effect of Cognitive Load (see Table 1, failing to find support for Hypothesis 1) or of Veracity, $F(1, 116) = 0.07$, $p = .785$, $d = 0.05$, 95% CI [-0.31, 0.41]. The Cognitive Load $\times$ Veracity interaction effect was not significant either, $F(1, 116) = 0.05$, $p = .826$, $\eta_p^2 = .000$.

The significant Cognitive Load main effects for (i) total questions and (ii) appropriate questions asked but not for the proportion of appropriate questions suggests that participants in the low Cognitive Load condition also asked more inappropriate questions than participants in the high Cognitive Load condition. This was indeed the case, but the difference was not significant (see Table 1).

Hypothesis Testing: Pre-registered Hypothesis 2

Truth/Lie Rated Judgements

A 2 (Cognitive Load) $\times$ 2 (Veracity) ANOVA with truth/lie judgements on a 7-point scale as the dependent variable revealed a significant main effect of Cognitive Load. The target persons were seen as more deceptive in the high cognitive load condition than in the low cognitive load condition (see Table 1). This supports Hypothesis 2. The Veracity main effect, $F(1, 118) = 0.95$, $p = .332$, $d = 0.17$, 95% CI [-0.19, 0.53], and the Cognitive Load $\times$ Veracity interaction effect were not significant, $F(1, 118) = 2.90$, $p = .091$, $\eta_p^2 = .024$.

Lie Judgements on a Dichotomous Scale

A log-linear regression was conducted on truth/lie judgements in the forced-choice question with Cognitive Load, Veracity and the

Dichotomous Judgements as factors. Both bivariate interactions that included Cognitive Load were statistically significant (interaction with Veracity: $p = .019$, $B = 1.49$; interaction with Judgement: $p = .009$, $B = 1.52$). The interaction between Veracity, Cognitive Load and Judgement was also statistically significant ($p = .003$, $B = -2.37$). For the truth and lie conditions combined, there was no significant difference between lie judgements in the high cognitive load condition and in the low cognitive load condition. This does not support Hypothesis 2. However, in the truth condition, there were more lie judgements in the high cognitive load condition than in low cognitive load condition, whereas no difference occurred in the lie condition between lie judgements in the high cognitive load condition and in the low cognitive load condition. Data are presented in Table 2.

Hypothesis Testing: Pre-registered Hypothesis 3

Truth Detection Accuracy

A chi-square test was performed to compare the truth detection accuracy in the high and low cognitive load conditions. The truth detection accuracy rate was significantly lower in the high cognitive load condition (20.00%) than in the low cognitive load condition (53.33%), $\chi^2(1, N = 60) = 7.18$, $p = .007$. Hypothesis 3 was therefore supported for truth detection accuracy.

Lie Detection Accuracy

A chi-square test was carried out to compare lie detection accuracy in the high and low cognitive load conditions. The lie detection accuracy rates in the high cognitive load condition (50.00%) and the low cognitive load condition (70.00%) did not differ significantly from each other, $\chi^2(1, N = 60) = 2.50$, $p = .114$. Hypothesis 3 was therefore not supported for lie detection accuracy.

Total Detection Accuracy

A log-linear regression was conducted on total detection accuracy with Cognitive Load, Veracity and Accuracy as factors. The interaction between Veracity, Cognitive Load and Accuracy was not statistically significant ($p = .398$, $B = -0.67$). Only the bivariate interaction between Accuracy and Cognitive Load was statistically significant ($p = .009$, $B = 1.52$): The accuracy rate in the low cognitive load group (61.70%) was significantly higher than the accuracy rate in the high cognitive load group (35.00%).

Hypothesis Testing: Non-registered Exploratory Hypothesis 4

We tested Exploratory Hypothesis 4 in the high cognitive load condition only, carrying out a binomial logistic regression with accuracy as the dependent variable and all ten cues as independent variables. The results showed statistically significant positive relationships between total accuracy and (i) the verbal cue "details", and (ii) the vocal cue "speech disturbances" (see Table 3). No significant relationship was observed between total accuracy and the other cues, including all visual cues (see Table 3). Hypothesis 4 was partly supported as in the high cognitive load condition total

Table 3. Binomial Logistic Regression with Accuracy as the Dependent Variable and Visual, Verbal, and Vocal Cues as Independent Variables – High Cognitive Load Condition

Predictor	Estimate	SE	Z	p	OR [95% CI]
Body language	-0.33	0.82	-0.40	.691	0.72 [0.14, 3.60]
Demeanour	0.87	0.91	0.96	.338	2.39 [0.40, 14.19]
Facial expressions	-1.20	0.86	-1.40	.163	0.30 [0.06, 1.62]
Laughing	-0.28	1.06	-0.26	.792	0.76 [0.09, 6.07]
Hands/arms movements	-2.11	1.42	-1.48	.139	0.12 [0.01, 1.98]
Leg movements	3.46	2.22	1.56	.119	31.86 [0.41, 2475.63]
Plausibility	-1.17	1.27	-0.92	.357	0.31 [0.03, 3.75]
Details	2.20	0.89	2.48	.013	9.05 [1.59, 51.57]
Style of presentation	-0.00	0.83	-0.00	1.000	1.00 [0.20, 5.12]
Speech disturbances	1.99	1.02	1.96	.050	7.32 [1.00, 53.60]

detection accuracy was higher when participants focused on a verbal cue (details) and a vocal cue (speech disturbances) than on visual cues.

By leaving out the visual cues, we could analyse the relationship between cues and accuracy for all participants (including those allocated to the low cognitive load condition). We conducted three binomial logistic regressions with total accuracy as the dependent variable and the four verbal and vocal cues (style of presentation, details, plausibility and speech disturbances) as the independent variables. One regression was conducted on the total sample, a second one on the low cognitive load condition and a third one on the high cognitive load condition.

In the total sample, statistically significant positive relationships emerged between total accuracy and perceptions of (i) speech disturbances and (ii) the number of details reported by the interviewee. The other results were not statistically significant (see Table 4).

In the high cognitive load condition, a statistically significant positive relationship emerged between total accuracy and perception of the quantity of details reported by the interviewee. None of the other results were statistically significant (see Table 5).

In the low cognitive load condition, a statistically significant positive relationship emerged between total accuracy and the perception of speech disturbances. None of the other results were significant (see Table 6).

Discussion

Participants in the low cognitive load condition asked more questions in total and more appropriate questions than participants in the high cognitive load condition. The latter shows that high cognitive load impairs the quality of questions asked. The difference between high and low cognitive load in the experiment was that participants in the high cognitive load condition could see the target person whereas participants in the low cognitive load condition could not. Research has shown that the relationship between visual cues and deception is weak (see Vrij et al., 2019 for a review). Since visual cues are not indicative to deceit and having access to them impairs the quality of the interview, the findings thus suggest that it may be beneficial not to observe visual cues too closely in interviews.

The number of correct details recalled by participants in the low cognitive load condition did not differ significantly from the number of correct details recalled by participants in the high cognitive load condition. This contrasts with the findings of Hanway et al. (2021) who found that participants in the higher cognitive load conditions were less able to accurately recall what the interviewee said than those in the control condition. Perhaps the task in the present experiment was insufficiently challenging to display a difference, creating a ceiling effect.

Participants in the high cognitive load condition judged the accounts as more deceptive than participants in the low cognitive load condition. This aligns with the idea that messages are perceived

Table 4. Binomial Logistic Regression with Accuracy as the Dependent Variable and Verbal and Vocal Cues as Independent Variables – Total Sample

Predictor	Estimate	SE	Z	p	OR [95% CI]
Plausibility	0.12	0.53	0.23	.820	1.13 [0.40, 3.17]
Details	1.13	0.45	2.50	.012	3.09 [1.28, 7.45]
Style of presentation	-0.27	0.45	-0.61	.543	0.76 [0.32, 1.84]
Speech disturbances	1.46	0.46	3.18	.001	4.31 [1.75, 10.60]

Table 5. Binomial Logistic Regression with Accuracy as the Dependent Variable and Verbal and Vocal Cues as Independent Variables – High Cognitive Load Condition

Predictor	Estimate	SE	Z	p	OR [95% CI]
Plausibility	-0.98	1.17	-0.84	.402	0.37 [0.04, 3.73]
Details	1.78	0.68	2.61	.009	5.88 [1.55, 22.30]
Style of presentation	-0.15	0.75	-0.20	.841	0.86 [0.20, 3.74]
Speech disturbances	0.91	0.76	1.21	.228	2.50 [0.56, 11.04]

Table 6. Binomial Logistic Regression with Accuracy as the Dependent Variable and Verbal and Vocal Cues as Independent Variables – Low Cognitive Load Condition

Predictor	Estimate	SE	Z	p	OR [95% CI]
Plausibility	0.25	0.74	0.33	.739	1.28 [0.30, 5.51]
Details	-0.55	0.80	-0.69	.489	0.56 [0.12, 2.75]
Style of presentation	-0.78	0.61	-1.28	.199	0.46 [0.14, 1.51]
Speech disturbances	1.33	0.64	2.06	.039	3.77 [1.07, 13.32]

as less truthful when they contain visual information, which generally leads people to focus on cues to deceit (Bond & DePaulo, 2006; Vrij et al., 2019). In the low cognitive load condition, truth detection accuracy (53.3%) was significantly higher than that in the high cognitive load condition (20%). Perhaps due to the poor truth detection in the high cognitive load condition, participants in the low cognitive load condition achieved a higher total accuracy rate (61.7%) than participants in the high cognitive load condition (35.00%). These results support the idea that access to visual information hampers lie detection.

In the high cognitive load condition (where participants had access to verbal, vocal, and visual cues), total accuracy was positively correlated with paying attention to a verbal cue (the quantity of details included in an account), and a vocal cue (speech disturbances) but not with any visual cue. The analysis of the verbal and vocal cues in the total sample confirmed that accuracy was positively related to paying attention to the number of details the interviewee reported and to speech disturbances. The importance of paying attention to details in lie detection was recently also found by Verschuere et al. (2023). They presented nine studies that showed that when individuals relied solely on the richness of detail, they consistently were able to discriminate lies from truths at levels well above chance (59–79% accuracy).

Limitations and Future Research

In the post-interview questionnaire, we examined the participants' perceived cognitive load and found no effect of the presentation modality. A possible explanation for this finding is that the stimulus material was relatively short and that therefore the different presentation modalities have affected the overall cognitive load to a lesser extent than would have been the case in longer interviews. In addition, we measured perceived cognitive load via self-reports and perhaps they lacked construct validity as self-reports often do (Nisbett & Wilson, 1977). This would explain why, despite the unsuccessful manipulation according to the self-reports, participants in the low cognitive load condition nevertheless outperformed participants in the high cognitive load condition in several ways predicted in the hypotheses.

We found that having no access to visual cues increased the quality of the questions asked and the ability to distinguish between truth tellers and lie tellers. We consider these strong arguments for ignoring visual cues in interviews. However, there may be at least two disadvantages when ignoring visual cues in practice. First, since interviewers typically have access to audiovisual cues, they may feel uncomfortable if there is no access to visual cues. If interviewers feel uncomfortable their performance may suffer. Second, perhaps interviewers use visual cues to judge issues other than deceit in interviews, including the extent to which an interviewee feels at ease. It is important that interviewees feel at ease in interviews as that enables them to talk more (Vrij et al., 2017), and obtaining information is an important aim in interviews (Loftus, 2011). If interviewers misjudge how uncomfortable an interviewee feels, their reaction may be inadequate, and interviewees can 'shut down' as a result. Future research can examine whether these two issues hold true. A possible option to overcome these limits could be having two interviewers in two separate rooms. The main interviewer in the room with the suspect and the second interviewer in another room with the possibility to listen to the suspect and to talk to the main interviewer to suggest follow-up questions.

Conclusion

Participants in the low cognitive load condition asked more questions in total, more high-quality questions, and obtained higher accuracy in distinguishing between truth tellers and lie tellers.

Accuracy in distinguishing between truths and lies was positively correlated with observations of a verbal cue (the quantity of details reported by the interviewee) and a vocal cue (speech disturbances) and not related to observing visual cues. The results suggest that access to visual cues hampers the quality of the interview and veracity assessments.

Conflict of Interest

The authors of this article declare no conflict of interest.

Note

¹We added the number of appropriate questions to the pre-registered hypothesis because this reflects better the performance of the interviewer in terms of asking high quality questions.

References

- Baddeley, A. (1992). Working memory. *Science*, 255(5044), 556–559. <https://doi.org/10.1126/science.1736359>
- Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, 4(11), 417–423. [https://doi.org/10.1016/s1364-6613\(00\)01538-2](https://doi.org/10.1016/s1364-6613(00)01538-2)
- Baddeley, A. D., & Hitch, G. (1974). Working memory. *Psychology of Learning and Motivation*, 8, 47–89. [http://doi.org/10.1016/S0079-7421\(08\)60452-1](http://doi.org/10.1016/S0079-7421(08)60452-1)
- Bond, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review*, 10(3), 214–234. https://doi.org/10.1207/s15327957pspr1003_2
- Bull, R. (2010). The investigative interviewing of children and other vulnerable witnesses: Psychological research and working/professional practice. *Legal and Criminological Psychology*, 15(1), 5–24. <https://doi.org/10.1348/014466509X440160>
- Camos, V., & Portrat, S. (2015). The impact of cognitive load on delayed recall. *Psychonomic Bulletin & Review*, 22(4), 1029–1034. <https://doi.org/10.3758/s13423-014-0772-5>
- Chabris, C., & Simons, D. (2011). *The invisible gorilla: How our intuitions deceive us*. Harmony.
- Deeb, H., Vrij, A., Leal, S., Fallon, M., Mann, S., Luther, K., & Granhag, P. A. (2022). Mapping details to elicit information and cues to deceit: The effects of map richness. *European Journal of Psychology Applied to Legal Context*, 14(1), 11–19. <https://doi.org/10.5093/ejpalc2022a2>
- Deeb, H., Vrij, A., Leal, S., & Mann, S. (2021). Combining the Model Statement and the sketching while narrating interview techniques to elicit information and detect lies in multiple interviews. *Applied Cognitive Psychology*, 35(6), 1478–1491. <https://doi.org/10.1002/acp.3880>
- Fisher, R. P. (2010). Interviewing cooperative witnesses. *Legal and Criminological Psychology*, 15(1), 25–38. <https://doi.org/10.1348/135532509X441891>
- Gombos, V. A. (2006). The cognition of deception: The role of executive processes in producing lies. *Genetic, Social, and General Psychology Monographs*, 132(3), 197–214. <https://doi.org/10.3200/MONO.132.3.197-214>
- Griffiths, A., & Milne, R. (2006). Will it all end in tiers? Police interviews with suspects in Britain. In T. Williamson (Ed.), *Investigative interviewing: Rights, research, regulation* (pp. 167–189). Willan Publishing.
- Hallgren K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology*, 8(1), 23–34. <https://doi.org/10.20982/tqmp.08.1.p023>
- Hanway, P., Akehurst, L., Vernham, Z., & Hope, L. (2021). The effects of cognitive load during an investigative interviewing task on mock interviewers' recall of information. *Legal and Criminological Psychology*, 26(1), 25–41. <https://doi.org/10.1111/lcrp.12182>
- He, Y., Guo, Z., Wang, X., Sun, K., Lin, X., Wang, X., Li, F., Guo, Y., Feng, T., Zhang, J., Li, C., Tian, W., Liu, X., & Wu, S. (2022). Effects of audiovisual interactions on working memory task performance-interference or facilitation. *Brain Sciences*, 12(7), Article 886. <https://doi.org/10.3390/brainsci12070886>
- Kleider-Offutt, H. M., Clevinger, A. M., & Bond, A. D. (2016). Working memory and cognitive load in the legal system: Influences on police shooting decisions, interrogation and jury decisions. *Journal of Applied Research in Memory and Cognition*, 5(4), 426–433. <https://doi.org/10.1016/j.jarmac.2016.04.008>
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174. <https://doi.org/10.2307/2529310>
- Levine, T. R. (2014). Truth-Default Theory (TDT): A theory of human deception and deception detection. *Journal of Language and Social Psychology*, 33(4), 378–392. <https://doi.org/10.1177/0261927X14535916>

- Levine, T. R., & Street, C. N. (2024). Lie-truth judgments: Adaptive lie detector account and truth-default theory compared and contrasted. *Communication Theory*, 34(3), 143-153. <https://doi.org/10.1093/ct/qtac008>
- Loftus, E. F. (2011). Intelligence gathering post-9/11. *American Psychologist*, 66(6), 532-541. <https://doi.org/10.1037/a0024614>
- Nisbett, R. E., & Wilson, T. D. (1977). The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology*, 35(4), 250-256. <https://doi.org/10.1037/0022-3514.35.4.250>
- Oxburgh, G., Myklebust, T., & Grant, T. (2010). The question of question types in police interviews: A review of the literature from a psychological and linguistic perspective. *International Journal of Speech, Language and the Law*, 17(1), 45-66. <http://doi.org/10.1558/ijsl.v17i1.45>
- Oxburgh, G., Ost, J., & Cherryman, J. (2012). Police interviews with suspected child sex offenders: Does use of empathy and question type influence the amount of investigation relevant information obtained? *Psychology, Crime & Law*, 18(3), 259-273. <https://doi.org/10.1080/1068316X.2010.481624>
- Phillips, L. H., Channon, S., Tunstall, M., Hedenstrom, A., & Lyons, K. (2008). The role of working memory in decoding emotions. *Emotion*, 8(2), 184-191. <https://doi.org/10.1037/1528-3542.8.2.184>
- Street, C. N., & Masip, J. (2015). The source of the truth bias: Heuristic processing? *Scandinavian Journal of Psychology*, 56(3), 254-263. <https://doi.org/10.1111/sjop.12204>
- Verschuere, B., Lin, Ch.-Ch., Huismann, S., Kleinberg, B., Willemse, M., Jia Mei, E. Ch., van Goor, T., Löwy, L. H. S., Appiah, O. K., & Meijer, E. (2023). The use-the-best heuristic facilitates deception detection. *Natural Human Behaviour*, 7(5), 718-728. <https://doi.org/10.1038/s41562-023-01556-2>
- Vrij, A., & Fisher, R. P. (2019). Nonverbal cues to deception in Title IX investigations. *Journal of Applied Research in Memory and Cognition*, 8(4), 417-419. <https://doi.org/10.1016/j.jarmac.2019.07.006>
- Vrij, A., & Fisher, R. P. (2020). Unraveling the misconception about deception and nervous behaviour. *Frontiers in Psychology, section Personality and Social Psychology*, 11, Article 1377. <https://doi.org/10.3389/fpsyg.2020.01377>
- Vrij, A., Fisher, R. P., & Leal, S. (2023). How researchers can make verbal lie detection more attractive for practitioners. *Psychiatry, Psychology, & Law*, 30(3), 383-396. <https://doi.org/10.1080/13218719.2022.2035842>
- Vrij, A., Granhag, P. A., Mann, S., & Leal, S. (2011). Outsmarting the liars: Towards a cognitive lie detection approach. *Current Directions in Psychological Science*, 20(1), 28-32. <https://doi.org/10.1177/0963721410391245>
- Vrij, A., Hartwig, M., & Granhag, P. A. (2019). Reading lies: Nonverbal communication and deception. *Annual Review of Psychology*, 70, 295-337. <https://doi.org/10.1146/annurev-psych-010418-103135>
- Vrij, A., Leal, S., Mann, S., Vernham, Z., & Brankaert, F. (2015). Translating theory into practice: Evaluating a cognitive lie detection training workshop. *Journal of Applied Research in Memory and Cognition*, 4(2), 110-120. <https://doi.org/10.1016/j.jarmac.2015.02.002>
- Vrij, A., & Mann, S. (2006). Criteria-based content analysis: An empirical test of its underlying processes. *Psychology, Crime, & Law*, 12(4), 337-349. <https://doi.org/10.1080/10683160500129007>
- Vrij, A., Mann, S., Leal, S., & Fisher, R. P. (2021). Combining verbal veracity assessment techniques to distinguish truth tellers from lie tellers. *European Journal of Psychology Applied to Legal Context*, 13(1), 9-19. <https://doi.org/10.5093/ejpalc2021a2>
- Vrij, A., Meissner, C. A., Fisher, R. P., Kassin, S. M., Morgan III, A., & Kleinman, S. (2017). Psychological perspectives on interrogation. *Perspectives on Psychological Science*, 12(6), 927-955. <https://doi.org/10.1177/1745691617706515>
- Walczyk, J. J., Igou, F. P., Dixon, A. P., & Tcholakian, T. (2013). Advancing lie detection by inducing cognitive load on liars: A review of relevant theories and techniques guided by lessons from polygraph-based approaches. *Frontiers in Psychology*, 4, Article 14. <https://doi.org/10.3389/fpsyg.2013.00014>
- Walsh, D., Bull, R., & Areh, I. (2025). *Routledge international handbook of investigative interviewing and interrogation*. Routledge.

Appendix

Univariate Main Effect of Cognitive Load for Incorrect Details

Dependent Variable	Low Cognitive Load		High Cognitive Load		F	p	d [95% CI]
	M (SD)	95% CI	M (SD)	95% CI			
Incorrect details	2.45 (2.78)	1.73, 3.17	3.13 (4.24)	2.04, 4.23	1.10	.299	-0.19 [-0.55, 0.17]