

Original Study

Open Access

Marko Uotinen^{1,*}, Petteri Simola², Pentti J. Henttonen³

Detecting deceit within a predominantly true statement using two parallel assessment methods: A pilot study

DOI 10.2478/jms-2023-0005

Received: January 05, 2023; Accepted: September 14, 2023

Abstract: In human intelligence, a verbal statement from a source is seldom 100% true or false, and not very often is the source a total liar or a truth teller. From this standing point, a simple dichotomy of a liar or a truth teller might not offer an adequate diagnostic value for the purposes of human intelligence. A more diagnostic approach would be to assess which parts of the predominantly truthful verbal statement are likely to be true and which parts are assessed to be doubtful. In addition, the use of two parallel methods to detect deceit should improve the diagnostic value of the results. A pilot study in laboratory conditions ($n = 8$, yielding 190 assessment points) utilising an applied mock crime scenario was conducted. Correlation calculations showed that a dual-method approach slightly improved the within-statement truth accuracy, and it was achieved mainly by decreasing the number of false positives. As the truth accuracy was increased, the lie accuracy within the test group slightly decreased. The results confirmed that by applying parallel orienting response (EDA) and cognitive load (speech-related indices)-based assessment methods, it is possible to detect embedded lies successfully in an information-gathering interview setup.

Keywords: human intelligence, deception detection, embedded lies, within-statement lie detection, parallel methods, information-gathering interview

1 Introduction

Lie detection research often concentrates on exploring or improving a method with which the test participants can be accurately categorised into two groups, liars and truth tellers. This kind of dichotomist categorisation at the group level is well suited for law enforcement purposes as the questioning of suspects is aided or enhanced with a certain lie detection or veracity assessment method. Such a well-known method is the polygraph. Additionally, other methods have been developed based on different scientific foundations.

From the human intelligence collection point of view, it is not essential to find out if the source is innocent or guilty, or a total liar or truth teller. The starting point should be that the human source is potentially at least partially deceitful. For a motivated liar, a totally fabricated statement would not be a preferred option, if at least some related background information is anticipated to be available for the interviewer (DePaulo and Kashy 1998, p. 64). Not often lies are complete fabrications, although research setups are usually built in this way (Verigin et al. 2020, p. 369). If the source chooses not to fabricate the entire statement, generalisation, exclusion of details, events, persons and locations, low-stake white lies and lies regarding the subject's strategic interests can be expected to take place (Vrij and Granhag 2012, p. 114; George et al. 2014, p. 2). It has been found that interviewees often mix truths and lies in their statements (Palena et al. 2019, p. 2). In addition, as the source is interviewed, the investigator might not have any prior knowledge how the truth or deceit should go, and both statements could be very much alike in their basic content (Van Koppen 2012, p. 125).

From this starting point, the effort should be put into assessing the veracity of the statement and its episodic content to find out which parts of intelligence should be considered likely true and which parts are doubtful. Interviews should be conducted professionally, and the used assessment methods must be scientifically reliable to obtain accurate information (Vrij et al. 2014, p. 129).

*Corresponding author: Marko Uotinen, (Postgraduate Student), Department of Leadership and Military Pedagogy, National Defence University, PL7, 00861 Helsinki, Finland, E-mail: marko.uotinen@mil.fi
Petteri Simola, (PhD, psych.), Military Psychology Research Group, Human Performance Division, Finnish Defence Research Agency, Tuusula, Finland
Pentti J. Henttonen, (MA), Department of Psychology and Logopedics, University of Helsinki, Helsinki, Finland

The motivation for this kind of approach comes from the human intelligence collection-related practical requirements concerning the assessment of the collected information. The intelligence interview should produce information that is reliable and usable for future investigative purposes (Nunan et al. 2020, p. 512). It perhaps does not happen very often that security or intelligence officials are faced with a task to decide which members of a group should be considered totally truthful or totally deceiving. Even if this task would be relevant and could be completed successfully, the underlying question still prevails: Will these sources retain their status in the future?

During the collection of human intelligence, several phases of the overall process would potentially benefit from successful lie detection and the assessment of the veracity of the given statement. For example, the motives and goals of a walk-in agent, intelligence provided by a covert human intelligence source, intelligence received during prisoner-of-war interrogations or after-action debriefings of friendly troops could be aided by a reliable and applicable veracity assessment. Moreover, during the investigation of a suspected foreign agent or a spy taken into custody, veracity assessment and lie detection methods could be used. These kinds of procedural and technical aids would not only be beneficial for the various military intelligence collection or counterintelligence-related purposes but also for other law enforcement officials who run human intelligence-related operations and share the same kind of interests of knowledge in regard to lie detection and veracity assessment procedures.

Therefore, instead of concentrating on the successful lie detection at a group level, the limitations and possibilities of lie detection on an individual, within-statement level should be taken under scrutiny. Real-life-related requirements expressed by lie detection practitioners also highlight the need for this kind of approach (Vrij et al. 2022, p. 7). The ability to detect embedded lies should be endorsed as it has been recognised to be a common strategy for deceivers to increase their perceived overall credibility by lying as little as possible and by embedding their previous experiences within the statement as a deception (Verigin et al. 2020, p. 369).

1.1 Theoretical background of lie detection

Currently, three major theories are reported to explain the observable signs of deceit: emotional arousal, cognitive load and behavioural control. Variables that can be measured are physiological, verbal and non-verbal

(Zuckerman et al. 1981, pp. 7–10; Granhag and Vrij 2005, p. 52, 65; Hart et al. 2009, p. 135).

According to the emotional arousal theory, liars are more excited than truth tellers. This excitement and resulting emotions such as fear, guilt and delight induce noticeable changes in one's verbal and non-verbal behaviours. The cognitive theory is based on observations that lying is cognitively more demanding than telling the truth. The truth is the norm, and it is automatically activated. For a deceiver, it requires extra effort to suppress the truth and express something that is in contradiction to that. The lie must be plausible and consistent so that the liar does not get caught. In addition to monitoring own behaviour, a liar must also monitor the receiver to get feedback from his or her success. People have quite stereotypical views about how a sincere person behaves. Good liars try to act by these preconceptions and thus must control their behaviour accordingly. This leads to the attempted control theory that explains why liars fail to act naturally and leave a rigid impression while trying to suppress what they think resembles a guilty person's behaviour (Kirchhübel and Howard 2013, p. 695).

In addition to the emotional arousal theory, the preliminary process theory (PPT) and orienting response explain the changes in physiological reactions during the instrumental assessment process. Stimuli, in this case the questions, initiate cognitive processes, which produce variances in physiological responses. The greater the relative significance of the stimuli is, the larger the variance in the monitored parameters become. The veracity assessment is based on the monitoring and recording of the reactions by measuring the subjects' electrodermal activity (EDA, Palmatier and Rovner 2015, p. 1).

In lie detection research, such speech-related variables as the pause length, voice pitch, hesitation and the response length have been studied as potential indicators of deception (Hart et al. 2009, p. 135). Some of them have shown to be more potential than others. For example, studies regarding high-stake police interviews with real suspects showed that lying induced increased pauses (Vrij and Granhag 2012, p. 112). Additionally, an increased latency period, 'ah' speech disturbances and speech rate have been found to correlate with deception both in laboratory studies and in real-life situations (Vrij et al. 2000, p. 251, 254). Longer latency periods, increased pauses, hesitations, speech errors and slower speech rate have been found to indicate heightened cognitive load, which are interpreted as an attempt to deceive (Vrij et al. 2008, p. 255).

There are also studies that show less correlation between lying and speech hesitations, speech rate and

speech errors. In a study concerning the effectiveness of increasing cognitive load to facilitate lie detection, it was found that in the control group, the aforementioned criteria showed no significant discriminating power (Vrij et al. 2008, p. 256, 259). It is perhaps noteworthy that the participants in the control group had received a thorough coaching and very detailed information about the actual event that they lied about and thus could confidently reproduce their version of the event.

As the understanding of the human behaviour that is related to lying and lie detection accumulates, new theories explaining this vast phenomenon are created. Often orienting response, arousal, cognitive load and behavioural control theories can be traced back to establish the foundation that explains observable and measurable human behaviour supporting practical lie detection efforts (Walczyk et al. 2013, pp. 2-3; Vrij et al. 2019, pp. 298-300).

1.2 Simultaneous use of two lie detection methods

The simultaneous use of two or more lie detection methods has been recognised as a potential avenue of approach to lie detection (Granhag and Vrij 2005, p. 79). Theoretically, the parallel use of two or more methods should increase the overall accuracy by reducing the number of false positives, if two or more simultaneous indications based on different assessment methods are required as a confirmed result. In addition, two indications taking place at the same time should increase the perceived confidence in the finding, when no other supporting information is available.

Thus, the paradigm change towards being able to detect deceit within a single person's statement and the simultaneous use of multiple assessment methods lead to the following research problem: Will the use of two parallel methods improve the reliability of the lie detection within a predominantly truthful verbal statement?

For this study, the two selected lie detection methods used were orienting response and PPT-based measurement of EDA and cognitive load theory-based changes in speech production including 'ah' speech, speech errors and slowed speech rate.

There were two primary reasons for selecting these methods. Firstly, they are widely studied, and there is a fair amount of published research to support their reliability and applicability when they are tested at the group level.

Secondly, the essential requirement for a successful within-statement lie detection is the temporal fidelity

of the measured signs of deception in relation to the actual deception. The measured variables must take place, if not immediately, but at least relatively timely with the actual deception. The temporal fidelity between the observation and the occurrence of the deception is highly desired because the aim of this approach is to be able to identify suspected deceitful parts from the likely truthful parts of the statement. If the method, like the general perceived credible demeanour or the immediacy of the sender, would not meet the requirement of the temporal fidelity, the assessment would again be made at the group level, and the within-statement approach would not be tested.

As the task was to detect deceit within freely produced statement containing multiple episodic events, the question-answer latency was excluded as a cognitive load indicator. Illustrating hand gestures was also excluded because the EDA signal was measured from the interviewees' left hand, and they were instructed to hold them as motionless as possible to insure a proper contact. This was also assessed to hinder the interviewees' bodily movement on the whole.

At the selection of the used methods, it was assessed that the EDA measurement and the decoding of the speech could be carried out with the least amount of interference from the recording and the person performing the ratings.

Audio-video recording produced a wealth of data, which were not used in this study. Vocal cues, facial expressions, blinking rates, content for the statement validity assessment or reality monitoring are available to be used in future studies. The exclusion was based on the availability of time and other resources reserved for this study.

As the test setup was slightly different from the more common ones in lie detection studies, the applicability of the methods in this study had to be tested first. This was carried out by testing the following hypotheses:

- H1: Truth-tellers exhibit less EDA-derived indices of orienting response or nervousness than deceivers over the entire interview. The aim of this hypothesis is to validate the use of the method in the current test setup.
- H2: Truth-tellers exhibit less speech-related indices of cognitive load than deceivers over the entire interview. The aim is the same as in H1.
- H3: The combined use of two parallel lie detection methods increases the overall temporal accuracy of lie detection and improves the accuracy mainly by decreasing the number of false positives. The aim of

this hypothesis is to test the feasibility of the dual-method approach in the current test setup.

2 Method

2.1 Secret mission – the creation of the self-experienced event

Some arguments against the claimed effectiveness of the polygraph testing address the make-shift nature of the test setups in laboratory studies and the analysis of the results. The benefit of a laboratory study, often utilising mock crimes, is the knowledge of the ground truth. However, laboratory studies lack the heightened personal stakes of real-life situations that concern both guilty and innocent suspects, and thus, laboratory tests can be claimed not to represent the real potential of the polygraph test (Iacono 2008, p. 25). However, even if the real fear of getting caught and thus being penalised for lying was missing, the orienting response, or PPT, and cognitive load theory-based approach would overcome the possible issue of the participants not being mentally committed to lying. It can be argued that a potential reward does not work as effectively as a potential punishment, but a reward-based approach is assessed to produce measurable Orienting Response and Cognitive Load derived indices.

During an interview, reports from self-lived activities are assessed to be more plausible than the ones which were pre-scripted by a third party and afterwards reported as own experiences (Verigin et al. 2020, p. 380). To increase the personal involvement, possibly heighten the excitement factor and to increase the test persons' uncertainty concerning what the interviewer might know, the test persons were put to perform a secret mission consisting of a series of tasks. In principle, the tasks were simple, but they had to be performed successfully without getting exposed to anyone prior to the completion of the mission. The motivation to perform, and if needed to lie, was assessed to be high among all participants. In this test setup, a successful deception would be considered as a positive act, which could lower the perceived negative feelings towards lying (Halevy et al. 2014, p. 56).

Another purpose for the created secret mission type of scenario was to simulate a simple series of events that have at least some resemblances to actual human intelligence collection or counterintelligence operations. Even though the setup was following a typical mock crime scenario, in effect, the test persons playing secret agents were instructed to get into contact with an agent handler or an

assisting agent, receive further instructions from them, visit a dead letter box, receive additional instructions and not get caught in the middle of their covert activities. This kind of setup was assessed to be applicable with a secret human asset's after-action debriefing or a suspected foreign agent's investigative efforts.

As the research problem dictated, the content of the statement obtained during the intelligence interview had to be mostly, or at least partially, true. The overall duration of the secret mission was estimated to be from 30 to 60 min. This would result in enough content for the intelligence interview to be completed successfully. The desired duration of the intelligence interview was 10–15 min.

Eight participants, six males and two females between 21 years and 41 years ($M = 32,4$ standard deviation $[SD] = 6,0$), were randomly divided into test and control groups by the test assistant. They were treated equally in terms of providing instructions and the completion of the intelligence interview. In the morning of the test day, all participants in both groups were given instructions about their role in the study. They were advised that at any circumstances were they not to disclose any information regarding the given mission, even if directly asked. Both the test and control group participants were informed that they would be later interviewed about the test day's activities. If they would complete the given tasks and keep the secrets successfully, a reward of 20 euros would be granted.

The difference between the groups was that the test group was given a secret mission to be performed during the lunch break. The control group's mission was to have lunch without any additional directions or tasks. The test group's secret mission was identical for every participant. The mission was to perform a series of tasks and avoid getting caught or otherwise exposed. The initial task was to establish contact with a person in a coffee corner on the way to the lunch restaurant. The second task was received from the contact person. Instructions were written on a small piece of paper, and they ordered the test group participants to navigate to a given nearby location, an old telephone booth serving as a janitor's storage, retrieve a wallet and act according to the instructions found in the wallet. A note found in the wallet instructed to perform the third task, which was to take the money and return the wallet to the reception in the building lobby. Depending on the random selection, the wallet was either empty or there was money, which the test person should have taken if the instructions were followed. The participants should also keep and hide the instructions until the secret mission and the test was officially ended.

A research assistant was used to act as the contact person with whom the test group participants were to

meet at a coffee corner. Also, the lobby receptionist who monitored the access in and out of the building and to the decommissioned phone booth was assisting the test and received the wallets which were returned as lost and found. After the lunch break, the participants would independently leave for their respective activities until they would be called upon for the interview.

2.2 Interviewing protocol

Common to all veracity assessment and lie detection methods is that they are based on observations. It has been suggested that questioning techniques that produce longer answers should be preferred. These techniques potentially produce more information and thus more opportunities for verbal cues of deception to occur (Vrij and Granhag 2012, p. 115).

An identified risk concerning data collection was the potential shortness of the statement. It was anticipated that for the control group, a regular lunch break would not become an especially eventful period. As the control group participants would not have much to talk about, the test group participants were anticipated to keep the story short and simple.

Cognitive interview is a questioning method that aims to increase the statement's accuracy and the level of detail. Different techniques are applied to make the memory retrieval easier and more accurate. These techniques are the contextual reinstatement, repeating the statement without preliminary screening, recalling the events from another participant's point of view and recalling the events in mixed or reversed order (Memon and Higham 1999, pp. 178-184).

An enhanced version of the cognitive interview (ECI) was developed with the emphasis on the building rapport and effective communication. During the interview, the free flow of speech and information should not be interrupted, and the interviewee's effort should be put into active listening. Initially, the communication is facilitated with open-ended questions about neutral topics. The actual interview is started with contextual reinstatement and interviewee's free narrative of events. The interviewer then reminds that it is important to give a full account of events in as much detail as possible (Memon et al. 2010, p. 5).

The Strategic Use of Evidence (SUE) is a questioning technique that is based on a carefully planned presentation of critical and possibly incriminating evidence. The presentation of the facts, which the interviewer knows and has verified to be true, is supposed to produce different verbal responses from truth tellers than liars. It is

thought that the guilty suspect will use more avoidance strategies during the discussion than the innocent suspect who is thought to be more forthcoming. Avoidance strategies include, for example, leaving parts or details of the event untold. Truth tellers are generally more likely to tell the truth as it happened. On the tactical level, the SUE consists of three elements: evidence tactics, question tactics and disclosure tactics. Evidence tactics is a part of the preparation before the questioning. Question tactics is used to systematically exhaust the suspect's alternative explanations for the presented incriminating evidence. Disclosure tactics is used to increase the diagnostics value of the evidence by assessing the strength of the source and the degree of the precision to be able to present the evidence in a certain way. Research shows that the step-by-step presentation of stronger and more detailed evidence produces more and stronger cues to deception than presenting the most incriminating pieces of evidence straight away (Vrij and Granhag 2012, p. 114).

An applied version of the ECI with an SUE protocol type of incriminating evidence presentation at the end of the interview was created. Compared to the full version of the ECI, it was decided to exclude the contextual reinstatement and recalling the events from another participant's point of view because the interview took place at the same day, and it was assessed that all the related events were readily available for the accurate memory retrieval. Although the increase in the cognitive load has been shown to increase the ability to detect deception (Vrij et al. 2008, p. 262), recalling the events in a mixed or reversed order was excluded as it could offer an opportunity for a deceiving interviewee to spend time on irrelevant and possibly non-compromising topics. To promote an unhampered flow of information and an undisturbed occurrence of cues to deceit, no measures to add pressure during the interview was taken (Verschuere et al. 2016, p. 916). Although no extra effort to build rapport between the interviewer and the interviewee was taken, a neutral and cooperative atmosphere was sought after to promote an effortless flow of information (Nunan et al. 2020, p. 513).

All the interviews were conducted in Finnish, which was the native language for all the participants.

2.3 Interviewing questions

The free telling information-gathering interview was identical to all the participants. The interviewer did not know whether the interviewee was from the control group or from the test group. The interview consisted of a warmup period covering non-pertinent neutral topics.

WARMUP PERIOD	FREE TELLING PHASE - "Tell freely..." - "Once more, in more detail..." - "Include previously excluded..." - "Anything to add..."	PRESENTATION OF POTENTIALLY COMPROMISING LIGHT EVIDENCE	PRESENTATION OF POTENTIALLY COMPROMISING HEAVY EVIDENCE	ENDING
---------------	---	---	---	--------

Fig. 1: Phrasing of the used intelligence interview protocol.

The pertinent part of the interview consisted of three main phases: *free telling* initiated with open-ended questions concerning the lunch break, the presentation of *potentially compromising light evidence* and the presentation of *potentially compromising heavy evidence* (Figure 1).

The warmup period was intended to last from 10 to 15 min. Light topics concerning hobbies, pets and summer vacation activities were covered. The aim of the warmup period was to get the interviewee talking, ease up the atmosphere and form a calm starting point for the assessment during the actual information-gathering interview.

During the free telling phase, the following open-ended questions were used to obtain the statement concerning the lunch break: 'Please tell me freely, with your own words, what happened during your lunch break today?' Once the statement was finished, the interviewee was asked to elaborate more by asking: 'Please cover your lunch break in more detail, chronologically as it happened?' A third iteration of the statement was initiated by asking 'Please tell a bit more, and include details and events which you might have excluded earlier as insignificant?' The last question, which ended the free telling phase, was 'Once more, did you miss to mention anything that you would wish to add at this point?' This type of repetition was necessary for two reasons: Firstly, it was anticipated that the statements were not going to be very long. An hour-long lunch break would not likely offer much to talk about. Secondly, deceivers were assessed to keep their statements as short as possible as part of their strategy of not getting caught. All four questions were stated in this way independent of the length or the level of detail of the previous statements. A set of contingency questions was also available in case the interviewee would start to argue against the questions or decided to go numb. Another anticipated scenario was that the interviewee would try to drain the time reserved for the interview by going side tracked or by giving a super detailed statement. This would have

been encountered by asking the interviewee first to cover the main events in a more general level and after that go into details. In this study, all participants were cooperative and did not begin to resist the interview.

The presentation of the potentially compromising light evidence was completed with three questions. These questions were presented to all interviewees as they were regardless of their answers or reactions. The first question was 'During the lunch break you were seen meeting with a person at the coffee corner by the canteen, who did you meet with?' It was followed by 'Why did you choose to meet this person?' The third question ended this phase: 'What did you talk about?' If the interviewee denied having met anyone, the following optional question was asked: 'You were witnessed to meet with someone at the coffee corner, we have certain information about that, who did you meet with?' If the interviewee still bluntly denied the meeting taking place, the interview transitioned to the next phase.

Potentially compromising heavy evidence was presented with one question: 'You were reported to have returned a lost wallet at the reception, did you take any money from it?' If the interviewee denied or gave an ambiguous answer, the follow-up question was asked: 'Did you take money from the wallet, please answer yes or no?' (Figure 2).

2.4 Three types of embedded lies

The aim of the *secret mission* design and the use of an applied version of the ECI with an SUE-protocol type of incriminating evidence presentation was to create suitable conditions to deduce three types of lies. During the free telling phase, the concealment or omission of a detail, fact, event or a person, i.e. not telling the whole truth and that way manipulate the context of the actions, was expected. During the presentation of light evidence,

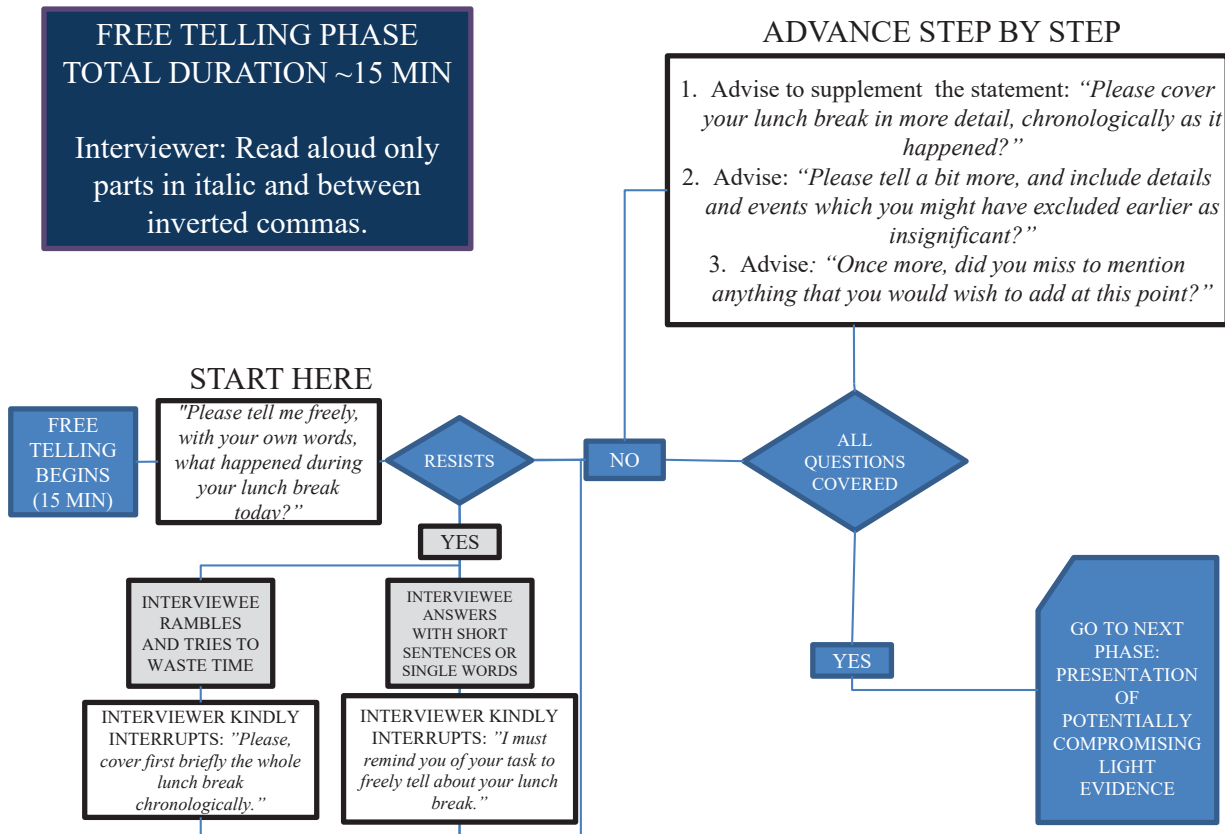


Fig. 2: Flowchart used by the interviewer during the free telling phase.

a quick fabrication of a plausible explanation, i.e. a replacement of truth with a small lie, was expected. If the deceitful interviewee was prepared, a credible cover story could have been made up and presented within the initial statement. Another option would be a complete denial of the compromising evidence. During the presentation of potentially compromising heavy evidence, a deliberate production of an untrue statement, i.e. a direct lie, was expected. Ideally, as the interview advances, the lies would develop from subtle lies to exaggerations and outright lies (Vrij et al. 2010, p. 9). All the anticipated lies would be embedded in their nature because they were expected to be told within the otherwise truthful statement (Vrij et al. 2022, p. 8).

2.5 Participants' level of preparation and coaching

All the participants were assessed to be motivated to perform well in the given tasks. They were all soldiers in reserve, non-commissioned officers and officers taking part in the refreshment training course. They did not have a unified military background or basic training. They had

not received any prior training related to lie detection, lie detection countermeasures or questioning resistance measures. The participation was voluntary, and it had to be conducted parallel to the refreshment training.

The participants in the test group were expected to be mentally prepared for the deception. They received the instructions in the morning, and they would complete the given tasks during the lunch break. The interview was told to be conducted on the same day. This arrangement would give enough time for the participants in the test group to come up with a deception strategy such as a cover story or a plausible explanation for their potentially unusual activities during the lunch.

Referring to the checklist by Walczyk et al. (2013, p. 7), truth-tellers in the control group were assessed to be motivated, they wanted to be perceived as believable and all the related details were easily available to be retrieved to working memory. In addition, the deceivers were expected to keep their statement internally and externally consistent, realistic in terms of time and events to get caught. It was also assessed that the planned deception was at least mentally rehearsed and that the deceivers were monitoring the interviewer for possible feedback concerning their credibility.

2.6 Data collection

The intelligence interview took place in a medium-size office with a window, a room enough to house three people and all the required equipment. The room had minimal furniture: three chairs for the interviewer, the interviewee and the technical assistant; a small table; non-intrusive auxiliary light; a video camera; and a NEXUS-10 biosignal recorder manufactured by Mind Media B.V., Netherlands. The interview was recorded, both video and audio. The equipment was assessed not to be intrusive, although their presence and purpose were neither concealed in any way nor was their function specifically explained. The general atmosphere was business-like neutral, and no deliberate efforts to increase mental pressure were taken.

Data collection produced 32 samples per second EDA time series along with a High Definition (HD) audio-video recording. The EDA measurement formed a master track to which the audio recording was synchronised using manual triggers. Only the audio track of the recording was used in this study.

The interview was the final part of the participants' secret mission. Before the test persons were allowed to leave, their activities during the test were double-checked from the assistants who were acting as the contact person and the receptionist. Also, the sequence of events was double-checked by asking the participants to state them in the chronological order. After that, their participation to the test was confirmed to be successful, and their rewards were paid.

2.7 Data analysis

The question–answer turns for each interview were coded. Every participant's statement was divided into 15-s slots. Each slot was analysed in terms of cognitive load indicators and nervousness or orienting response indicators. The slot was flagged as a positive indication, if cognitive load speech indicators, including 'ah' speech, speech errors and slowed speech rate, occurred. The assessment was based on the verbatim interview transcripts. The speech rate, words per minute, baseline was established within the pertinent part of the interview (Vrij 2016, p. 1114). Individual speech rates were calculated from the transcripts over 5-s intervals, and if the measured speech rate dropped under the set threshold, the corresponding 15-s slot was flagged as a positive indication of deceit. The threshold for slow speech was the participant-specific words per minute average subtracted with the SD calculated over the pertinent part of the interview.

All participants' measured EDA time series was pre-processed and decomposed into tonic (slow) and phasic (fast) components using Ledalab software (Benedek and Kaernbach 2010, p. 82). The last 3 min of the non-pertinent part of the interview was taken as a reference point, or the threshold, of a non-deceiving statement, for the pertinent part of the interview. Phasic EDA, which had occurred during the last 3 min of the warmup period, formed an envelope of normal variation. Phasic EDA above that individual threshold was considered as an indication of deceit during the interview. The deception was assessed to have ended once the measured EDA had come down to the half of the peak value (Braithwaite et al. 2015, p. 6–7). All the 15-s slots containing deception indicating EDA markers were flagged as positive indications.

Blind to the actual status of the participants, both the EDA and speech-related indicator data were timely synchronised. If an EDA and a speech indicator occurred at the same time, the corresponding 15-s slot for the combined EDA and speech was flagged as a positive indication of deceit. This process produced three columns of indicator data for each participant: independent EDA, independent speech and combined EDA and speech. All these three data sets were compared with the actual deceit during the intelligence interview.

As described earlier, every participant's statement was broken down to 15-s slots and analysed accordingly. This resulted in a total of 190 assessment points, which were labelled as truths or lies, which was indicated by the studied veracity assessment methods. The correlation calculations were conducted over those 190 assessment points. Instead of having just eight test participants, who would traditionally be assessed as liars or truth-tellers, the used method multiplied this study's effective sample by over 20.

3 Results

Based on the test design, three types of lies were expected: concealments, fabrications and direct lies. All three types of lies were present in the three deceiving test persons' statements. In total, the test setup yielded 41 lies or attempted deceptions created by the test group participants. The most common form of deception was a fabrication, which was used 23 times. Two deceivers adopted a cover story, which resulted in a high number of fabrications. A concealment, skipping an event or excluding a related person, was used 11 times. A direct lie was told seven times. Judging by the content of their answers, the truth-tellers reported their activities during

the lunch break as they happened. In these terms, the test setup was successful.

The research problem was to find out if the use of two parallel methods improves the reliability of the lie detection within a predominantly truthful verbal statement. Before getting answers to the research problem, the hypotheses were tested.

As predicted in H1, a *t*-test confirmed that the truth-tellers scored significantly lower in EDA ratings ($M = 0.76$, $SD = 1.36$) than the deceivers ($M = 1.80$, $SD = 1.26$; $t(67) = -3.11$, $p < 0.01$) throughout the whole interview. The deceivers exhibited more signs of nervousness or orienting response in every phase of the interview, and the total number of indicated deception was higher among the deceivers than among the truth-tellers.

H2 predicted that the number of speech-related cognitive load indicators would be lower among the truth-tellers than among the deceivers. A *t*-test did not confirm this hypothesis. The truth-tellers did not score significantly lower in speech ratings ($M = 1.69$, $SD = 2.52$) than the deceivers ($M = 2.36$, $SD = 2.97$; $t(67) = -0.98$, $p = 0.16$).

To see if the cognitive load differed between the two groups during the presentation of potentially compromising light and heavy evidence, H2 was tested without the data from the free telling phase of the interview. A *t*-test confirmed that during the presentation of potentially compromising light and heavy evidence, the truth-tellers scored lower ($M = 0.23$, $SD = 0.42$) in speech indicator ratings than the deceivers ($M = 0.85$, $SD = 1.14$; $t(35) = -2.30$, $p < 0.05$).

The conclusion from these findings is that the free telling phase of the interview was cognitively tasking for both groups, although the reasons might have differed.

Perhaps the confrontation with potentially compromising light and heavy evidence required the deceivers to concentrate on their cover story or plausible denials as the truth-tellers basically bluntly denied or explained that there must be a mistake.

Regarding H3, it is first emphasised that the analysis entails only the free telling part of the interview. The two latter parts of the interview, where the participants were confronted with light or heavy compromising evidence, were excluded because the dual-method approach needed to be tested against narrative statements instead of simple yes and no answers or blunt denials.

H3 stated that the parallel use of two methods, which are based on different scientific approaches to lie detection, would improve the lie detection accuracy, especially by decreasing the number of false alarms within all participants.

To see if the within-statement assessment method would prove to be successful in detecting deception among the mostly true statement, a two-tailed Spearman's correlation coefficient between the actual deception and individual EDA, individual speech and combined EDA and speech indicators was calculated. The results are shown in Table 1.

The data show that among the truth-tellers, the use of combined EDA and speech indicators cancelled the general poor performance of the speech indicators and increased the correlation by decreasing the number of false positives when only EDA indicators were applied. Among the deceivers, the use of combined indicators did not increase the correlation. With two deceivers, the speech indicators yielded low correlation results,

Tab. 1: Correlation Between the Assessment and the Actual Status of the Statement Using EDA Indicators, Speech Indicators and Combined EDA & Speech Indicators.

	EDA			Speech			EDA & Speech		
	<i>rs</i>	<i>p</i> *	<i>N</i>	<i>rs</i>	<i>p</i> *	<i>N</i>	<i>rs</i>	<i>p</i> *	<i>N</i>
Truth-tellers									
Test Subject 3	0.27	0.131	33	0.21	0.250	33	0.48	0.005	33
Test Subject 5	1.00	0.000	13	-1.00	0.000	13	1.00	0.000	13
Test Subject 7	0.34	0.163	18	0.24	0.330	18	0.54	0.020	18
Test Subject 9	0.29	0.096	34	0.13	0.469	34	0.31	0.070	34
Test Subject 10	0.67	0.016	12	0.21	0.516	12	0.67	0.016	12
	<i>rs</i>	<i>p</i> *	<i>N</i>	<i>rs</i>	<i>p</i> *	<i>N</i>	<i>rs</i>	<i>p</i> *	<i>N</i>
Deceivers									
Test Subject 4	0.56	0.006	23	0.17	0.435	23	0.46	0.026	23
Test Subject 6	0.49	0.002	37	0.49	0.002	37	0.40	0.014	37
Test Subject 8	0.62	0.004	20	0.15	0.518	20	0.49	0.027	20

*2-tailed

EDA, electrodermal activity.

which also lowers the correlation results of the combined method, when compared to the use of EDA indicators only. This is explained by the high number of ‘ah’ speech indications, which resulted in a high number of false positives. In this study, the use of ‘ah’ speech was not found to support the accurate lie detection. However, the use of combined EDA and speech indicators showed moderate to strong correlations among four out of five truth-tellers and all three deceivers with statistical significance, which confirmed the general applicability of the used method (see Table 1).

The accuracy of the three different approaches was calculated during the free telling part of the interview. The combined use of two parallel lie detection methods increased the truth accuracy during the free telling phase in both groups. As predicted in H3, the number of the false positives decreased compared to the use of EDA or speech indicators only (see Table 2). For all the eight participants, the truth accuracy increased or stayed the same as the combined EDA and speech indicators were applied. Especially the reduction of speech-related false positives was clearly notable.

The predicted increase in truth accuracy came at the expense of the deceivers’ deceit accuracy. Although the deceivers’ EDA and speech indicator overall accuracy results ranging from 67% to 78% were at the expected numbers (Granhag and Vrij 2005, p. 56), the deceit accuracy was lower than the single use of either EDA or speech indicators. The highest overall accuracy ranging from 73% to 82% was achieved by applying the EDA indicators only. This again highlights the potential issue of including the ‘ah’ speech as an indicator in the detection as they seem to have been very common to all participants, regardless of their status in the study.

The accuracy of the three different approaches during the presentation of potentially compromising light evidence was calculated as a single group using the data from all the participants. The total number of presented light evidence-related questions for all the participants was 22. Thirteen answers were truthful, and nine were deceiving. The application of EDA indicators resulted in 79% overall accuracy (69% truth accuracy, 89% deceit accuracy). The application of speech indicators resulted in 64% overall accuracy (61% truth accuracy, 67% deceit accuracy). The use of EDA and speech indicators resulted in 70% overall accuracy (85% truth accuracy, 55% deceit accuracy). Again, the EDA and speech increased the truth accuracy at the expense of the deceit accuracy. The highest overall accuracy during the confrontation with light evidence was achieved by applying EDA indicators only.

When confronting the test persons with potentially compromising heavy evidence, the anticipated answer was either yes or no. For this reason, only the EDA indicator data were included, although some of the participants wanted to offer longer explanations. The results were calculated using data from both groups. The total number of presented heavy evidence-related questions for all the participants was 12. Nine answers were truthful, and three were deceiving. The application of individual EDA indicators resulted in 83% overall accuracy (67% truth accuracy, 100% deceit accuracy). The number of deceiving answers was only three, which slightly makes this result anecdotal.

The overall accuracy results from the first two phases of the interview confirmed that the parallel use of two lie detection methods improved the overall accuracy mainly by decreasing the number of false positives, as stated in H3.

Tab. 2: Accuracy Rates for Deception and Truth Using EDA Indicators, Speech Indicators and Combined EDA & Speech Indicators.

	Truth			Deceit			Overall		
	EDA	Speech	EDA & Speech	EDA	Speech	EDA & Speech	EDA	Speech	EDA & Speech
Truth-tellers									
Test Subject 3	70%	58%	88%	na	na	na	na	na	na
Test Subject 5	100%	0%	100%	na	na	na	na	na	na
Test Subject 7	67%	50%	83%	na	na	na	na	na	na
Test Subject 9	92%	33%	92%	na	na	na	na	na	na
Test Subject 10	74%	35%	76%	na	na	na	na	na	na
	EDA	Speech	EDA & Speech	EDA	Speech	EDA & Speech	EDA	Speech	EDA & Speech
Deceivers									
Test Subject 4	92%	25%	92%	73%	91%	64%	82%	58%	78%
Test Subject 6	88%	76%	92%	58%	75%	42%	73%	76%	67%
Test Subject 8	73%	36%	82%	89%	78%	67%	81%	57%	74%

EDA, electrodermal activity.

4 Discussion

The correlation between the assessment and the actual status of the statement was mainly moderate within the test and the control group when the dual-method approach was applied. The results showed that during a free telling information-gathering interview, the parallel use of orienting response and cognitive load theory-based approaches slightly increased the within-statement truth accuracy compared to the use single use of orienting response indicators. Compared to the cognitive load indicators, the truth accuracy increase was more significant. However, the used dual-method approach did not increase the lie accuracy, and both individual methods performed better.

Additionally, when the test persons were confronted with potentially compromising evidence, the dual-method approach proved to be successful with some limitations. Generally, the test persons' responses were short or single-word answers, which posed a limitation to the use of speech indicators. However, the overall accuracy was at a promising level.

The comparison between the EDA indicator and speech indicator results during the free telling phase showed that EDA indicators returned higher correlation with all test persons. Also, the truth accuracy was higher within all the test persons. Interestingly, within the deceiving test persons, the average speech indicator deceit accuracy was higher than the EDA indicator deceit accuracy. This result is mainly explained by the used speech indicator markers, especially 'ah' speech disturbances, which occurred very frequently in all the test persons' statements, and thus is more likely to co-occur with the deceit. This also partially explains a low speech indicator truth accuracy average within all the test persons.

As a laboratory study, this setup represents a low-stake situation. It could be argued that in a high-stake real-life situation, the motivation of not getting caught lying, or the fear of getting unjustly accused as a liar, would be significantly higher (Granhag and Vrij 2005, p. 74). On the other hand, it could be argued that a well-prepared and experienced deceiving interviewee would be seemingly calm and feel relaxed at the information gathering-oriented intelligence interview, which largely resembles the used test setup. In addition, an extensive meta-analysis shows no difference in the lie detection accuracy between high-stake and low-stake situations (Verschuere et al. 2016, p. 917).

By the design of the secret mission, both the test group and the control group participants had all the event-related information readily available. All the control

group participants exhibited signs of orienting response and cognitive load. Since the control group participants were not deceiving during the interview, the existence of measured signs of deception can be argued to be caused by general excitement, efforts to think hard while answering or some other personal reasons which were left unidentified.

The test group participants were all consistently deceiving during the free telling part of the interview when the repetition of the statement was requested. Their statements were based on self-experienced real-life events which were readily available. Their lies were embedded among the true parts of the statement, and these lies were developed by the participants themselves, not fabricated by someone else. From this standing point, the test was largely measuring lies and deception, as was the intention of the test setup, but it must be also recognised that other sources for the measured signs of deception can exist.

4.1 Limitations

Some limitations must be addressed. The sample size is small ($n = 8$), of which a test group of three participants and a control group of five participants was established.

However, as the measurement and the analysis were conducted on an individual, within-statement level, the number of the assessment points was much higher than the number of the participants (Levine et al. 2022, p. 191). In a dichotomist approach, one participant constitutes one sample as the choice is between a truth teller or a deceiver. In this study, each statement was sampled in 15-s slots, where measurement and data analysis were applied. Thus, in this setup, eight participants constituted a total of 190 assessment points, which were analysed during the free telling phase. This resulted in an average of 23 assessment points per participant.

The correlation calculations did not interfere with each other as Spearman's correlation coefficient was calculated individually for every test person. Similarly, accuracy rates were calculated on an individual level over within-statement assessment points. This type of measurement and analysis method is assessed to increase the reliability of the results and to decrease the effect of the small number of the test participants to the overall validity of the results.

Setting the threshold of the phasic EDA on a level that occurred during the last 3 min of non-pertinent part of the interview was based on an assumption that the test persons were not deceiving at that time. In this test setup, the assumption was assessed to be relatively accurate and

reliable, but it can be argued to become a significant factor of unreliability outside of the laboratory environment.

As mentioned before, the prevalence of frequent ‘ah’ speech disturbances might be related to the interviewees’ native tongue or some other personal features, which were not known. This phenomenon might have saturated the cognitive load indicators or at least made them less indicative in this study. In addition, in the Finnish language, the sources, meaning and purpose of different interferences are likely to be unique, or their relevance regarding lie detection is unknown. How lying affects the production of Finnish language is not widely studied. To improve the diagnostic value of Finnish language disfluencies to facilitate lie detection, more extensive studies are required. However, disfluencies observed in this study’s interviews follow the general guidelines found in the Finnish population (Penttilä et al. 2018, p. 159, 161). The next challenge is to find out how these data could be operationalised to aid verbal lie detection.

The application of this type of intelligence interview and veracity assessment setup requires the interviewee to be cooperative. Firstly, the interviewee must be willing and able to reply to the open-ended questions and produce relatively long narrative answers. The interviewee must also be willing to repeat the statement as requested and not to suddenly get frustrated or anxious about the repetition. In addition, the interviewee must be willing to get wired to a polygraph or an EDA recorder and thus be able to control his or her bodily movement throughout the interview.

5 Conclusion

The aim of this study was to test the feasibility of the orienting response and cognitive load-based lie detection methods in identifying the deceiving part of the statement within a predominantly truthful statement. The results show that by applying concurrent orienting response (EDA) and cognitive load (speech-related indices)-based assessment methods, it is possible to detect embedded lies successfully, which can be indicated by the correlation calculations.

The used dual-method approach slightly improved the truth accuracy in both the test and the control group but, in return, worsened the deceit accuracy in the test group. However, despite the worsening effect on the deceit detection, the overall accuracy remained beyond the level of chance in the test group, too.

For the tested method to be applicable, the interviewee needs to be cooperative, answer to the questions and not to resist by being numb, resorting to not to comment or

only agreeing to talk about irrelevant topics. Why would then the interviewee agree to answer with the best of his or her knowledge? With the truth-tellers, the rationale might be simpler. Their motive could be argued to be to present themselves as straightforward, cooperative and trustworthy. A truthful person might also have doubts about the interview and the interviewer, and he or she might be afraid of not to be taken seriously. This could be observed as nervousness, excitement or reservedness. On the other hand, the deceiver’s motive could be argued to be able to avoid being seen as suspicious and unreliable. Suspicions may arise if the deceiver gives away compromising pieces of information, the statement appears to be inconsistent or implausible or he or she refuses to cooperate or, in some other way, behaves in contradiction to the situation-related expectations. From this standpoint, if both the truth-tellers and the deceivers were motivated to succeed, the tested method could be argued to be applicable in detecting deceit from cooperative interviewees.

From the human intelligence collection point of view, these results are promising. High truth accuracy is desirable in two ways: Firstly, recognised pieces of valuable intelligence can be exploited with moderate to high reliability, and secondly, future intelligence collection and lie detection efforts can be directed effectively towards doubtful elements or topics of the statement. High truth accuracy also means that only little valuable intelligence, which the human source provides, is discarded as lies during the process. Evidently with room for improvement, the currently demonstrated moderate deceit accuracy is high enough to set guidelines for the future investigative efforts accepting the fact that few lies, smaller or bigger, are left undetected.

So far, 100% accurate lie detection or veracity assessment methods are yet to be discovered, and some level of uncertainty, whether it is methodological or statistical, must be accepted. A within-statement approach offers benefits over dichotomist lie detection approaches.

From the human intelligence collection point of view, be it military intelligence or other security agency-related, it is of value to be able to assess the validity of collected information. A dishonest human source with a bad reputation of telling many lies might still give away or even voluntarily offer valuable information, regardless of his or her previous history as a liar. This makes the within-statement assessment approach more usable than dichotomist approaches. The method itself enables practical application and the exploitation of the retrieved information, even if 30%–20% of results are incorrect. In total, 70%–80% overall accuracy means that most of the valuable truthful information can be exploited and only some useful information is discarded as lies. It also means that almost all

attempted lies or efforts to deceive are recognised. Eventually, these results would lead to the use of other intelligence sources, and further investigative efforts would be conducted, all which would likely make up for the missing 20%–30% of overall accuracy.

The results of this study enable the development of the current and future human intelligence methods and supporting techniques and procedures, within the Finnish Defence Forces and other organisation working with the human source intelligence collection efforts. Not only counterintelligence operations but also active human intelligence collection operations would benefit from improved understanding on the strengths and weaknesses of studied veracity assessment and lie detection methods. The studied methods could be used to assess the motives and goals of the future walk-in agents, intelligence retrieved from active human intelligence sources, intelligence collected from prisoners of war interrogations or after-action debriefings, to mention few.

There is still room for improvement with the questioning protocol, data collection and data processing. The used questioning protocol did not benefit from the use of control and irrelevant questions when the interviewees were confronted with potentially compromising heavy evidence. The data collection method concerning cognitive load was mainly manual. Automated speech recognition, recording the speech errors and measuring the speech rate, would enable almost simultaneous recognition of the selected speech indicators. This pilot study proved that the used concept is feasible, and it is worthwhile to continue the development of the method. For the future research, it is recommended to repeat the study with a larger sample size. To test the effect of the heightened cognitive load, a statement in the reverse order could be added at the end of the free telling part of the interview.

Ethical standards

Disclosure statement

The authors report there are no competing interests to declare.

Data availability statement

The data that support the findings of this study are available from the corresponding author, (Uotinen, M), upon reasonable request.

Ethical approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee (National Defence University, MPKK_AP5977_2019) and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent

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

Funding

This research is part of the doctoral work of the first author and was supported by the Werner Hacklin Foundation under Grant (23042019) to the first author.

Acknowledgements

The authors wish to thank everyone who assisted and otherwise took part in the preparations of the data collection and other supporting tasks.

Author biography

Marko Uotinen is a postgraduate student at the National Defence University.

Petteri Simola (PhD, psych.) is a research manager at the Finnish Defence Research Agency and an associate professor (military psychology) at the National Defence University.

Pentti J. Henttonen (MA) is a project planner at the Swedish School of Social Science, University of Helsinki.

References

- Benedek, M., & Kaernbach, C. (2010). A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods*, 190(1), pp. 80-91. doi: 10.1016/j.jneumeth.2010.04.028.
- Braithwaite, J. J., Watson, D. G., Jones, R., & Rowe, M. (2015). *A Guide for Analysing Electrodermal Activity (EDA) & Skin Conductance Responses (SCRs) for Psychological Experiments (Revised version: 2.0)*. Technical Report, 2nd version. Selective Attention & Awareness Laboratory (SAAL) Behavioural Brain Sciences Centre, University of Birmingham, UK. URL: <https://>

- www.birmingham.ac.uk/documents/college-les/psych/saal/guide-electrodermal-activity.pdf
- DePaulo, B. M., & Kashy, D. A. (1998). Everyday lies in close and casual relationships. *Journal of Personality and Social Psychology*, 74(1), pp. 63-79. doi: 10.1037/0022-3514.74.1.63.
- George, J. F., Tilley, P., & Giordano, G. (2014). Sender credibility and deception detection. *Computers in Human Behavior*, 35, pp. 1-11. doi: 10.1016/j.chb.2014.02.027.
- Granhag, P. A., & Vrij, A. (2005). Deception detection. In: Brewer, N., & Williams, K. D. (eds.), *Psychology and Law: An Empirical Perspective*. The Guilford Press, New York City, pp. 43-92.
- Halevy, R., Shalvi, S., & Verschuere, B. (2014). Being honest about dishonesty: Correlating self-reports and actual lying. *Human Communication Research*, 40(1), pp. 54-72. doi: 10.1111/hcre.12019.
- Hart, C. L., Fillmore, D. G., & Griffith, J. D. (2009). Indirect detection of deception: Looking for change. *Current Research in Social Psychology*, 14(9), pp. 134-142.
- Iacono, W. G. (2008). Accuracy of polygraph techniques: Problems using confessions to determine ground truth. *Psychology and Behavior*, 95(2), pp. 24-26. doi: 10.1016/j.physbeh.2008.06.001.
- Kirchhübel, C., & Howard, D. M. (2013). Detecting suspicious behaviour using speech: Acoustic correlates of deceptive speech – An exploratory investigation. *Applied Ergonomics*, 44(5), pp. 694-702. doi: 10.1016/j.apergo.2012.04.016.
- Levine, T. R., Daiku, Y., & Masip, J. (2022). The number of senders and total judgments matter more than sample size in deception-detection experiments. *Perspectives on Psychological Science*, 17(1), pp. 191-204. doi: 10.1177/1745691621990369.
- Memon, A., & Higham, P. A. (1999). A review of the cognitive interview. *Psychology Crime & Law*, 5(1-2), pp. 177-196. doi: 10.1080/10683169908415000.
- Memon, A., Meissner, C. A., & Fraser, J. (2010). The cognitive interview: A meta-analytic review and study space analysis of the past 25 years. *Psychology, Public Policy, and Law*, 16(4), pp. 340-372. doi: 10.1037/a0020518.
- Nunan, J., Stanier, I., Milne, R., Shawyer, A., & Walsh, D. (2020). Eliciting human intelligence: Police source handlers' perceptions and experiences of rapport during covert human intelligence sources (CHIS) interactions. *Psychiatry, Psychology and Law*, 27(4), pp. 511-537. doi: 10.1080/13218719.2020.1734978.
- Palena, N., Caso, L., & Vrij, A. (2019). Detecting lies via a theme-selection strategy. *Frontiers in Psychology*, 9, p. 2775. doi: 10.3389/fpsyg.2018.02775.
- Palmatier, J. J., & Rovner, L. (2015). Credibility assessment: Preliminary process theory, the polygraph process, and construct validity. *International Journal of Psychophysiology*, 95(1), pp. 3-13. doi: 10.1016/j.ijpsycho.2014.06.001.
- Penttilä, N., Korpipää-Huuhka, A.-M., & Kent, R. D. (2018). Tavallista sujumattomuutta: Aikuisten puheen sujuvuuden kvantitatiivinen analyysi. *Puhe ja kieli*, 38(3), pp. 153-173. doi: 10.23997/pk.77384.
- Van Koppen, P. J. (2012). Deception detection in police interrogations: Closing in on the context of criminal investigations. *Journal of Applied Research in Memory and Cognition*, 1(2), pp. 124-125. doi: 10.1016/j.jarmac.2012.04.005.
- Verigin, B., Meijer, E. H., Vrij, A., & Zauzig, L. (2020). The interaction of truthful and deceptive information. *Psychology Crime & Law*, 26(4), pp. 367-383. doi: 10.1080/1068316X.2019.1669596.
- Verschuere, B., Meijer, E., & Vrij, A. (2016). Increasing pressure does not benefit lie detection: A reply to Ten Brinke et al. (2015). *Psychology Crime & Law*, 22(9), pp. 915-920. doi: 10.1080/1068316X.2016.1202249.
- Vrij, A. (2016). Baseline as a lie detection method. *Applied Cognitive Psychology*, 30(6), pp. 1112-1119. doi: 10.1002/acp.3288.
- Vrij, A., Edward, K., Roberts, K. P., & Bull, R. (2000). Detecting deceit via analysis of verbal and nonverbal behavior. *Journal of Nonverbal Behavior*, 24(4), pp. 239-263. doi: 10.1023/A:1006610329284.
- Vrij, A., Ennis, E., Farman, S., & Mann, S. (2010). People's perceptions of their truthful and deceptive interactions in daily life. *Open Access Journal of Forensic Psychology*, 2, pp. 6-42.
- Vrij, A., Fisher, R. P., & Leal, S. (2022). How researchers can make verbal lie detection more attractive for practitioners. *Psychiatry, Psychology and Law*, 30, pp. 383-396. doi: 10.1080/13218719.2022.2035842.
- Vrij, A., & Granhag, P. A. (2012). Eliciting cues to deception and truth: What matters are the questions asked. *Journal of Applied Research in Memory and Cognition*, 1(2), pp. 110-117. doi: 10.1016/j.jarmac.2012.02.004.
- Vrij, A., Hartwig, M., & Granhag, P. A. (2019). Reading lies: Nonverbal communication and deception. *Annual Review of Psychology*, 70, pp. 295-317. doi: 10.1146/annurev-psych-010418-103135.
- Vrij, A., Hope, L., & Fisher, R. P. (2014). Eliciting reliable information in investigative interviews. *Policy Insights from the Behavioral and Brain Sciences*, 1(1), pp. 129-136. doi: 10.1177/2372732214548592.
- Vrij, A., Mann, S. A., Fisher, R. P., Leal, S., Milne, R., & Bull, R. (2008). Increasing cognitive load to facilitate lie detection: The benefit of recalling events in reverse order. *Law and Human Behavior*, 32(3), pp. 253-265. doi: 10.1007/s10979-007-9103-y.
- Walczyk, J. J., Igou, F. P., Dixon, A. P., & Tcholokian, 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, p. 14. doi: 10.3389/fpsyg.2013.00014.
- Zuckerman, M., DePaulo, B. N., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. *Advances in Experimental Social Psychology*, 14, pp. 1-59. doi: 10.1016/S0065-2601(08)60369-X.