

RUNNING HEAD: Drawing-based deception detection
techniques

1

Drawing-based deception detection techniques: A state-of-the-art review

Erik Mac Giolla^a

Pär Anders Granhag^{a, b}

Zarah Vernham^c

^aDepartment of Psychology,

University of Gothenburg, Gothenburg, Sweden

^bNorwegian Police University College,

Oslo, Norway

^cDepartment of Psychology,

University of Portsmouth, United Kingdom

(Word count: 5957)

Author Note

Corresponding author:

Erik Mac Giolla, Department of Psychology, University of Gothenburg, P.O. Box 500, 405 30
Gothenburg, Sweden

E-mail: Erik.Mac.Giolla@psy.gu.se

Phone: +46 31 786 1934

Fax: +46 31 786 4628

Abstract

The current article presents a concise overview of the emerging literature on drawing-based deception detection techniques. We cover the theoretical rationale of such techniques as well as the main results from the extant empirical studies. These studies have primarily looked at differences in the drawings between truth tellers and liars in terms of quality (e.g., detail, plausibility) and consistency (both within-group, and between-statement). The findings highlight drawings as a promising tool to elicit differences between truth tellers and liars on such cues. The paper also examines more practical aspects, such as practitioners' experience of the approach and preference for the approach in training studies. Finally, the susceptibility of the approach to counter-measures and directions for future research are discussed. Although research on drawing-based deception detection techniques is still very much in its infancy, results of this first round of studies are promising. They indicate the potential of incorporating drawings into real-life investigative interviews as a cheap, effective, and easy to use approach to deception detection.

Keywords: drawing; deception; cognitive load; consistency

Introduction

Recent years have seen a marked shift in research on deception detection. Many researchers acknowledge that the simple search for, and measurement of, cues to deceit is insufficient. This has ushered in a growing body of research examining strategic interviewing techniques (Vrij & Granhag, 2012). In such contexts, the lie-catcher is active and attempts to elicit cues to deception through the strategic use of questioning during an interview. Recent years has seen the development of a variety of strategic interviewing techniques, such as the strategic use of evidence (Hartwig, Granhag, Strömwall, & Kronkvist, 2006), the cognitive load approach (Vrij, Fisher, Mann, & Leal, 2008), asking unanticipated questions (Vrij et al., 2009), and the assessment criteria indicative of deception (Colwell, Hiscock-Anisman, Memon, Taylor, & Prewett, 2007). Here we provide a review of studies examining the use of drawings to elicit cues to deception. Drawings have great practical potential as a lie detection tool as they are cheap and easy to produce, understand, implement, and analyse (Vrij et al., 2010). Drawings have other potential benefits. For instance, they may overcome some of the problems that arise when an interviewer and interviewee do not speak the same language. We begin this review by outlining the theoretical rationale of drawing-based deception detection approaches. This is followed by a review of the extant literature and suggestions for future research.

Theoretical Rationale

Basic research on memory is at the heart of using drawing as a lie detection tool. Research on spatial memory indicates that humans, like other animals, are adept at remembering spatial aspects of an episodic event (Postma, Kessels, & van Asselen, 2008). This ability is invaluable both from an evolutionary perspective, as it allows animals to efficiently find food or avoid dangers in their environment, but also for modern everyday tasks, such as remembering where one left one's keys. Debate exists with regard to the automaticity of the encoding of spatial

information. Some argue that it is largely an automatic process, requiring little intentional effort (Hasher & Zacks, 1979), while others argue it is not as automatic as it may first appear (Light & Zelinski, 1983; Naveh-Benjamin, 1987). Regardless of the exact extent to which our encoding of spatial information is automatic, it seems that this ability is at least partly hardwired and is carried out with relative ease (Postma et al., 2008; Pouliot & Gagnon, 2005). The ease with which people encode and remember spatial information is supported by other areas of memory research. For instance, the picture-superiority effect demonstrates that people have better recognition and recall for pictures compared to words (Mulligan, 2014).

Perhaps the most important line of work for the current overview is research on reality monitoring (Johnson & Raye, 1981). Reality monitoring is concerned with how people distinguish between memories of events that actually occurred and memories of fabricated or imagined events. Perceptual information, such as spatial or visual information, is central to this task. In brief, memories of events that actually occurred are typically richer and contain more perceptual information than memories of imagined events. This is because when events are actually experienced different channels of encoding are engaged, which results in the acquisition of different forms of information—such as sensory and contextual information—that form the building blocks of rich episodic memories (Johnson, Foley, Suengas, & Raye, 1988).

This is important for deception detection contexts since only truth tellers describe actual experienced events. Liars present fabricated stories that should not naturally be accompanied by the perceptual information associated with autobiographical memories (Masip, Sporer, Garrido, & Herrero, 2005). In order to make their lies more convincing, liars, when given the opportunity, will prepare their lies (Hines et al., 2010). However, since we can assume that most liars will expect to present their answers verbally, it follows that their preparations should primarily be verbal in nature, taking the form of lie scripts (DePaulo et al., 2003; Köhnken, 1989). Liars'

reliance on rehearsed verbal scripts should put them at a disadvantage when asked to draw specific aspects of their lies for at least two reasons. First, liars may find it more difficult than truth tellers to produce drawings of events because drawings encourage specification and detail. Hence, if liars have prepared vague verbal statements (e.g., a brief description of a lunch date), it may be difficult to competently convert this to the detail required for pictorial drawings (e.g., showing the locations of windows, counters, entrances, and other customers). Second, liars may be less cognitively flexible than truth tellers. As noted above, genuine autobiographical memories often include a great amount of perceptual information, such as the spatial layout of a location. For this reason, truth tellers should be readily able to switch between verbal and pictorial recall modes. In fact, varying recall modes in this manner may even increase the amount of information recalled by a truthful cooperative speaker (Fisher, Schreiber Compo, Rivard, & Hirn, 2014), albeit at the expense of accuracy (Otgaar, van Ansem, Pauw, & Horselenberg, 2016). In contrast, liars who base their lies on a verbal description are perhaps less likely to have considered their lies from other perceptual vantage points. This should make liars less cognitively flexible. That is, translating such rather limited lie-scripts into a drawing may be a considerably difficult task.

Quality of Drawings

Based on the theoretical arguments outlined above, it is plausible that, when asked to draw a scene of an event, a number of differences in the drawings of truth tellers and liars should emerge. These differences can be in terms of quality (e.g., detail), but also in terms of more idiosyncratic aspects (e.g., the inclusion of bystanders). To date, five studies have compared qualitative aspects of drawings produced by truth tellers and liars (Calderon, Mac Giolla, Ask, & Granhag, 2017; van Veldhuizen et al., 2017; Vrij et al., 2009; Vrij et al., 2010; Vrij, Mann, Leal, & Fisher, 2012).

Overview of Studies

These five studies have addressed a variety of practical contexts. Vrij et al. (2009), for example, employed a mock-crime paradigm typical of deception studies. Pairs of truth tellers had lunch together, while pairs of liars performed a mock crime, and used a lunch date as their cover story. During the interview, all participants were asked to draw the layout of the restaurant that they had claimed to have had lunch in. This drawing was to include details about where they and their friend sat, where the toilets were, where the front door was, where the kitchen was, and where the closest diners were. Coders, blind to the experimental conditions, rated the drawings in terms of detail on 7-point scales (1 = not at all detailed, 7 = very detailed). The inter-rater reliability was acceptable ($r = .74$).

A different scenario was examined by Vrij et al. (2010) whereby participants (police and army officers) performed a mission consisting of collecting and delivering a package to another individual. Following the mission, they were stopped and interviewed about their activities. Half of the participants provided honest answers, the other half were told to fabricate their answers in order to mask their true objectives. During the interview, participants were asked to sketch the location of the place where they received the package. Coders, blind to the experimental conditions, rated the drawings in terms of (1) detail and (2) plausibility, on 7-point scales (1 = not at all detailed/plausible, 7 = very detailed/plausible). The inter-rater reliability for detail was excellent ($r = .87$). However, the inter-rater reliability for the plausible measure was low ($r = .48$). This was perhaps due to the operationalisation of the plausible measure which simply asked coders to what extent they “could imagine the person standing there”. In addition, drawings were rated for (3) whether or not bystanders were included and (4) whether an over-the-shoulder camera perspective or an overhead camera perspective was used. The inter-rater agreements between coders on these dichotomous variables were excellent (100% agreement for the number of bystanders and 96% agreement for the camera-perspective).

Yet another setup was used in the study by Vrij et al. (2012). This study involved participants being interviewed about their occupation. The goal was to mimic identity theft situations. Half of the participants honestly described their occupations, whilst the other half deceptively described an occupation that they had never worked in. In addition to other questions, participants were asked to sketch the layout of their workplace. Coders, blind to the experimental conditions, rated the drawings in terms of (1) detail and (2) plausibility, on 7-point scales (1 = not at all detailed/plausible, 7 = very detailed/plausible). The inter-rater reliability for these two measures were modest ($r = .66$ for detail; $r = .72$ for plausibility). In addition, coders counted the number of bystanders included in the sketch and rated the detail in which the bystanders were drawn. The inter-rater reliability for these two measures were excellent ($r = .95$ for number of details; $r = .96$ for detail of bystanders).

The study by van Veldhuizen et al. (2017) focused on issues of relevance for asylum interviews. The specific focus was on how to determine the veracity of someone's claimed country of origin, an important issue in many asylum interviews. Participants were students who were either residents of Tilburg, a city in the Netherlands, Maastricht, another city in the Netherlands, or Gothenburg, a city in Sweden. All participants were to claim during interviews that they were residents of Tilburg. In addition to a number of other questions, participants were asked to sketch the main square in Tilburg (see Appendix 1A)—which only the students of Tilburg were expected to be familiar with. For an example of a drawing of the square by a resident of Tilburg (i.e., a truth teller) see Appendix 1B. For an example of a drawing of the square by a resident of Gothenburg (i.e., a liar) see Appendix 1C. Coders, blind to the experimental conditions, counted the number of details included in the sketches of the square. The inter-rater agreement for this coding was excellent (intraclass correlation coefficient [ICC] = .99).

Finally, the study by Calderon et al. (2017) examined situations of true and false intent (for a review of this topic see Granhag & Mac Giolla, 2014). Truth tellers and liars were interviewed about their future intentions. Truth tellers provided honest descriptions of their intentions to carry out a shopping task. In order to conceal their intentions to perform a mock-crime, liars were instructed to provide a cover-story, that is, a false statement of intent. The cover-story was identical to the shopping task truth tellers were to perform. During the interview, participants were asked if they had experienced any mental image of the shopping task, and if so, to draw their most prominent mental image. A large number of independent coders ($N > 100$), blind to the experimental conditions, rated these images for level of abstractness on a 7-point scale, ranging from 1 (very concrete) to 7 (very abstract). Definitions of concrete and abstract were adapted from work on Construal Level Theory (Trope & Liberman, 2010). Specifically, drawings were rated as more concrete when they represented complex scenes of a specific context with a focus on surface aspects of the task. Drawings were rated as more abstract when they represented simple decontextualised scenes that focused on core aspects of a task. The inter-rater agreement for this coding was excellent ($ICC = .935$).

Main Findings

To reiterate, four out of five of the studies examined the drawings produced by truth tellers and liars in terms of overall detail. In the studies by Vrij et al. (2009), Vrij et al. (2012), and van Veldhuizen et al. (2017) truth tellers produced significantly more detailed drawings than liars. A plausible explanation for this is that truth tellers' memories included perceptual information that could readily be transformed into a pictorial drawing. In contrast, liars relied primarily on verbal lie-scripts, making them less cognitively flexible, reducing the amount of detail that they could provide in a pictorial drawing. However, Vrij et al. (2010) did not find significant differences between truth tellers' and liars' drawings in terms of detail (for an

overview of the standardised mean differences see Table 1). The explanation the authors gave was that liars based their lies around an experienced event. This countermeasure of embedding one's lies in true memories meant that liars had access to perceptual information, and hence, could provide drawings of a similar detail to truth tellers. This highlights a possible countermeasure to drawing-based lie detection techniques (this issue is discussed further in the Future Directions and Limitations section below).

Additional results further complicate the findings around drawing and detail. For instance, Vrij et al. (2009) found that although truth tellers produced more detailed drawings than liars, they also produced more detailed verbal statements than liars. In fact, the largest difference observed between truth tellers and liars in the study was based on verbal answers to questions on spatial aspects of the event. Although this provides support for the basic underlying mechanism that truth tellers have more perceptual information to base their answers on than liars, it speaks against drawings as being the only, or even the most effective way to elicit such differences.

We turn now to the results of the other qualitative measures that have been examined. In two studies (Vrij et al., 2010; Vrij et al., 2012), the drawings from truth tellers and liars, and their corresponding verbal statements, were rated in terms of plausibility. In both studies, truth tellers' drawings were rated as significantly more plausible than liars' drawings. In contrast, for truth tellers' and liars' verbal descriptions, no discernable differences in terms of plausibility were observed. Although these results are promising, the interrater reliability of the plausibility measure was modest across the two studies, which is likely to be due to the highly subjective nature in which plausibility was measured (for further discussion on this issue see the Future Directions and Limitations section below). However, other qualitative aspects of drawings, which were considerably more reliable to code, also showed notable differences between truth tellers and liars. For instance, Vrij et al. (2012) found that truth tellers included more people in their

drawings, and drew the people they included in more detail, compared to liars. Vrij et al. (2010) also found that truth tellers were more likely to include bystanders in their drawings than liars. In addition, they found that truth tellers were more likely than liars to use an over-the-shoulder camera perspective (as opposed to an overhead camera perspective) in their drawings. These latter findings are of particular note since, as noted above, Vrij et al. (2010) did not find differences in a more global measure of detail. In other words, in some situations, the type of details might be more important than the number of details.

The study on true and false intentions by Calderon et al. (2017) provides another example of the importance of focusing on types of details. To reiterate, during this study truth tellers and liars were required to provide drawings of the most prominent mental image related to their stated future intention (a shopping task). These images were coded on a 7-point scale ranging from 1 (very concrete) to 7 (very abstract). This coding was chosen because of research demonstrating that future events which are more likely to occur are represented more concretely compared to future events that are less likely to occur (Wakslak, Trope, Liberman, & Alony, 2006). Therefore, since false intentions are by definition less likely to occur, they should be represented more abstractly than true intentions. In line with this reasoning, liars' drawings were rated as more abstract than truth tellers' drawings (though it should be noted that this difference was quite modest in size, see Table 1).

Taken together, the five studies examining the differences in detail and quality between truth tellers' and liars' drawings are promising. Importantly, these findings are largely independent of the interviewees' drawing ability or skill. This is because it is not the excellence of the drawing that aids lie detection. Rather, it is aspects such as the spatial layout (e.g., of aspects in relation to one another) or the elements included within the drawing (e.g. other people or CCTV cameras) that differentiate between truth tellers and liars. Additionally, drawings can be

used in conjunction with other lie detection methods, such as the verifiability approach, whereby details in the drawing can be checked (see Nahari, Vrij, & Fisher, 2014), and so the excellence of the drawing becomes further irrelevant. We turn now to consistency cues, which are perhaps even more important for drawing-based deception detection techniques.

(insert Table 1 about here)

Drawings, Deception, and Consistency

Consistency comes in many different forms. For example, the information provided in a single statement can be examined for *within-statement consistency*. Alternatively, in situations when someone is interviewed more than once, the overlap of their multiple statements can be compared for *between-statement consistency*. Similarly, in situations where there are groups of suspects or witnesses, it is possible to compare the overlap of statements taken from different people to measure *within-group consistency* (Vredeveldt, van Koppen, & Granhag, 2014). Regardless of the type of consistency, laymen and professionals alike, appear to be in agreement that liars are typically more inconsistent than truth tellers (Akehurst, Köhnken, Vrij, & Bull, 1996; Greuel, 1992; Strömwall & Granhag, 2003; The Global Deception Research Team, 2006). However, despite these commonly held beliefs, research demonstrates that liars can be as, or even more consistent than truth tellers with regards to both between-statement consistency (Granhag, Mac Giolla, Sooniste, Strömwall, & Liu-Jönsson, 2016; Granhag & Strömwall, 1999) and within-group consistency (Granhag, Strömwall, & Jonsson, 2003). These results can be explained by the repeat vs reconstruct hypothesis (Granhag & Strömwall, 1999, 2001). In brief, this position holds that liars, in an explicit attempt to maintain consistency will repeat rehearsed answers in different interviews and across time. In contrast, truth tellers who are expected to take their innocence for

granted (Jordan & Hartwig, 2013), will simply recall unrehearsed answers from memory. Importantly, this process of spontaneous recollection can lead to memory errors, such as omissions and commissions resulting in natural inconsistencies in truth tellers' statements (Baddeley, 1997; Schacter, 1999). The differing interview strategies of truth tellers and liars result in similar levels of statement consistency. However, liars' repeat strategy is dependent on correctly anticipating the types of questions that interviewers will ask. Drawing-based deception detection techniques are designed to disrupt liars' repeat strategy by changing the mode of recall from the anticipated verbal recall to the unanticipated pictorial recall.

Drawings and Between-Statement Consistency

Leins et al. (2011) examined the potential for drawings to elicit inconsistencies from liars over repeated interview situations. In two experiments, the authors examined the between-statement consistency of truth tellers and liars when the modality of reporting (verbal vs. pictorial) was varied between interviews. The first experiment used data from the study by Vrij et al. (2009) mentioned above. Pairs of truth tellers had lunch together, whilst pairs of liars imagined doing so. All participants were then asked to provide a sketch of the restaurant. In a subsequent interview, participants were asked questions regarding the spatial layout of the restaurant. Coders, blind to the experimental conditions, rated the consistency (overlap) between the sketch and the verbal answers on a 7-point Likert scale (1 = not at all consistent; 7 = completely consistent). The inter-rater agreement between the coders was modest, albeit acceptable ($r = .71$). As predicted, truth tellers provided more consistent answers than liars. Based on the measure of consistency, a discriminant analysis correctly classified 80% of truth tellers and 70% of liars.

Leins et al.'s (2011) second experiment differed from the first in terms of (1) the task (truth tellers were asked to perform five simple tasks in a room, while liars were asked to imagine

performing these tasks), (2) the report mode order (participants provided a verbal description of the room in a first interview, and a sketch of the room in the second interview), and (3) the measurement of consistency (a more fine-grained, object level, measure of consistency was used instead of a Likert scale). With regards to the latter point, coders, blind to the experimental conditions, dichotomously rated all responses within statements as either consistent or inconsistent. For example, if a participant stated that ‘the blocks were to the right of the suitcase’ and subsequently drew the blocks to the right of the suitcase, this would be rated as consistent. The dependent variable was overall consistency, calculated by dividing the consistent responses by the total number of responses (i.e., consistent + inconsistent responses). Again, the inter-rater agreement between the coders was modest, but acceptable ($r = .70$). Despite the methodological differences from the first experiment, the results of the second experiment replicated the initial findings. Truth tellers were again rated as more consistent than liars. Based on the consistency measure, a discriminant analysis correctly classified 100% of truth tellers and 77% of liars.

Leins et al. (2011) offered two explanations for the results. First, liars may have been less consistent because they found the drawing task inherently more difficult than truth tellers (cf. cognitive load approach, Vrij, Fisher, et al., 2008). Alternatively, liars, due to a lack of cognitive flexibility, may have found it more difficult than truth tellers to switch between different recall modalities. In a subsequent study, Leins, Fisher, and Vrij (2012) addressed the respective merits of these two explanations. In this study, truth tellers performed five tasks, while liars imagined doing so. All participants were interviewed twice. Half of the participants were requested to provide their answers in the same modality across interviews (verbally and then again verbally, or pictorially and then again pictorially). The other half provided their answers in different modalities across interviews (verbally first, then pictorially, or pictorially first, then verbally). Coders, blind to the experimental conditions, rated participants’ answers across interviews for

consistency. A fine-grained measure of consistency, similar to the second experiment of Leins et al. (2011), was used. That is, the dependent variable was overall consistency, calculated by dividing the consistent responses by the total number of responses. The inter-rater agreement was excellent ($r = .95$). The main results showed that when modality did not change across interviews, liars achieved similar levels of consistency as truth tellers. In other words, the unanticipated nature of providing a pictorial drawing was not sufficient to disrupt a liars' repeat strategy. In contrast, liars were significantly less consistent than truth tellers when modality differed across interviews. This result supports the cognitive flexibility explanation.

Drawings and Within-Group Consistency

Two studies have examined the influence of drawing-based deception techniques on the within-group consistency of statements provided by groups of truth tellers and liars. The first was the aforementioned study by Vrij et al. (2009) that examined pairs of truth tellers who described a lunch date, and pairs of liars who described an imagined lunch date. In addition to the analyses mentioned above, within-group consistency was also measured. Specifically, coders, blind to the experimental conditions, rated on 7-point scales how much the drawings provided by each pair overlapped (1 = low correspondence, 7 = high correspondence). The inter-rater reliability between the coders was good ($r = .82$). Results showed that pairs of truth tellers provided considerably more consistent drawings than pairs of liars. Furthermore, there was no difference in within-group consistency between pairs of truth tellers and pairs of liars in the verbal statements they provided to the open questions (e.g., *Tell me in as much detail what you did at the restaurant*).

The second study examining drawing and within-group consistency was conducted by Roos af Hjelmsäter, Öhman, Granhag, and Vrij (2014). In this study, adolescents (aged 13-14 years), in groups of three, met with an adult stranger or imagined a meeting with an adult

stranger. Participants provided a verbal description of the event as well as a visual description. The visual description consisted of marking aspects of the event on a sketch of the area. This included marking on the sketch where they themselves stood, where the stranger stood, and where their group members stood. The verbal and visual descriptions were independently rated by a large group of coders ($N = 200$), all of whom were blind to the experimental conditions. A measure of inter-rater reliability was not reported. The main results showed that groups of truth tellers and groups of liars had similar levels of within-group consistency for the general description of the event. However, truth tellers showed considerably higher levels of within-group consistency compared to liars with regards to the drawing task.

The studies examining drawing-based deception detection techniques and consistency offer further support for this lie detection tool. A glance at Table 2 shows effect sizes that are among the highest in the published literature on cues to deceit (DePaulo et al., 2003). The review thus far has examined laboratory-based experimental studies. We turn next to training studies on the topic.

(insert Table 2 about here)

Training Studies

Training studies are crucial in bridging the gap from the laboratory to the field. These studies typically involve (1) training practitioners in techniques developed in the lab, and (2) examining these techniques in controlled environments by assessing their efficacy and practitioners' adherence to them. As of yet, no training study has exclusively examined drawing-based deception techniques. However, two training studies have examined a package of strategic interviewing techniques, where drawing was included as one of the ten techniques taught (Vrij,

Leal, Mann, Vernham, & Brankaert, 2015; Vrij, Mann, Leal, Vernham, & Vaughan, 2016). In both of the studies, half of the practitioners (experienced police officers) received one half day of training on strategic interviewing techniques. Ten specific techniques were taught, including the reverse order technique (Vrij, Mann, et al., 2008); the use of a model statement (Leal, Vrij, Warmelink, Vernham, & Fisher, 2015); asking unanticipated questions (Vrij et al., 2009); and asking the suspect to provide drawings or sketches during the interview. In subsequent mock-suspect interviews, untrained practitioners interviewed in any way they wished whereas trained practitioners were encouraged to use the techniques they had been taught. Half of the mock-suspects lied during the interview while the other half told the truth. The two studies used equivalent set-ups, but differed in two important ways. First, Vrij et al. (2015) used a repeated-measures, pre-training post-training design, whereas Vrij et al. (2016) used an independent-measures design. Second, in Vrij et al. (2015), the authors of the paper trained the practitioners, while in Vrij et al. (2016) an experienced police officer, who was initially trained in the techniques, taught the practitioners.

In terms of effectiveness, results from the two studies were ambiguous. Although the training package increased accuracy of deception judgements in the study by Vrij et al. (2015), it did not do so in the study by Vrij et al. (2016). Furthermore, even when the training package was effective in increasing accuracy rates, it was not possible to discern the effectiveness of any individual technique. This was because the practitioners were taught the different strategic interviewing techniques as a package. As such, the studies say little about the individual effectiveness of drawing-based deception detection techniques.

The results do, however, provide insights into other aspects of practical importance. For example, out of the ten techniques taught, asking the mock-suspects to provide drawings during the interview was the most commonly employed strategic interviewing technique in both studies.

Furthermore, none of the untrained participants used the drawing approach. This suggests that drawing-based deception detection techniques are not just common sense (cf. Masip, Herrero, Garrido, & Barba, 2011). Although the practitioners' opinions on the techniques were not collected, the frequency of use suggests a preference for the drawing approach compared to the other nine techniques. A plausible explanation for this preference is that drawing-based deception detection techniques are both easy to understand and to implement in the real-world. This idea is further supported in a recent study examining police officers' perceptions of deception cues based on different measures of consistency (Deep, Vrij, et al., 2017). Amongst other questions, police officers were asked if they had ever attempted to determine the veracity of a suspect by asking him/her to provide drawings during a suspect interview. Forty-eight (approximately 68%) of the 71 officers in the study reported having used drawings at least once in this manner. Of these 48 officers, 34 focused on some form of consistency cue to determine veracity. Taken together, these studies suggest that practitioners find drawing based deception techniques relatively easy to understand and implement. This is noteworthy since understanding and ease of use are central in translating scientific findings into practice (Rogers, 2010).

Future Directions and Limitations

We envision that future research on drawing-based deception detection techniques can advance the field at both a theoretical and a practical level. At the theoretical level, more can be done to make use of basic psychological research in fields such as spatial processing and spatial memory. For example, this research shows that spatial orientation is processed from one of two primary reference points. Locations can be framed with respect to one's own position (egocentric framing) or with respect to external objects or landmarks (alocentric framing; Postma et al., 2008). The study by van Veldhuizen et al. (2017) examined participants drawings of the main square in Tilburg. In this study, it is likely that participants relied primarily on alocentric framing.

In contrast, in the study by Vrij et al. (2009), participants sketched a lunch date. In this situation, it is perhaps more likely that participants relied on egocentric framing. By taking such differences into account it may be possible to further sharpen drawing-based deception detection techniques. The allocentric-egocentric distinction is just one example of how basic research can further this area of enquiry. At worst, such basic research can provide a more nuanced language to discuss drawing-based deception detection techniques. At best, it can inform future research on the topic and increase our understanding of this promising lie detection tool.

From a practical perspective, a more pressing issue is the effectiveness of countermeasures. A first step is simply to gauge the effectiveness of drawing-based deception detection techniques when liars have been made aware of the technique prior to an interview. A second step is to examine specific countermeasures. One such strategy of particular importance is the use of embedded lies. Embedded lies refer to those largely true statements where crucial details have been changed or omitted. Such lies are common and are particularly difficult to uncover (Leins, Fisher, & Ross, 2013; Vrij, 2008). Since they are based on mostly true memories, drawings may be ineffective at uncovering them. This is because embedded lies should be accompanied by the same visual and spatial information that inform truth tellers' statements and drawings.

A recent study has examined such questions and demonstrates the importance of examining countermeasures (Deep, Granhag, et al., 2017). In this study, truth tellers bought lunch from a café, whereas liars performed a mock-crime and afterwards also bought lunch from a café. Truth tellers and liars were questioned about their whereabouts in a subsequent suspect-styled interview. Liars were to conceal their involvement in the mock-crime by using the café as an alibi. In other words, since they had in fact been to the café, liars provided embedded lies. In addition, half of the liars were informed before they visited the café that they may be asked to

draw the layout of the café during an interview (*informed liars*). The other half did not receive this information before visiting the café (*uninformed liars*). During the interviews, all participants were required to provide drawings of the layout of the café. Two primary findings were observed. First, truth tellers and the uninformed liars provided highly similar drawings in terms of details and consistency. This result highlights the difficulties in uncovering embedded lies. Second, the informed liars provided significantly more detailed drawings than both the truth tellers and the uninformed liars. From the perspective of the lie-catcher, this finding is more positive as it highlights the difficulties in correctly deploying countermeasures. More research of this kind is needed to fully understand the boundary conditions of this technique.

Another important practical issue is the reliability of the measurements of the different deception cues. Simply put, cues that can be measured with more objectivity and reliability will be most relevant for practice. One way in which reliability is reduced is in how the cue is operationalised and defined. Ambiguous and poorly defined cues will typically have lower reliability than clearly defined ones. *Plausibility*, for example, is arguably one of the more ambiguous cues discussed in this review. The ambiguity of this cue is likely a strong contributing factor to the low inter-rater reliability observed in its measurement. This can be contrasted with an easily defined cue, such as the presence or absence of bystanders in a drawing. This cue showed a near perfect inter-rater reliability. Measurement, rather than definitional issues, can also affect the reliability of a cue. For example, within deception detection research a distinction is often made between (1) measurements that avail of Likert-type scale ratings and (2) more fine-grained measurements that count the frequency of the occurrence of criteria in raw scores. Recent research indicates the frequency count method to be more reliable than the scale rating method (Nahari, 2016). A glance at the reliability of the measurements covered in this review concurs with Nahari's (2016) results. That is, the inter-rater reliabilities were typically lower, and

fluctuated more, for measurements based on scale ratings compared to measurements based on a frequency count method. In order to increase the practical value of findings, future research should focus on cues that have a high degree of measurement reliability. Based on the above review, this will likely be cues that are clearly and easily defined and are measured on a more fine-grained frequency count level.

Practical Relevance and Legal Implications

Although this is a new line of research, we believe that drawing as a lie detection method has high practical relevance. In brief, it is a method that is easy to use, it comes with little costs, it can be used in tandem with other techniques, and it overcomes language barriers. Like most other lie detection methods, we believe that drawings will have most value during the investigative phase. Furthermore, we believe that the method should be used as a complement, not as an alternative, to more traditional and established lie detection methods, such as the strategic use of evidence (Hartwig, et al., 2006), the verifiability approach (Nahari et al., 2014), and the imposing cognitive load approach (Vrij et al., 2008). The outcome of a drawing-task should not be used as a stop-rule for deciding on the veracity of a statement. The outcome should be seen as a possible pointer for more in-depth questioning. Finally, the use of drawings is not limited to traditional police interviews, the method could be used (and is in fact sometimes used) during, for example, asylum interviews and interviews with sources and informants in human intelligence contexts.

Conclusions

The use of drawings as a lie detection technique has significant promise. The studies mentioned within this review demonstrate that when used appropriately (e.g. alongside other recall modalities) and when the correct cues to deceit are measured and interpreted correctly (e.g. consistency) then the differences between truth tellers and liars become more pronounced. However, our understanding of drawing as a lie detection tool is still in need of further research

and development. Until more studies (1) examine the impact of countermeasures and (2) are conducted and tested in practice, we will not be able to conclude about the true effectiveness of incorporating drawings into real-life investigative interviews. With that said, given that accuracy rates for detecting deception seem to improve with the implementation of drawings and that practitioners seem to be able to implement drawings into practice with ease and confidence, drawing-based deception detection techniques are certainly worthy of future deception detection research.

References

- Akehurst, L., Köhnken, G., Vrij, A., & Bull, R. (1996). Lay persons' and police officers' beliefs regarding deceptive behaviour. *Applied Cognitive Psychology, 10*(6), 461-471. doi: 10.1002/(sici)1099-0720(199612)10:6<461::aid-acp413>3.0.co;2-2
- Baddeley, A. D. (1997). *Human memory: Theory and practice* (revised ed.). Hove, UK: Psychology Press.
- Borenstein, M., Hedges, L. V., Higgins, J., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester: Wiley
- Calderon, S., Mac Giolla, E., Ask, K., & Granhag, P. A. (2017). *Drawing what lies ahead: False intentions are more abstractly depicted than true intentions*. Manuscript submitted for publication.
- Colwell, K., Hiscock-Anisman, C. K., Memon, A., Taylor, L., & Prewett, J. (2007). Assessment Criteria Indicative of Deception (ACID): An integrated system of investigative interviewing and detecting deception. *Journal of Investigative Psychology and Offender Profiling, 4*(3), 167-180. doi: 10.1002/jip.73
- Deep, H., Granhag, P. A., Vrij, A., Strömwall, L. A., Hope, L., & Mann, S. (2017). *Visuospatial countermeasures by liars familiar with the alibi setting*. Manuscript submitted for publication.
- Deep, H., Vrij, A., Hope, L., Mann, S., Granhag, P. A., & Strömwall, L. A. (2017). *Police officers' perceptions of statement inconsistency*. Manuscript submitted for publication.
- DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129*(1), 74-118. doi: 10.1037/0033-2909.129.1.74

- Fisher, R. P., Schreiber Compo, N., Rivard, J., & Hirn, D. (2014). Interviewing witnesses. In T. J. Perfect & D. S. Lindsay (eds) *The SAGE handbook of applied memory* (559-578). London: SAGE.
- Granhag, P. A., & Mac Giolla, E. (2014). Preventing future crimes: Identifying markers of true and false intent. *European Psychologist, 19*(3), 195-206. doi: 10.1027/1016-9040/a000202
- Granhag, P. A., Mac Giolla, E., Sooniste, T., Strömwall, L. A., & Liu-Jönsson, M. (2016). Discriminating between statements of true and false intent: The impact of repeated interviews and strategic questioning. *Journal of Applied Security Research 11*(1), 1-17. doi: 10.1080/19361610.2016.1104230
- Granhag, P. A., & Strömwall, L. A. (1999). Repeated interrogations—stretching the deception detection paradigm. *Expert Evidence, 7*(3), 163-174. doi: 10.1023/A:1008993326434
- Granhag, P. A., & Strömwall, L. A. (2001). Deception detection: Examining the consistency heuristic. In C. M. Bruer, M. M. Kommer, J. F. Nijboer & J. M. Reijntjes (Eds.), *New trends in criminal investigation and evidence* (Vol. 2, pp. 309-321). Antwerp, Belgium: Inresentia.
- Granhag, P. A., Strömwall, L. A., & Jonsson, A.-C. (2003). Partners in crime: How liars in collusion betray themselves. *Journal of Applied Social Psychology, 33*(4), 848-868. doi: <http://dx.doi.org/10.1111/j.1559-1816.2003.tb01928.x>
- Greuel, L. (1992). Police officers' beliefs about cues associated with deception in rape cases. In F. Lösel, D. Bender & T. Blieseneer (Eds.), *Psychology and law: International perspectives* (pp. 234-239). Berlin: Walter de Gruyter.

- Hartwig, M., Granhag, P. A., Strömwall, L. A., & Kronkvist, O. (2006). Strategic use of evidence during police interviews: When training to detect deception works. *Law and Human Behavior, 30*(5), 603-619. doi: 10.1007/s10979-006-9053-9
- Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General, 108*(3), 356-358. doi: 10.1037/0096-3445.108.3.356
- Hines, A., Colwell, K., Hiscock-Anisman, C., Garrett, E., Ansarra, R., & Montalvo, L. (2010). Impression management strategies of deceivers and honest reporters in an investigative interview. *The European Journal of Psychology Applied to Legal Context, 2*(1), 73-90.
- 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*(4), 371-376. doi: 10.1037/0096-3445.117.4.371
- Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review, 88*(1), 67-85. doi: 10.1037/0033-295x.88.1.67
- Jordan, S., & Hartwig, M. (2013). On the phenomenology of innocence: The role of belief in a just world. *Psychiatry, Psychology and Law, 20*(5), 749-760. doi: 10.1080/13218719.2012.730903
- Köhnken, G. (1989). Behavioral correlates of statement credibility: Theories, paradigms, and results. In H. Wegener, F. Lösel & J. Haisch (Eds.), *Criminal behavior and the justice system* (pp. 271-289). London: Springer.
- Leal, S., Vrij, A., Warmelink, L., Vernham, Z., & Fisher, R. P. (2015). You cannot hide your telephone lies: Providing a model statement as an aid to detect deception in insurance telephone calls. *Legal and Criminological Psychology, 20*(1), 129-146. doi: 10.1111/lcrp.12017

- Leins, D. A., Fisher, R. P., & Ross, S. J. (2013). Exploring liars' strategies for creating deceptive reports. *Legal and Criminological Psychology, 18*(1), 141-151. doi: 10.1111/j.2044-8333.2011.02041.x
- Leins, D. A., Fisher, R. P., & Vrij, A. (2012). Drawing on liars' lack of cognitive flexibility: Detecting deception through varying report modes. *Applied Cognitive Psychology, 26*(4), 601-607. doi: 10.1002/acp.2837
- Leins, D. A., Fisher, R. P., Vrij, A., Leal, S., & Mann, S. (2011). Using sketch drawing to induce inconsistency in liars. *Legal and Criminological Psychology, 16*(2), 253-265. doi: 10.1348/135532510x501775
- Light, L. L., & Zelinski, E. M. (1983). Memory for spatial information in young and old adults. *Developmental Psychology, 19*(6), 901-906. doi: 10.1037/0012-1649.19.6.901
- Masip, J., Herrero, C., Garrido, E., & Barba, A. (2011). Is the behaviour analysis interview just common sense? *Applied Cognitive Psychology, 25*(4), 593-604. doi: 10.1002/acp.1728
- 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*(1), 99-122. doi: 10.1080/10683160410001726356
- Mulligan, N. W. (2014). Memory for pictures and actions. In T. J. Perfect & D. S. Lindsay (Eds.), *The Sage handbook of applied memory* (pp. 20-36). London: Sage Publications Ltd.
- Nahari, G. (2016). When the long road is the shortcut: A comparison between two coding methods for content-based lie-detection tools. *Psychology, Crime & Law, 22*(10), 1000-1014. doi: 10.1080/1068316x.2016.1207770
- Nahari, G., Vrij, A., & Fisher, R.P. (2014). Exploiting liars' verbal strategies by examining the verifiability of details. *Legal and Criminological Psychology, 19*, 227-239. doi: 10.1111/j.2044-8333.2012.02069.x

- Naveh-Benjamin, M. (1987). Coding of spatial location information: An automatic process? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(4), 595-605. doi: 10.1037/0278-7393.13.4.595
- Otgaar, H., van Ansem, R., Pauw, C., & Horselenberg, R. (2016). Improving children's interviewing methods? The effects of drawing and practice on children's memories for an event. *Journal of Police and Criminal Psychology*, *31*(4), 279-287. doi: 10.1007/s11896-016-9190-0
- Postma, A., Kessels, R. P. C., & van Asselen, M. (2008). How the brain remembers and forgets where things are: The neurocognition of object–location memory. *Neuroscience & Biobehavioral Reviews*, *32*(8), 1339-1345. doi: <http://dx.doi.org/10.1016/j.neubiorev.2008.05.001>
- Pouliot, S., & Gagnon, S. (2005). Is egocentric space automatically encoded? *Acta Psychologica*, *118*(3), 193-210. doi: 10.1016/j.actpsy.2004.10.016
- Rogers, E. M. (2010). *Diffusion of innovations*. New York: Simon and Schuster.
- Roos af Hjelmsäter, E., Öhman, L., Granhag, P. A., & Vrij, A. (2014). 'Mapping' deception in adolescents: Eliciting cues to deceit through an unanticipated spatial drawing task. *Legal and Criminological Psychology*, *19*(1), 179-188. doi: 10.1111/j.2044-8333.2012.02068.x
- Schacter, D. L. (1999). The seven sins of memory: Insights from psychology and cognitive neuroscience. *American Psychologist*, *54*(3), 182-203. doi: 10.1037/0003-066x.54.3.182
- Strömwall, L. A., & Granhag, P. A. (2003). How to detect deception? Arresting the beliefs of police officers, prosecutors and judges. *Psychology, Crime & Law*, *9*(1), 19-36. doi: 10.1080/10683160308138
- The Global Deception Research Team. (2006). A world of lies. *Journal of Cross-Cultural Psychology*, *37*(1), 60-74. doi: 10.1177/0022022105282295

- Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review*, *117*, 440-463. doi: 10.1037/a0018963
- van Veldhuizen, T. S., Horselenberg, R., Stel, M., Landström, S., Granhag, P. A., & van Koppen, P. J. (2017). The Provenance of Émigrés: The Validity of Measuring Knowledge of Places. *Psychology, Crime & Law*, Advance Online Publication.
- Vredeveltdt, A., van Koppen, P. J., & Granhag, P. A. (2014). The inconsistent suspect: A systematic review of different types of consistency in truth tellers and liars. In R. Bull (Ed.), *Investigative Interviewing* (pp. 183-207). New York: Springer.
- Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities (2nd ed.)*. New York, NY: John Wiley & Sons Ltd.
- Vrij, A., Fisher, R. P., Mann, S., & Leal, S. (2008). A cognitive load approach to lie detection. *Journal of Investigative Psychology and Offender Profiling*, *5*(1-2), 39-43. doi: 10.1002/jip.82
- 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), 110-117. doi: <http://dx.doi.org/10.1016/j.jarmac.2012.02.004>
- 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*(2), 159-166. doi: 10.1007/s10979-008-9143-y
- Vrij, A., Leal, S., Mann, S., Warmelink, L., Granhag, P. A., & Fisher, R. P. (2010). Drawings as an innovative and successful lie detection tool. *Applied Cognitive Psychology*, *24*(4), 587-594. doi: 10.1002/acp.1627
- Vrij, A., Leal, S., Mann, S., Vernham, Z., & Brankaert, F. (2015). Translating theory into practice: Evaluating a cognitive lie detection training workshop. *Journal of Applied*

Research in Memory and Cognition, 4(2), 110-120. doi:

<http://dx.doi.org/10.1016/j.jarmac.2015.02.002>

Wakslak, C. J., Trope, Y., Liberman, N., & Alony, R. (2006). Seeing the forest when entry is unlikely: Probability and the mental representation of events. *Journal of Experimental Psychology: General*, 135(4), 641-653. doi: 10.1037/0096-3445.135.4.641

Vrij, A., Mann, S., Leal, S., & Fisher, R. (2012). Is anyone there? Drawings as a tool to detect deceit in occupation interviews. *Psychology, Crime & Law*, 18(4), 377-388. doi: [10.1080/1068316X.2010.498422](https://doi.org/10.1080/1068316X.2010.498422)

Vrij, A., Mann, S., Leal, S., Vernham, Z., & Vaughan, M. (2016). Train the trainers: A first step towards a science-based cognitive lie detection training workshop delivered by a practitioner. *Journal of Investigative Psychology and Offender Profiling*, 13(2), 110-130. doi: 10.1002/jip.1443

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 an event in reverse order. [journal article]. *Law and Human Behavior*, 32(3), 253-265. doi: 10.1007/s10979-007-9103-y

Table 1. Effect sizes (Cohen's d) for the differences in truth tellers and liars drawings on cues of detail and quality. Positive values indicate higher scores for truth tellers.

Que	Study	Vrij et al. (2009)	Vrij et al. (2010)	Vrij et al. (2012)	Van Veldhuizen et al. (2017)	Calderon et al. (2017)
Detail		0.48*	0.25	0.79*	2.60*	-
Plausibility		-	0.93*	0.96*	-	-
No. of Bystanders		-	1.84 * [†]	0.59*	-	-
Detail of Bystanders		-	-	1.21*	-	-
Camera Perspective		-	0.88 * [†]	-	-	-
Level of Concreteness		-	-	-	-	0.40*

* $p < .05$

[†]Odds ratios were calculated and converted to d values according to Borenstein, Hedges, Higgins, and Rothstein (2009).

Table 2. Effect sizes (Cohen's d) for the differences in truth tellers and liars drawings for cues of consistency.

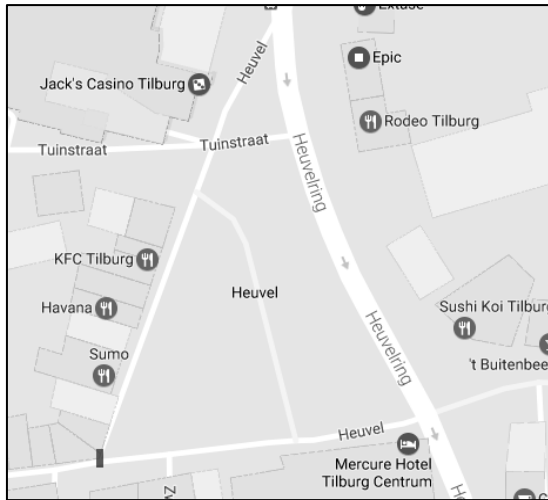
Positive values indicate more consistent answers for truth tellers.

Study	Vrij et al. (2009)	Leins et al. (2011. Exp 1.)	Leins et al. (2011. Exp 2.)	Leins et al. (2012)	Roos af Hejlsäter et al. (2014)
Que					
Between-statement consistency	-	1.25*	2.06*	1.13* [†]	-
Within-group Consistency	1.10*	-	-		1.29*

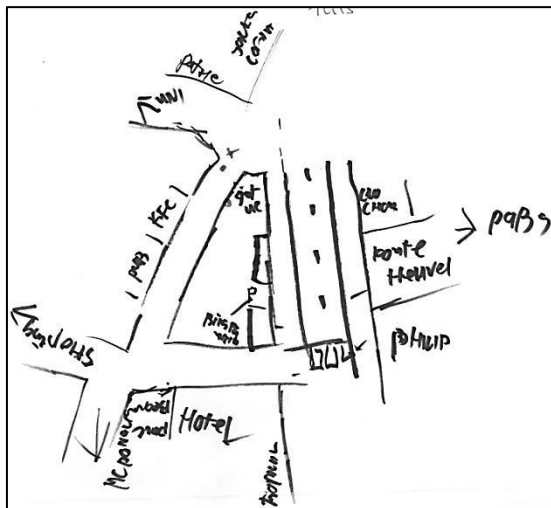
* $p < .05$

[†]This analysis is only based on the measures of between-statement consistency where the recall format (drawing vs. verbal) varied between interviews (i.e., the last four columns of Figure 1, in Leins et al. 2012). The data were provided by the authors of the paper. The pooled means and sds were used to calculate the d value.

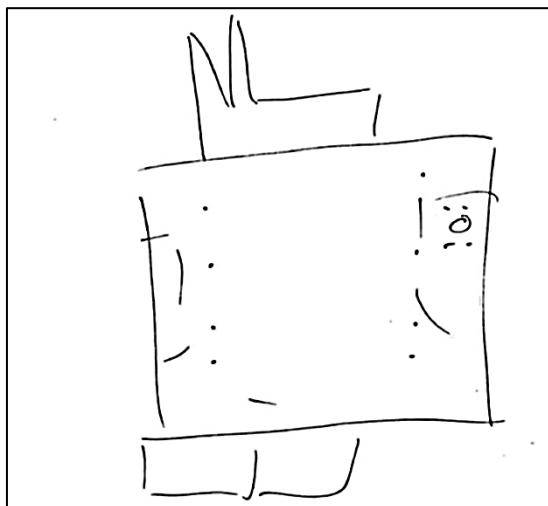
Appendix 1



1A. A map of the main square in Tilburg. The map is copied from Google Maps ©.



1B. Drawing of the main square in Tilburg by a truth teller (i.e., a resident of the city). Drawing provided by van Veldhuizen et al. 2017.



1C. Drawing of the main square in Tilburg by a liar (i.e., not a resident of the city). Drawing provided by van Veldhuizen et al. 2017.