

**Systematic Errors (Biases) in Applying Verbal Lie Detection Tools: Richness in
Detail as a Test Case**

Galit Nahari¹

Department of Criminology, Bar-Ilan University

Email: galit.nahari@biu.ac.il

Aldert Vrij

Psychology Department, University of Portsmouth

Email: aldert.vrij@port.ac.uk

¹ Correspondence concerning this article should be addressed to Galit Nahari,
Department of Criminology, Bar-Ilan University, Ramat Gan, 52900, Israel. E-mail:
galit.nahari@biu.ac.il.

Abstract

The current paper describes potential systematic errors (or biases) that may appear while applying content-based lie detection tools, by focusing on richness in detail – a core indicator in verbal tools - as a test case. Two categories of biases are discussed: those related to the interviewees (i.e., interviewees with different characteristics differ in the number of details they provide when lying or telling the truth) and those related to the tool expert (i.e., tool experts with different characteristics differ in the way they perceive and interpret verbal cues). We suggested several ways to reduce the influence of these biases, and emphasized the need for future studies in this matter.

Systematic Errors (Biases) in Applying Verbal Lie Detection Tools: Richness in Detail as a Test Case

Already around 900 B. C. content quality was mentioned as indicator for distinguishing truths from lies. A papyrus of the Vedas stated that a poisoner ‘does not answer questions, or gives evasive answers; he speaks nonsense’ (Troville, 1939, p. 849). A systematic search for verbal cues to deceit, often examined as part of a verbal veracity assessment tool, has accelerated since the 1950s (DePaulo et al., 2003; Hauch, Blandón-Gitlin, Masip, & Sporer, in press; Masip, Sporer, Garrido, & Herrero, 2005; Vrij, 2008).

The three verbal tools most frequently used by scholars or practitioners are Criteria-Based Content Analysis (CBCA), Reality Monitoring (RM) and Scientific Content Analysis (SCAN). Each tool consists of a list of criteria for discriminating truths from lies. CBCA is the core part of Statement Validity Analysis (SVA), which originates from Sweden (Trankell, 1972) and Germany (Undeutsch, 1982), and was designed to determine the credibility of *child* witnesses’ testimonies in trials for *sexual offences*. CBCA comprises 19 criteria and CBCA-trained evaluators judge the strength of presence of each of these criteria in an interview transcript. According to Köhnken (1996) several CBCA criteria are likely to occur in truthful statements for cognitive reasons as they are typically too difficult to fabricate, while other criteria are more likely to occur in truthful statements for motivational reasons, as liars may leave out information that, in their view, will damage their image of being honest. The presence of each criterion strengthens the hypothesis that the account is based on genuine personal experience. Therefore, truthful statements are expected to generate higher CBCA scores than false statements. CBCA is the most frequently researched verbal veracity tool to date and more than 50 studies have been published examining the working of this tool. It

is used as evidence in criminal courts in North American and several West-European countries including Germany, the Netherlands, Spain and Sweden (Vrij, 2008).

RM is, to our knowledge, never used in real life but it is popular amongst scholars, presumably because it has a solid theoretical background originating from the field of memory (Johnson & Raye, 1980). The RM lie detection tool consists of a set of eight content criteria assessing the presence of contextual, perceptual and cognitive operation attributes as well as the realism of the described event, the general clarity of the description, and the ability to reconstruct the event based on the given information (Sporer, 1997, 2004). According to the RM lie detection approach, truths (which are based on perceptual experience) are likely to be richer in detail and contain *perceptual information* (details of sound, smell, taste, touch, or visual details) and *contextual information* (spatial details about where the event took place, and details about how objects and people were situated in relation to each other, e.g., "Fred stood behind me" and temporal details about time order of the events, e.g., "First he switched on the video-recorder and then the TV", and details about duration of events). These memories are usually clear, sharp and vivid. In contrast, lies (which are based on self-generated thoughts or imagination) are likely to contain *cognitive operations*, such as thoughts and reasoning ("I must have had my coat on, as it was very cold that night"). They are usually vaguer and less concrete. Similar to the CBCA coding protocol, oral statements are transcribed and trained RM coders judge the strength of presence of the RM criteria in the transcripts. More than 30 RM deception studies have been published to date.

SCAN was developed by Avinoam Sapir, a former polygraph examiner in the Israeli police. However, despite its name, *Scientific Content Analysis*, no theoretical justification is given as to why truth tellers and liars would differ from each other regarding the suggested parameters. In this method, the examinee is asked to write

down in detail all his/her activities during a critical period of time in such a way that a reader without background information can determine what actually happened. The handwritten statement is then analysed by a SCAN expert on the basis of a list of criteria. It is thought that some SCAN criteria are more likely to occur in truthful statements than in deceptive statements, whereas other criteria are more likely to occur in deceptive statements than in truthful statements (Sapir, 1987/2000).

There is not a fixed list of SCAN criteria and different experts seem to use different sets of criteria. A list of 12 criteria is mostly used in workshops on the technique (Driscoll, 1994), in research (Smith, 2001), or by SCAN users in a field observation (Bogaard, Meijer, Vrij, Broers, & Merckelbach, 2014). SCAN is popular in the field and probably frequently and widely used. It is used in countries such as Australia, Belgium, Canada, Israel, Mexico, UK, US, the Netherlands, Qatar, Singapore, and South Africa (Vrij, 2008), and by federal law enforcement (including the FBI), military agencies (including the US Army Military Intelligence), secret services (including the CIA) and other types of investigators (including social workers, lawyers, fire investigators and the American Society for Industrial Security) (Bockstaele, 2008; www.lsiscan.co.il). <http://www.lsiscan.com/id29.htm> provides a full list of past participants of SCAN Courses. According to (the American version of) the SCAN website (www.lsiscan.com) SCAN courses are mostly given in the US and Canada (on a weekly basis in those countries). In addition, online courses are also available. However, in contrast with its popularity in the field, SCAN has hardly been researched.

In terms of accuracy in distinguishing lies from truths, CBCA and RM achieved similar accuracy rate of around 70% (Vrij, 2008). Due to the little research, accuracy rate of SCAN is unknown, but seems to be much lower than that of CBCA and RM (see Nahari & Vrij, 2012). Obviously, verbal lie detection tools' decisions are not free from

errors and interviewees are sometimes wrongly classified as truth-tellers or liars. The very first step in the attempt to reduce such errors is to recognise them. In the current paper, we discuss systematical errors (known as 'biases') that may occur when applying verbal tools to determine veracity, and suggest several ways to decrease their effect.

Biases in veracity verbal tools: richness in detail as a test case

Systematic errors, known as 'biases', are related to external factors that affect measurements and decisions. Examples of such external factors in relation to verbal tools are the personality of the expert, his/her prior expectations about the veracity of a statement or the verbal style of the interviewee. Since these factors are not part of the tool, and thus are not considered in the application of the tool, they may lead to biased decisions. In the current paper, we discuss two categories of biases – those related to the interviewee and those related to the tool expert. Our discussion focuses on 'richness in details' as a test case. Richness in detail is an important component of RM and CBCA tools, and refers to the number of details mentioned by the interviewee regarding spatial and temporal information, descriptions of people, objects and places, conversations, emotions, senses, tastes, smells, and the like. Four of the 19 criteria of CBCA (quantity of detail, contextual embedding, descriptions of interactions, reproduction of conversation) and five of the eight RM criteria (clarity, perceptual information, spatial information, temporal information affects; see Sporer, 2004; Vrij, 2008) are measuring aspects of richness in details. Furthermore, the CBCA and RM individual criteria that were found to be the most diagnostic (Vrij, 2008) are measuring richness in detail (quantity of details, contextual embedding and reproduction of conversation). None of the 12 SCAN criteria refers directly to richness in detail. However, one can assume that several of them are indirectly affected by the

amount of details provided. For example, it might be difficult to analyze *Objective and subjective time* and *change of language* in a statement that is poor in detail.

In the current section, we discuss potential biases in determining richness in detail in a statement. We describe several external factors – related either to the interviewee or to the expert - that affect richness in detail and subsequently may lead to incorrect veracity judgements.

Biases related to the interviewees

Richness in detail of a statement is affected by the interviewees' style of language, quality of their memory, type of lie they use, and their awareness of the working of the tool.

Style of Language

There are differences among interviewees in the number of details they provide when lying or telling the truth (Nahari & Vrij, 2014). For example, Nahari et al. (2012) reported standard deviations of 91.35 (truth tellers) and 53.60 (liars) in the number of words spoken, which were large compared to the total number of words spoken (243.14 by truth tellers and 129.20 by liars). Nahari & Vrij (2014) further showed that the tendency to provide rich or poor statements is stable across situations (i.e., stable when discussing different topics), which implies that this tendency is not random but related to personal characteristics. Indeed, the few studies that have examined individual differences have revealed significant effects. For example, public self-consciousness and ability to act were negatively correlated with RM scores (Vrij, Edward, & Bull, 2001), and high and low fantasy prone individuals gave different descriptions (in terms of content qualities) of incidents they had experienced

(Merckelbach, 2004; Schelleman-Offermans & Merckelbach, 2010). Newman, Groom, Handelman, and Pennebaker (2008), who analyzed 14,000 texts collected from females and males, showed that texts provided by females included more senses (e.g., touch, hold, feel), sound details (e.g., heard, listen, sounds), motion verbs (e.g., walk, go) and emotions than texts provided by males. In accordance with this, Nahari and Pazualo (2015) found that females' truthful accounts were richer in detail than males' truthful accounts.

Individual differences in richness in detail may lead to systematic errors in RM and CBCA decisions, especially for interviewees who provide relatively few or many details. Thus, a liar who tends to provide relatively many perceptual and contextual details (e.g., a high fantasy prone individual) might be misattributed as a truth-teller, whereas a truth teller who tends to provide relatively few perceptual and contextual details (e.g., a man) may be misattributed as liar. Furthermore, if interviewees differ in the amount of details they include in their truthful statements, it is difficult to know how many details to expect in a truthful statement, and consequently impossible to establish 'norms' or 'cut-off points' for veracity assessments based on richness in detail. Cut-off points are crucial in applying a tool in the field, where someone needs a criterion to decide whether the score obtained for a single statement is high enough (i.e., rich in details) to conclude that the interviewee is telling the truth, or low enough (i.e., poor in details) to conclude that the interviewee is lying. This may be a serious limitation, especially for RM, which is based mainly on assessing richness in detail.

A potential solution is to develop within-subject lie detection tools that allow for idiosyncrasies and whereby the richness of the statement under investigation is assessed relatively to the richness of a truthful statement given by the same

interviewee. Within-subject evaluations are widely applied in psycho-physiological lie detection, including in the Concealed Information Test (CIT). The CIT utilises a series of multiple-choice questions, each having one relevant alternative (e.g., a feature of the crime under investigation that should be known only to the perpetrator) and several neutral (control) alternatives, selected so that an innocent suspect would not be able to discriminate them from the relevant alternative (Lykken, 1959, 1960, 1998). Due to individual differences in physiological responses during a polygraph test (Ben-Shakhar, 1985; Lacey & Lacey, 1958), suspects' responses to the relevant items are compared to their responses to the neutral items. Liars are thought to display stronger physiological responses to the relevant alternatives than to the control alternatives, whereas no differences in responses are expected in truth tellers.

Such a within-subject measurement may be difficult to achieve in many situations in which verbal tools can be applied. Indeed, an investigator could ask the interviewee to provide a truthful statement which will function as a "baseline statement", and then compare the statement under investigation with that baseline statement. For example, if the investigator wants to assess the veracity of the interviewee's statement regarding his/her activities on a Tuesday night, the investigator can first ask the interviewee to discuss his/her Monday night activities (baseline statement), and subsequently ask about the Tuesday night activities (statement under investigation). Obviously, the deceptive interviewee may realise what the investigator is trying to achieve and may therefore deliberately give a baseline and investigation statements that are similar in the number of details. Despite these problems in applying it, a within-subjects tool may reduce errors resulting from individual differences between interviewees, and is thus worthwhile to develop in future studies.

Quality of memory

The number of perceptual and contextual details in a text can be influenced by factors other than the veracity of the account. One such factor is memory. When providing a statement, truth tellers are dependent on their memory, which decreases over time (Carmel, Dayan, Naveh, Raveh & BenShakhar, 2003; Nahari & Ben-Shakhar, 2011). Therefore, truth tellers provide fewer details when they are interviewed about an event that occurred in the distant past compared to an event that occurred recently (Sporer & Sharman, 2006; Vrij et al., 2009). This implies that one should avoid assessing events that happened a long time ago, or at least take the time factor into account, and set appropriate norms for assessing richness in detail in distant past events.

Type of lie

Not all lies are pure fabrications, liars frequently embed true details in a false statement, so called embedded lies (Leins, Fisher, & Ross, 2013; Vrij, 2008; Vrij, Granhag, & Porter, 2010). Such embedded lies can be largely truthful (Nahari, Vrij, & Fisher, 2014b). Richness in detail might be affected by the liar's strategy to use embedded lies, because such lies contain, by definition, truthful perceptual and contextual details (Nahari, Vrij, & Fisher, 2014b). It is difficult to naturalize liars' strategy to report an embedded lie, but it is usually possible to recognize whether the liar has an opportunity to do so. For example, Nahari et al. (2014b) showed that liars sometimes presented their criminal activities (e.g., copying a stolen exam in the library) as a non-criminal activity (e.g., copying an article in the library). Liars could report such a false account relatively easily, because their presence in the library at the time that the crime was committed was legitimate (it was during opening times of the library), and liars were therefore able to report true details, such as real conversations

they had in the library or paying by credit card for the copies. It will be much more difficult for liars to apply such a strategy in situations where their presence at the crime scene at a specific time is not legitimate (e.g., when a crime has been committed in the library after closing time). In that situation liars cannot admit that they were in the library at the time the crime occurred, but need to pretend they were somewhere else. It may be more difficult for them to tell an embedded lie when they claim to have been somewhere else. Thus, an investigator should consider whether the presence of the suspect at the crime scene when the crime took place is legitimate. If this is the case, the investigator should be aware that a suspect could tell an embedded lie (see Nahari & Vrij, 2015), and thus should be careful with interpreting their richness in detail scores.

Awareness to the working of the tool

The amount of details provided in a statement can further be influenced by the interviewees' knowledge regarding the mechanism of verbal tools. Participants who received insight into certain tool criteria, and who were instructed to include those criteria in their statements provided more details in their statements and thus improved their CBCA (Vrij, Akehurst, Soukara, & Bull, 2002, 2004) and RM (Nahari & Pazualo, 2015) scores. . Another study showed a similar effect for less explicit informing: Mere exposure of interviewees to an audiotape of a detailed account, as a model example, doubled the amount of information provided by both truth tellers and liars (Leal, Vrij, Warmelink, Vernham, & Fisher, 2015). Presumably, liars understood that providing a statement rich in detail was expected from them and they managed to add false details to their statements.

A simple tactic, and an integral part of the verifiability approach (Nahari, Vrij, & Fisher, 2014a, 2014b), may help to decrease liar's ability to deliberately add false details to their statements: The investigator warns the interviewee, before s/he starts presenting his/her account, that the investigator may check the truthfulness of all or some of the details provided by the interviewee. Obviously, this tactic cannot eliminate the possibility of adding false details, but may reduce the number of false details provided.

Biases related to the tool expert

Richness in detail of a statement is further affected by the personal characteristics (e.g., gender, experience etc.) of the tool experts (i.e, the coders or analyst) and by their cognitive biases.

Individual differences

People may differ from each other in the way they perceive and interpret verbal cues in the same statement (Nahari, Glicksohn, & Nachson, 2010). For example, in a study by Granhag and Strömwall (2000), participants disagreed about the degree of richness of a specific statement, whereby some considered the statement as poor in detail and others as rich. Such disagreements may result in low internal reliability of the verbal lie detection tool (Masip, sporer, et al., 2005; Sporer, 1997).

Differences in judgments between coders may be related to receiver characteristics. Nahari (2012) found that the very same statement was assessed by professional lie detectors (e.g., police officers) as being poor in detail and by laypersons (students) as being rich in detail, presumably because of the tendency of the former to be suspicious (Masip, Alonso, Garrido, & Anton, 2005). In another

study, Nahari and Vrij (2014) found that the perception of richness of other people's statements depended upon the tendency of the receivers to tell a rich story themselves. Specifically, the richer their own statements were compared to the other person's statement, the more critical they were when evaluating the other person's statement. A likely explanation for these disagreements is that individuals use different scales for judging richness in detail, and thus arrive at different conclusions. An individual's own behavior may determine his/her "decision threshold" (Glicksohn, 1993-94), thus it is possible that individuals who tend to provide rich accounts themselves may have higher expectations regarding the amount of details a statement of another person should contain to be judged as truthful.

These findings suggests that richness in detail assessments do not only depend on the quality of the statement (as it should) but also on characteristics of the tool expert which could lead to biases. Such biases can be minimized in several ways. First, assessing richness in details by counting the details instead of rating the richness on a scale should be preferred. Although counting is more time-consuming, it is much less sensitive to individual differences than scale rating (Nahari, 2015). Second, it is better to have several coders rather than one coder, not only in research, but also in real-life practice. By using several coders, biases related to the subjectivity of the coders can be detected. Third, perhaps human assessment could be accompanied by a computerized analysis. However, it is far from certain that human analyses can be replaced by computer software, since content analysis should consider the context of words which is difficult to achieve via a computer. For example, if an interviewee says "I went to the supermarket because I will cook dinner tomorrow", someone should not consider "tomorrow" as a time detail and "cook" as a perceptual detail. "I will cook dinner tomorrow" is an explanation for the shopping and not part of the

activities at the time under investigation. To develop a software package that picks up such subtleties is challenging.

Cognitive biases

Judgments of tool experts can also be affected by cognitive biases, such as primacy effect (the disproportional influence of information acquired early in the process on the final judgment; e.g., Bond, Carlson, Meloy, Russo, & Tanner, 2007; Nickerson, 1998) or confirmation bias (the tendency to unconsciously seek and interpret behavioral data in a way that verifies the first impression or prior expectations about the object in question; see Kassin, Goldstein, & Savitsky, 2003). For example, in Nahari and Ben-Shakhar (2013) participants read one of two versions of the same story. Both versions contained the same amount of perceptual and contextual details, but differed in their structure. One version began in a poor manner and ended in a rich manner, that is, most of the perceptual and contextual details appeared in the second half of the story. The other version began in a rich manner and ended in a poor manner, that is, most of the perceptual and contextual details appeared in the first half of the story. Results showed that participants who were exposed to the version that began in a rich manner rated the story as being richer in perceptual and contextual details compared to those who were exposed to the version that began in a poor manner. Presumably, participants interpreted information that came later in accordance with the impressions they formed based on the initial information, demonstrating a primacy effect. Thus, their assessments did not just reflect the actual attributes of the text, they were also influenced by prior impressions and expectations.

Another example of a cognitive bias was provided by Bogaard, Meijer, Vrij, Broers, and Merckelbach (2014). They showed that ratings of RM criteria were

affected by extra-domain information, so that statements which were accompanied by positive information (e.g., a positive eye-witness identification) were judged as being richer in RM criteria than statements which were accompanied by negative information (e.g., a personal background that implied a history of lying). In other words, the accompanied information *contaminated* the ratings of the RM criteria, demonstrating a confirmation bias. Ben-Shakhar (1991) reasoned that contamination of a decision due to external information will be more likely to occur when the veracity assessment tool is subjective and has no well-defined decision rules. Considering the lack of standardization in verbal content analysis tools (e.g., Vrij, 2008), and disagreements between coders regarding the rating of content criteria (Masip, Sporer, et al., 2005; Sporer, 1997), contamination may be an actual risk when using these tools, in spite of their "face value" as objective techniques.

Awareness of the potential influence of cognitive biases may reduce its biasing impact. For example, Schuller & Hastings (2002) found that the influence of information about the sexual history between a complainant and defendant on the perceived credibility of the complainant was moderated when participants were instructed not to use the sexual history in their evaluations. However, a meta-analysis, focusing on the ability of juries to ignore inadmissible evidence, showed that instructing jurors to ignore evidence was effective only when the judicial reasoning was provided as to why the evidence was unreliable (Stebly, Hosch, Culhane, & McWethy, 2006). Therefore, in order to minimize the vulnerability of content analysis tools to cognitive biases, one should provide coders with a good explanation of the potential biases and their mechanisms. For example, coders should be aware that details can be unevenly distributed throughout statements. In addition, the influence of extra-domain information or additional evidence on verbal criteria coding can be

counteracted by leaving the coder unaware of that extra-domain information or additional evidence.

Conclusions

In this paper we discussed several systematic errors that occur when using verbal tools to determine veracity and we suggested several ways to reduce their influence. Those who apply verbal veracity tools should consider the mechanism of these tools, their suitability to a specific expert or interviewee, and should take into account alternative explanations for the results. Specifically, when there is a large time gap between the investigation and the criminal event or when an interviewee has difficulties in expressing him/herself verbally, one may avoid using verbal tools to determine veracity. In addition, it is important to take into account that liar's may apply strategies to hamper the effectiveness of the tool, especially if they are aware of the working of the tool or when the situation allows embedding true details into their lies. One should also take steps to reduce subjective influences on decisions (e.g., counting details rather than using a scale) and to avoid cognitive biases (e.g., to keep the coder blind to prior information). We would like to see more studies in the important but largely neglected domain of examining systematic errors in applying verbal veracity tools and to develop solutions to overcome such errors.

References

- Ben-Shakhar, G. (1985). Standardization within individuals: A simple method to neutralize individual differences in psychophysiological responsivity. *Psychophysiology*, *22*, 292-299.
- Ben-Shakhar, G. (1991). Clinical judgment and decision-making in CQT-polygraphy. *Integrative Psychological and Behavioral Science*, *26*, 232-240. doi: 10.1007/BF02912515
- Bockstaele, M. (2008). Scientific Content Analysis (SCAN). Een nutting instrument by verhoren? In L. Smets & A. Vrij (Eds.), *Het analyseren van de geloofwaardigheid van verhoren* (pp. 105-156). Brussels: Politeia.
- Bogaard, G., Meijer, E. H., Vrij, A., Broers, N. J., & Merckelbach, H. (2014). Contextual bias in verbal credibility assessment: Criteria-Based content analysis, Reality Monitoring and Scientific Content Analysis. *Applied Cognitive Psychology*, *28*, 79 - 90. doi: 10.1002/acp.2959
- Bond, S. D., Carlson, K. A., Meloy, M. G., Russo, J. E., & Tanner, R. J. (2007). Information distortion in the evaluation of a single option. *Organizational Behavior and Human Decision Processes*, *102*, 240-254. doi: 10.1016/j.obhdp.2006.04.009
- Carmel, D., Dayan, E., Naveh, A., Raveh, O., & Ben-Shakhar, G. (2003). Estimating the validity of the Guilty Knowledge Test from simulated experiments: The external validity of Mock Crime studies. *Journal of Experimental Psychology: Applied*, *9*, 261-269. doi: 10.1037/1076-898X.9.4.261
- DePaulo, B. M., Lindsay, J. L., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin*, *129*, 74-118.

- Driscoll, L. N. (1994). A validity assessment of written statements from suspects in criminal investigations using the SCAN technique. *Police Studies, 17*, 77-88.
- Glicksohn, J. (1993-1994). Rating the incidence of an altered state of consciousness as a function of the rater's own absorption score. *Imagination, Cognition and Personality, 13*, 225-228.
- Granhag, P. A., & Strömwall, L. A. (2000). Effects of preconceptions on deception detection and new answers to why lie-catchers often fail. *Psychology, Crime & Law, 6*, 197-218.
- Granhag, P. A., Strömwall, L. A., & Landström, S. (2006). Children recalling an event repeatedly: Effects on RM and CBCA scores. *Legal and Criminological Psychology, 11*, 81-98. doi: 10.1348/135532505X49620
- Hauch, V., Blandón-Gitlin, I., Masip, J., Sporer, S. L. (in press). Are computers effective lie detectors? A meta-analysis of linguistic cues to deception. *Personality and Social Psychology Review*.
- Johnson, M. K. (2006). Memory and reality. *American Psychologist, 61*, 760-771. doi: 10.1037/0003-066X.61.8.760
- Johnson, M. K., Bush, J. G., & Mitchell, K. J. (1998). Interpersonal reality monitoring: Judging the sources of other people's memories. *Social Cognition, 16*, 199-224. doi: 10.1521/soco.1998.16.2.199
- Johnson, M. K., Foley, M. A., Suengas, A. G., & Raye, C. L. (1988). Phenomenal characteristics of memories for perceived and imagined autobiographical events. *Journal of Experimental Psychology: General, 117*, 371-376. doi: 10.1037/0096-3445.117.4.371

- Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review*, 88, 67-85. doi: 10.1037/0033-295X.88.1.67
- Kassin, S. M., Goldstein, C. C., & Savitsky, K. (2003). Behavioral confirmation in the interrogation room: On the dangers of presuming guilt. *Law and Human Behavior*, 27, 187-203. doi: 10.1023/A:1022599230598
- Köhnken, G. (1996). Social psychology and the law. In G. R. Semin, & K. Fiedler (Eds.), *Applied Social Psychology* (pp. 257-282). London, Great Britain: Sage Publications.
- Leal, s., Vrij, A., Warmelink,L., & Fisher, R. P. (2015). You cannot hide your telephone lies: Priming as an aid to detect deception in insurance telephone calls. *Legal and Criminological Psychology*, 20, 129–146.
<http://dx.doi.org/10.1111/lcrp.12017>
- Lacey, J. I., & Lacey, B. C. (1958). Verification and extension of the principle of autonomic response-stereotypy. *The American Journal of Psychology*, 71, 50-73.
- Leins, D. A., Fisher, P. A., & Ross, S. J. (2013). Exploring liars' strategies for creating deceptive reports. *Legal and Criminological Psychology*, 18, 141-151. doi: 10.1111/j.2044-8333.2011.02041.x
- Lykken, D.T. (1959). The GSR in the detection of guilt. *Journal of Applied Psychology*, 43, 385-388.
- Lykken, D.T. (1960). The validity of the guilty knowledge technique: The effects of faking. *Journal of Applied Psychology*, 44, 258-262.
- Lykken, D.T. (1998). *A Tremor in the Blood: Uses and Abuses of the Lie Detector*. New York: Plenum Trade.

- Masip, J., Alonso, H., Garrido, E., & Anton, C. (2005). Generalized communicative suspicion (GCS) among police officers: Accounting for the investigator bias effect. *Journal of Applied Social Psychology, 35*, 1046-1066.
- Masip, J., Sporer, S. L., Garrido, E., & Herrero, C. (2005). The detection of deception with the reality monitoring approach: A review of the empirical evidence. *Psychology, Crime & Law, 11*, 99-122. doi: 10.1080/10683160410001726356
- Merckelbach, H. (2004). Telling a good story: Fantasy proneness and the quality of fabricated memories. *Personality and Individual Differences, 37*, 1371-1382. doi: 10.1016/j.paid.2004.01.007
- Nahari (2015). Reality Monitoring (RM) coding: A comparison between two methods. Paper presented at the 25th European Association of Psychology and Law.
- Nahari, G. (2012). Elaborations on Credibility Judgments by Professional Lie Detectors and Laypersons: Strategies of Judgment and Justification. *Psychology, Crime, & Law, 18*, 567-577. doi: 10.1080/1068316X.2010.511222
- Nahari, G., & Ben-Shakhar, G. (2011). Psychophysiological and behavioral measures for detecting concealed information: The role of memory for crime details. *Psychophysiology, 48*, 733-744. doi: 10.1111/j.14698986.2010.01148.x
- Nahari, G., & Ben-Shakhar, G. (2013). Primacy effect in credibility judgments: The vulnerability of verbal cues to biased interpretations. *Applied Cognitive Psychology, 27*, 247-255. doi: 10.1002/acp.2901
- Nahari, G., Glicksohn, J., & Nachson, I. (2010). Credibility judgments of narratives. *American Journal of Psychology, 123*, 319-335.

- Nahari & Pazuelo (in press). Telling a convincing story: Richness in detail as a function of gender and information. *Journal of Applied Research in Memory and Cognition*. <http://dx.doi.org/10.1016/j.jarmac.2015.08.005>.
- Nahari, G., & Vrij, A. (2014). Are you as good as me at telling a story? Individual differences in interpersonal reality monitoring. *Psychology, Crime & Law*, *20*, 573 - 583. doi: 10.1080/1068316X.2013.793771
- Nahari, G., Vrij, A., & Fisher, R. P. (2012). Does the truth come out in the writing? Scan as a lie detection tool. *Law and Human Behavior*, *36*, 68-76. doi: 10.1037/h0093965
- Nahari, G., Vrij, A., & Fisher, R. P. (2014a). Exploiting liars' verbal strategies by examining unverifiable details. *Legal and Criminological Psychology*, *19*, 227 - 239. doi: 10.1111/j.2044-8333.2012.02069.x
- Nahari, G., Vrij, A., & Fisher, R. P. (2014b). The verifiability approach: Countermeasures facilitate its ability to discriminate between truths and lies. *Applied Cognitive Psychology*, *28*, 122 - 128. doi: 10.1002/acp.2974
- Newman, M. L., Groom, C. J., Handelman, L. D., & Pennebaker, J. W. (2008). Gender differences in language use: An analysis of 14,000 text samples. *Discourse Processes*, *45*, 211-236. doi:10.1080/01638530802073712
- Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*, 175-220. doi: 10.1037/1089-2680.2.2.175
- Schelleman-Offermans, K., & Merckelbach, H. (2010). Fantasy proneness as a confounder of verbal lie detection tools. *Journal of Investigative Psychology and Offender Profiling*, *7*, 247-260. doi: 10.1002/jip.121

- Schuller, R. A., & Hastings, P. A. (2002). Complainant sexual history evidence: Its impact on mock jurors' decisions. *Psychology of Women Quarterly*, 26, 252-261. doi: 10.1111/1471-6402.00064
- Sapir, A. (1987/2000). *The LSI course on scientific content analysis (SCAN)*. Phoenix, ZA: Laboratory for Scientific Interrogation.
- Smith, N. (2001). *Reading between the lines: An evaluation of the scientific content analysis technique (SCAN)*. Police research series paper 135. London: UK Home Office, Research, Development and Statistics Directorate.
- Sporer, S. L. (1997). The less travelled road to truth: Verbal cues in deception detection in accounts of fabricated and self experienced events. *Applied Cognitive Psychology*, 11, 373-397. doi: 10.1002/(SICI)1099-0720(199710)11:5<373::AID-ACP461>3.0.CO;2-0
- Sporer, S. L. (2004). Reality monitoring and detection of deception. In P. A. Granhag & L. A. Stromwall (Eds.), *The Detection of Deception in Forensic Contexts* (pp. 64-102). UK: Cambridge university press.
- Sporer, S. L., & Sharman, S. J. (2006). Should I believe this? Reality monitoring of accounts of self-experienced and invented recent and distant autobiographical events. *Applied Cognitive Psychology*, 20, 837-854. doi: 10.1002/acp.1234
- Stebly, N., Hosch, H. M., Culhane, S. E., & McWethy, A. (2006). The impact on juror verdicts of judicial instruction to disregard inadmissible evidence: A meta-analysis. *Law and Human Behavior*, 30, 469-492. doi: 10.1007/s10979-006-9039-7
- Trankell, A. (1972). *Reliability of evidence*. Stockholm, Sweden: Beckmans
- Trovillo, P. V. (1939). A history of lie detection, I. *Journal of Criminal Law and Criminology*, 29, 848-881.

- Undeutsch, U. (1982). Statement reality analysis. In A. Trankell (Ed.), *Reconstructing the past: The role of psychologists in criminal trials* (pp. 27-56). Deventer, the Netherlands: Kluwer.
- Vanderhallen, M., Jaspert, E., & Vervaeke, G. (in press). SCAN as an investigative tool. *Police Practice and Research*.
- Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities*: Chichester: John Wiley and Sons.
- Vrij, A., Akehurst, L., Soukara, S., & Bull, R. (2002). Will the truth come out? the effect of deception, age, status, coaching, and social skills on CBCA scores. *Law and Human Behavior, 26*, 261-283. doi: 10.1023/A:1015313120905
- Vrij, A., Akehurst, L., Soukara, S., & Bull, R. (2004). Let me inform you how to tell a convincing story: CBCA and Reality Monitoring scores as a function of age, coaching, and deception. *Canadian Journal of Behavioural Science, 36*, 113-126. doi: 10.1037/h0087222
- Vrij, A., Edward, K., & Bull, R. (2001). Stereotypical verbal and nonverbal responses while deceiving others. *Personality and Social Psychology Bulletin, 27*, 899-909. doi: 10.1177/0146167201277012
- Vrij, A., Granhag, P. A., & Porter, S. B. (2010). Pitfalls and opportunities in nonverbal and verbal lie detection. *Psychological Science in the Public Interest, 11*, 89-121. doi: 10.1177/1529100610390861
- Vrij, A., Leal, S., Granhag, P. A., Mann, S., Fisher, R. P., Hillman, J., & Sperry, K. (2009). Outsmarting the liars: The benefit of asking unanticipated questions. *Law and human behavior, 33*, 159 - 166. doi: 10.1007/s10979-0089143-y