

# Self-Reported Beliefs About Verbal Cues Correlate with Deception-Detection Performance

Citation for published version (APA):

Bogaard, G., & Meijer, E. H. (2018). Self-Reported Beliefs About Verbal Cues Correlate with Deception-Detection Performance. *Applied Cognitive Psychology*, 32(1), 129-137. <https://doi.org/10.1002/acp.3378>

## Document status and date:

Published: 01/01/2018

## DOI:

[10.1002/acp.3378](https://doi.org/10.1002/acp.3378)

## Document Version:

Publisher's PDF, also known as Version of record

## Document license:

Taverne

## Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

## Take down policy

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.

Download date: 25 Apr. 2024

## Self-Reported Beliefs About Verbal Cues Correlate with Deception-Detection Performance

GLYNIS BOGAARD\*  and EWOUT H. MEIJER

Forensic Psychology Section, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands

*Summary:* In this study, we investigated whether people who hold more correct beliefs about verbal cues to deception are also better lie detectors. We investigated police officers and undergraduates' beliefs about (i) cues to deception via an open-ended question and (ii) 17 specific verbal cues, after which participants were asked to judge the truthfulness of eight video fragments. Results showed that undergraduates and police officers still hold wrongful beliefs about nonverbal cues, but have better insight into verbal cues. Moreover, a better insight in verbal cues was related to an increased accuracy for identifying truthful statements, showing that verbal cues do drive credibility judgments to some extent. Copyright © 2017 John Wiley & Sons, Ltd.

### INTRODUCTION

People typically perform around chance level when trying to detect deception (Bond & DePaulo, 2006, 2008). Some authors have argued that this poor performance can be explained by the misconceptions people have about what cues expose liars (e.g., Strömwall, Granhag, & Hartwig, 2004). The Global Deception Research Team (2006), for example, investigated the beliefs about deception cues of 2320 participants from 58 different countries. Gaze aversion was most frequently mentioned; it was reported by 64% of all participants (also Bogaard, Meijer, Vrij, & Merckelbach, 2016). However, gaze aversion has been shown to be unrelated to deception (DePaulo et al., 2003).

Moreover, people believe far more nonverbal cues are associated with lying than actually is the case (Akehurst, Köhnken, Vrij, & Bull, 1996; DePaulo et al., 2003; Strömwall et al., 2004; Vrij, Akehurst, & Knight, 2006). A meta-analysis by DePaulo et al. (2003) showed that 118 of the 158 examined cues did not show any relationship with deception, and only three of the nonverbal cues showed a small relationship: liars were more tense, had larger pupils, and kept their chin slightly up. Still, people report to heavily rely on nonverbal cues when making credibility judgments. Mann, Vrij, and Bull (2004), for example, asked police officers to detect lies in video clips, and 78% of all the cues reported to make a deception judgment were nonverbal. More recently, Bogaard et al. (2016) asked police officers and undergraduates to list cues that they thought were diagnostic of deception. Eighty-six percent of all cues mentioned were nonverbal. In line with previous studies, cues such as 'gaze aversion' and 'nervous behavior' were mentioned most often.

Thus, people predominantly rely on nonverbal cues and largely ignore verbal cues when detecting deceit, even though the latter are more diagnostic (Levine & McCormack, 2014; Masip & Herrero, 2015; Vrij, 2008c). Nevertheless, which beliefs people hold about verbal cues are only investigated sporadically. The Global Deception Research Team (2006), for example, reported that people generally believe

liars' stories are longer and more inconsistent than are truth tellers' stories. Furthermore, police officers, prosecutors, and judges also believed that liars' stories are less detailed and less consistent (Strömwall & Granhag, 2003). Research investigating criminals' beliefs about deception cues reported that prison inmates primarily use plausibility as a cue to deception, while lay people mostly rely on inconsistencies (Granhag, Andersson, & Strömwall, 2004). Moreover, prison inmates believe that verbal cues are more diagnostic when detecting deceit, while prison personnel and students believed the opposite to be true (Hartwig, Granhag, Strömwall, & Andersson, 2004).

Unlike with nonverbal cues, beliefs about verbal cues often correspond to what is known from empirical research. Bogaard et al. (2016) recently questioned 95 police officers and 105 undergraduates about their beliefs on lie-detection cues in general and about 28 specific verbal cues. Most often, the usual stereotypical, but nondiagnostic, behavioral cues were mentioned (i.e., gaze aversion, movements, and fidgeting). Only 14% of the mentioned cues pertained to verbal cues. However, when specifically asked about verbal cues, participants were less inclined to overestimate the relationship with deception. Overall, both groups judged the relationship with deception correctly for the majority of the verbal cues, or judged them as is hypothesized in the deception literature.

The finding that people hold more correct beliefs about verbal cues to deception than about nonverbal cues has important implications. It might, for example, explain the increased deception-detection performance in situations where nonverbal cues are actively excluded. Burgoon, Blair, and Strom (2008) presented participants with either video or audio clips of truthful and fabricated statements. Accuracy in the audio condition (60%) was significantly higher than that in the audio-visual condition (47%). Similarly, Mann, Vrij, Fisher, and Robinson (2008) asked observers to evaluate the credibility of police-suspect interviews. Observers received the footage without audio (visual only), heard the suspects (audio only), or both heard and saw the suspects (audio-visual). With 65% accuracy, observers in the audio and audio-visual conditions outperformed the participants in the visual only condition, who achieved 44% accuracy. More recently, Leach et al. (2016) showed that observers were more accurate at assessing credibility in witnesses

\*Correspondence to: Glynis Bogaard, Forensic Psychology Section, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands.  
E-mail: glynis.bogaard@maastrichtuniversity.nl

whose faces were covered by either a head or face veil (e.g., hijabs or niqabs) compared with uncovered faces.

Still, having correct beliefs about cues to deception does not necessarily mean increased deception-detection accuracy. Using Brunswik's lens model, Hartwig and Bond (2011; also Hartwig & Granhag, 2014) showed large dissociations between the cues that people list when asked how they can detect deception, and what they base their veracity judgment on when actually trying to detect deception. Commonly self-reported cues such as gaze aversion and fidgeting contributed only little to veracity judgments, while actual verdicts were based on impressions of incompetence, impressions of ambivalence, and lack of spontaneity. Recently published data from Wright and Wheatcroft (2017) supported Hartwig and Bond's (2011) finding that self-reported cues were only minimally related to lie-detection accuracy.

However, most of the studies included in these meta-analyses focused on behavioral cues to deception and their relation to deception accuracy. Interestingly, older research (Mann et al., 2004) found that good lie detectors reported more verbal cues than nonverbal cues when making credibility judgments, a finding that fits comfortably with the scientific literature (for a review, see Vrij, 2008c). Previous studies have shown a relationship between the number of verbal cues used for credibility judgments and lie-detection accuracy, but little research has included information on the validity of these cues. Moreover, given the small relation between self-reported and objective behavioral cues and deception judgments, the accuracy of those deception judgments might be better explained by participants' insight in verbal cues.

Therefore, in this study, our main aim was to investigate whether people who hold more correct beliefs about verbal cues to deception are better lie detectors. Furthermore, our study examined two additional aims, which are identical to Bogaard et al. (2016), namely, (i) to investigate police officers and undergraduates' beliefs about cues to deception via an open-ended question and (ii) to investigate their beliefs about 17 verbal cues. Finally, we investigated to what extent presentation mode (transcript or video) would influence lie-detection accuracy.

## METHOD

### Participants

The sample consisted of 138 police officers [ $M_{\text{age}} = 34$ , standard deviation ( $SD$ ) = 11.57] and 108 undergraduate students ( $M_{\text{age}} = 21$ ,  $SD = 2.35$ ). The police officers were recruited by approaching as many police stations as possible, both by phone and by email. Officers reported a mean experience of 9 years ( $SD = 11.92$ ) ranging from 1 to 38 years. Undergraduates were mainly first-year and second-year psychology students, as they had not yet received any information on lie-detection or interviewing techniques in their curriculum. Undergraduates received credit points, whereas police officers did not receive any compensation. The study was approved by the ethical

committee of the Faculty of Psychology and Neuroscience of Maastricht University.

### Materials

#### Statements

To measure lie-detection accuracy, we provided all participants with eight statements. These statements were selected from 13 structured interviews. The interviewees were all acquaintances of either the experimenter or the research assistant. Six of the respondents were interviewed about an event they actually experienced. The ground truth of this statement was known; the interviewees filed a report, and multiple witnesses were available to support their story. The seven remaining interviewees were interviewed about a fabricated event and were asked to base this lie on a realistic scenario. Interviewees were told that their statements were going to be judged on their credibility by observers and were asked to convince the observers they were telling the truth. Thus, six statements were truthful, and seven statements were fabricated. The interview consisted of the following questions: 'You told me that you ... [different for every interviewee]. Can you describe in your own words what exactly happened?', 'What was your first reaction when this happened?', 'How has this incident affected you?', 'Have you talked to others about this incident?', and 'Is there anything that you would like to add to your statement?' Each interview was recorded audio visually and showed the interviewee sitting at a table.

Fifteen undergraduate psychology students rated the statements' credibility, sound quality, and image quality on a 7-point scale ranging from completely false to completely true. Two statements (one truthful and one fabricated) scored low on sound quality and were therefore excluded. We also excluded two statements (one truthful and one fabricated) because their credibility score was below 3 (too implausible) or more than 5 (too credible). To obtain an equal number of truthful and fabricated statements, we excluded one more fabricated statement, the one in which content was least similar to the truthful statements. Of the remaining statements, average credibility scores of the four truthful statements ( $M = 4.01$ ,  $SD = 1.84$ ) did not differ significantly from the four fabricated statements ( $M = 4.09$ ,  $SD = 2.21$ ). The truthful statements were stories about a car theft (victim), two traffic accidents (both as witnesses), and a carjacking (victim). The fabricated statements described two traffic accidents (one victim and one witness) and two statements on random violence (one victim and one witness). Each interview was fully transcribed, including mistakes, slips, and stop words. The statements had an average word count of 720 ( $SD = 121$ ).

#### The questionnaire

The questionnaire measuring beliefs about cues to deception was identical to that of Bogaard et al. (2016). Participants were first asked the open-ended question 'What do you think are good cues for detecting lies?' and given unlimited space to respond. Next, they were asked to indicate their opinion about 17 content cues that were derived from two existing verbal credibility assessment tools, namely, criteria-based

content analysis (CBCA; Steller & Köhnken, 1989) and reality monitoring (RM; Johnson & Raye, 1981). For RM, we included all eight items described by Sporer (1997), and for CBCA, we included 12 of the 19 items. Seven items were omitted because they have shown to be only rarely present in statements, meaning little is known about their validity. We also excluded items that pertain primarily to children's statements (Vrij, 2005). As there is an overlap between three items from both methods, we included these items only once. This resulted in a list of 17 separate items, which were presented in the order listed in Appendix A.

For every item, we gave a short description and an example. For example, for temporal information, we gave the following description: 'This cue refers to all descriptions about when the event took place and how long it took (e.g., "it was early in the morning" or "an hour after I got up")'. Next, participants were asked to indicate their opinion on forced-choice answer scales. Respondents could choose between two directed (e.g., 'this cue is used less by liars' or 'this cue is used more by liars'), one neutral (e.g., 'there is no difference between liars and truth tellers regarding this cue'), and a 'don't know' alternatives. The complete questionnaire can be found in Appendix B.

Next, participants were asked to rate the veracity of the eight statements on a 10-point Likert scale (1 = not credible to 10 = very credible). The statements were presented as either transcripts or as video fragments. Additionally, participants answered questions about their background, function, and experience and whether they ever had training in lie detection. Finally, participants were asked to indicate on a 7-point scale ranging from 1 (very poor) to 7 (very good) how well they knew the literature about lie detection.

The questionnaire was administered online via the *thesistools.com* or in a classroom as a pen and paper test. In case of the latter, the group consisted of 63 police officers that were all allocated to the video condition, with the videos being presented on a large screen in a lecture hall. Police officers were seated separately, so they could not discuss their answers with each other. Like in participants who signed up for the online questionnaire, participation was voluntarily. After all participants were finished with the questions for each video, the next video was shown on screen. All other participants were randomly allocated to either the video or transcript condition. Although 138 officers and 108 students originally started the study, three officers and five students did not complete the questionnaire, and their data were therefore removed from the sample. Of the remaining participants, 102 police officers were in the video condition and 33 in the transcript condition. For the students, 46 were in the video condition and 57 in the transcript condition.

## RESULTS

Five of the police officers reported having received a training in interrogation skills, which included some information on lie detection, but this training did not focus on verbal lie-detection cues. In response to the question 'how well do you know the literature about lie detection?', the police

officers reported they were not very knowledgeable about the literature ( $M = 2.51$ ,  $SD = 1.91$ ), which was similar to what students reported ( $M = 2.8$ ,  $SD = 2.00$ ) [ $t(244) = -1.04$ ,  $p = ns$ ].

### Open question

The responses to the open question ('What do you think are good cues for detecting lies?') were grouped by two raters in two different categories: verbal and nonverbal cues. Within the nonverbal category, the answers were grouped into more specific categories such as speech characteristic (e.g., response latency and voice pitch), facial behaviors (e.g., blushing and gaze aversion), and body movements (e.g., illustrators and moving feet) with a total of 48 different possibilities (Bogaard et al., 2016). Average percentage agreement between the two raters was 98.80%.

Undergraduates reported 319 nonverbal and 72 verbal signals. The five most mentioned signals were (i) gaze aversion (57.69%), (ii) nervous behavior (33.65%), (iii) bodily signals (27.88%), (iv) amount of details (20.19%), and (v) sweating (17.31%). Police officers reported a total of 530 nonverbal and 125 verbal signals. The five most mentioned signals were (i) gaze aversion (47.47%), (ii) sweating (33.55%), (iii) nervous behavior (31.61%), (iv) inconsistencies/contradictions (29.03%), and (v) bodily signals (25.81%).

### Closed questions

#### *Don't know answers*

Table 1 shows to what extent police officers and students endorsed the directions of the nonverbal cues. Before we

Table 1. Percentages of chosen alternatives for police officers and undergraduates

Item	Police officers				Undergraduates			
	-	0	+	?	-	0	+	?
Coherence	55.8	12.3	23.9	8	64.8	1.9	25	3.7
Clarity	33.3	13.8	46.4	6.5	41.7	14.8	36.1	2.8
Spontaneous corrections	39.1	8.7	46.4	5.8	35.2	3.7	50.9	5.6
Inconsistencies	13	7.2	75.4	4.3	15.7	2.9	70.4	3.7
Perceptual information	71	10.9	15.2	2.9	67.6	11.1	12	4.6
Emotions	65.9	15.9	12.3	5.8	65.7	11.1	11.1	7.4
Amount of details	51.4	11.6	33.3	3.6	47.2	7.4	38	2.8
Spatial information	56.5	16.7	18.8	8	54.6	21.3	13.9	4.6
Unstructured production	36.2	15.2	44.2	4.3	31.5	13.9	42.6	7.4
Description of interaction	56.5	17.4	15.9	10.1	56.5	18.5	9.3	11.1
Temporal details	46.4	21	21.7	10.9	37	32.4	21.3	4.6
Superfluous details	37	8	52.2	2.9	37	3.7	50.9	3.7
Reproduction of conversation	54.3	15.9	17.4	10.1	45.4	18.5	24.1	7.4
Reconstructability	51.4	13.8	23.2	8.7	46.3	9.3	29.6	10.2
Unusual details	58.7	4.3	29	5.1	56.5	7.4	25.9	5.6
Plausibility	41.3	29	18.1	8.7	40.7	25.9	16.7	12
Cognitive operations	58	11.6	21	6.5	60.2	13.9	14.8	6.5

Note: - indicates less for liars, 0 indicates no difference between liars and truth tellers, + indicates more for liars, and ? indicates a 'don't know' answer.

could investigate the directional answers to the closed questions, we first examined whether students and police officers applied the same decision criteria for choosing a directional answer. To this end, we investigated how often both groups chose the don't know alternative. On average, police officers chose the don't know answer 6.6% of the time, against 6.1% for students. To allow for  $\chi^2$  tests, the don't know answer was coded as '0', while all other answers were scored as '1'. One  $\chi^2$  test was performed per question, and none of the 17  $\chi^2$  tests were significant (all  $p > .26$ ). This means both groups did not differ in their decision criteria and did not significantly differ in their frequency of choosing the don't know answer.

#### Directional and neutral answers

Data analysis was identical to that of Bogaard et al. (2016) and Strömwall and Granhag (2003). We first recoded directional and neutral answers as  $-1$  (i.e., less when lying),  $0$  (i.e., no difference), and  $1$  (i.e., more when lying) and excluded the don't know alternative from the following analyses. Next, we analyzed the data with 17 one-sample sign tests (one for every item in the questionnaire) to investigate whether the average mean value of every item was significantly different from 0. This allowed us to investigate whether there was a preference for either one of the directional answers (i.e., more or less). Means and  $p$ -values for both groups are presented in Table 2. To correct for multiple testing, we have adjusted the  $\alpha$  level to .01. We also investigated whether police officers and students significantly differed in their beliefs about the verbal cues by means of multiple Mann-Whitney tests (for every item separately). Again, we adjusted the  $\alpha$  level to .01, but no differences appeared.

Table 2. Average mean value for every item ranging between  $-1$  (less for liars) and  $1$  (more for liars), indicating the beliefs about verbal cues of police officers and undergraduates

Items	Police officers			Undergraduates			Between-groups $p$ -value	Correct answer
	Mean	$p$ -value	Effect size <sup>a</sup>	Mean	$p$ -value	Effect size <sup>a</sup>		
Coherence	<b>-0.35<sup>b</sup></b>	<.01	0.74	<b>-0.43<sup>b</sup></b>	<.01	0.44	.12	<
Clarity	0.14 <sup>b</sup>	.11	0.14	-0.06 <sup>b</sup>	.59	0.06	.05	—
Spontaneous corrections	0.08	.41	0.08	0.18	.10	0.18	.47	<
Inconsistencies	<b>0.65<sup>b</sup></b>	<.01	0.66	<b>0.60<sup>b</sup></b>	<.01	0.60	.63	>
Perceptual information	<b>-0.57<sup>b</sup></b>	<.01	0.58	<b>-0.61<sup>b</sup></b>	<.01	0.62	.96	<
Emotions	<b>-0.57</b>	<.01	0.56	<b>-0.62</b>	<.01	0.62	.78	—
Amount of details	-0.19	.03	0.18	-0.10	.35	0.10	.60	<
Spatial information	<b>-0.41<sup>b</sup></b>	<.01	0.40	<b>-0.44<sup>b</sup></b>	<.01	0.44	.62	<
Unstructured production	0.08	.34	0.08	0.13	.22	0.12	.43	>
Description of interaction	<b>-0.45</b>	<.01	0.46	<b>-0.56</b>	<.01	0.56	.64	—
Temporal details	<b>-0.28<sup>b</sup></b>	<.01	0.28	-0.17	.04	0.18	.91	<
Superfluous details	0.16 <sup>b</sup>	.07	0.16	0.15 <sup>b</sup>	.15	0.16	.88	—
Reproduction of conversation	<b>-0.42</b>	<.01	0.42	<b>-0.24</b>	.01	0.24	.34	<
Reconstructability	<b>-0.32</b>	<.01	0.32	-0.20 <sup>b</sup>	.06	0.34	.32	—
Unusual details	<b>-0.32</b>	<.01	0.32	<b>-0.34</b>	<.01	0.34	.96	—
Plausibility	<b>-0.26</b>	.01	0.26	<b>-0.29</b>	<.01	0.28	.77	<
Cognitive operations	<b>-0.41</b>	<.01	0.40	<b>-0.51</b>	<.01	0.56	.50	—

#### Note:

<sup>a</sup>Minus sign indicates that the item is less present when lying. Significant deviations from 0 are in bold. To correct for multiple testing, we used an  $\alpha$  of .01. > indicates that the cues are more present in fabricated statements, — indicates that there is no relationship between the cue and deception, and < indicates that the cues are less present in fabricated statements.

<sup>b</sup>Participants judged this item correctly. Reported effect size is the  $\delta$ -index for stochastic difference (Vargha & Delaney, 2000). A  $\delta$ -value of 0.11 is considered a small effect, 0.28 a medium effect, and 0.43 a large effect.

#### Relationship with extant literature

Based on the assumptions of CBCA and RM, all items are expected to be more present in truthful than fabricated statements, but vice versa for the cognitive operations item. However, empirical evidence has shown that these assumptions are not always correct. Therefore, Table 2 summarizes which verbal cues are valid cues to deception based on the extant empirical literature. We derived this validity estimate as in Bogaard et al. (2016). The validity of RM items was established by following the results of DePaulo et al. (2003) and Masip, Sporer, Garido, and Herrero (2005). An RM item was scored as diagnostic if the item was significantly more present in truthful statements compared with fabricated statements (or vice versa for cognitive operations) in more than 65% of the included studies. CBCA items were scored on the basis of Amado, Arce, and Fariña (2015), DePaulo et al. (2003), and Vrij (Vrij, 2005, 2008b). A CBCA item was scored as diagnostic if the item was significantly more present in truthful statements compared with fabricated statements in more than 65% of the studies included in the meta-analyses of Vrij (2005, 2008b) and/or showed an effect size of at least  $d = 0.50$  in Amado et al. (2015) and/or a significant difference in DePaulo et al. (2003). In total, 10 cues were evaluated as diagnostic of deceit, while evidence for seven items was mixed and hence was evaluated as nondiagnostic.

Students evaluated six out of 10 diagnostic cues (60%) correctly, while police officers judged seven out of 10 cues (70%) correctly (indicated with 'b' in Table 2).

Students evaluated three out of seven nondiagnostic cues correctly (42%), while police officers judged two out of seven cues (29%) correctly (indicated with 'b' in Table 2). We also investigated to what extent our participants' beliefs

were in agreement with the assumptions of CBCA and RM. For four of these cues police officers followed the assumed direction of the CBCA and RM instruments (i.e., criteria are less present when lying), and students did so for three of these cues. This means that for police officers, 13 out of 17 items were judged correctly (nine items) or as assumed (four items) from the deception literature; for students, 12 out of 17 items (nine items correct and three items as assumed).

### Credibility judgments of statements

Participants rated the credibility of eight statements (four truthful and four fabricated), which were presented either as video clips or as transcripts. All credibility judgments were made on a 10-point scale ranging from 1 (completely fabricated) to 10 (completely truthful). We calculated the average credibility score for the truthful and fabricated statements separately.

Given that 63 police officers were tested *in vivo*, and 70 took the online questionnaire, we first checked whether questionnaire format influenced our results by conducting a  $2 \times 2$  repeated-measures analysis of variance (ANOVA) with questionnaire (online vs. *in vivo*) and presentation mode (transcript vs. video) as between-subject variables and veracity (truth vs. lie) as a within-subject variable. Results showed no significant interactions of questionnaire format with any of the other variables, and questionnaire format was therefore dropped from the design.

Next, we investigated how presentation mode and profession influenced our participants' credibility judgments. To this end, we conducted a  $2 \times 2 \times 2$  repeated-measures ANOVA with group (police officers vs. students) and presentation mode (transcript vs. video) as between-subject variables and veracity (truth vs. lie) as a within-subject variable. In addition, we conducted a Bayesian repeated-measures ANOVA with JASP and report the Bayesian factors (BF; for interpretation of BF, see Jarosz & Wiley, 2014; Lee & Wagenmakers, 2013). The interaction model with JASP always includes both main effects and interaction. Therefore, evidence for the interaction model is calculated as (interaction model)/(main factors) (Wagenmakers et al., 2016).

First, results showed a significant interaction of group and presentation mode [ $F(1, 232) = 8.37, p < .001, \eta_p^2 = 0.035$ ]. The accompanying Bayes factor shows the data are 6.35 times more likely under the interaction model than under the two main effects model. Police officers' credibility judgments for the transcripts were significantly higher ( $M = 6.93, SD = 1.01$ ) compared with their judgments on the videos ( $M = 5.82, SD = 1.49$ ), while this difference was not significant for undergraduates (transcript  $M = 6.43, SD = 1.30$  vs. video  $M = 6.17, SD = 1.33$ ). Furthermore, results showed that fabricated statements ( $M = 6.05, SD = 1.39$ ) were scored as more credible than true statements were ( $M = 6.31, SD = 1.43$ ) [ $F(1, 232) = 5.96, p = .02, \eta_p^2 = 0.025$ ]. Yet the Bayes factor showed only weak support ( $BF_{10} = 1.21$ ). Moreover, participants scored the statements presented as transcripts ( $M = 6.68, SD = 1.23$ ) as more credible than the statements presented as videos ( $M = 6.00, SD = 1.45$ ) [ $F(1, 232) = 22.11, p = .00, \eta_p^2 = 0.087$ ] and was strongly

supported by the Bayes factor ( $BF_{10} = 4550.13$ ). There was no significant main effect for group [ $F(1, 232) = 0.27, p = .60, BF_{10} = 0.40$ ], no interaction between group and veracity [ $F(1, 232) = 0.01, p = .94$ ], and no significant three-way interaction [ $F(1, 232) = 0.29, p = .59$ ]. These results were supported by the Bayes factors; the main effects model is 1.35 times more likely than the group  $\times$  veracity interaction model and 4.67 times more likely than the three-way interaction model.

### Relationship between beliefs and lie-detection accuracy

Finally, to answer our main question, we tested whether participants who hold more correct beliefs about verbal cues to deception were better at deception detection. We first calculated the exact number of correct beliefs for each participant based on the criteria outlined earlier.<sup>1</sup> Next, we correlated this score with the average credibility score of the truthful and fabricated statements. Having more correct beliefs showed a small positive correlation with a higher credibility for truthful statements ( $r = .19, p < .004$ ). No significant correlation was found between correct beliefs and the credibility score for lies ( $r = -.06$ ).

Additionally, we investigated whether years of experience for police officers was correlated with lie-detection accuracy. More experience was related to a higher credibility for both truthful ( $r = .33, p < .001$ ) and fabricated ( $r = .36, p < .001$ ) statements. Thus, more experience led to a criterion shift: better detection of truthful statements, but worse detection of fabricated statements. Results further showed that experience was not related to a better insight into useful verbal cues ( $r = .02$ ).

## DISCUSSION

The current study examined the relationship between beliefs about verbal cues to deception and deception accuracy. Subsidiary aims were to investigate police officers' and undergraduates' general beliefs about cues to deception, beliefs about 17 specific verbal cues, and to what extent presentation mode influenced this relationship. To this end, participants were asked to make credibility judgments about eight statements that were presented as either videos or verbatim transcripts. As expected, participants primarily reported nonverbal cues—most notably gaze aversion and nervous behavior—as useful for detecting deception, replicating our previous study (Bogaard et al., 2016), as well as many others (e.g., Akehurst et al., 1996; Masip & Herrero, 2015; Strömwall et al., 2004; Strömwall & Granhag, 2003; Taylor & Hick, 2007; Vrij et al., 2006; Vrij & Semin, 1996). Although verbal signals were reported to a lesser extent, the cues that were reported (i.e., inconsistencies, details, and coherence) were largely supported by empirical research (Amado et al., 2015; Amado, Arce, Fariña, & Vilaríño, 2016; Masip et al., 2005; Oberlader et al., 2016; Vrij, 2005).

<sup>1</sup> One observation could be considered an outlier (mean  $\pm 2SD$ ). However, removing this observation from the sample did not influence our results, and we therefore included all observations in the analyses.

More importantly, participants with more correct beliefs about verbal cues were more accurate at detecting deception, showing that the verbal cues drive credibility judgments to some extent. The finding that this relation was limited to the truthful statements can most likely be explained by the fact that the verbal cues we included in our questionnaire are primarily truth indicators. This correlation might also explain why lie-detection training that focuses primarily on shifting attention from nonverbal to verbal cues is more successful than training that focus on nonverbal, paraverbal, or multichannel cues (Hauch, Sporer, Michael, & Meissner, 2014). When people attend to verbal cues to detect deception, their subjective (i.e., cues they think they use) and objective (i.e., cues they actually use) indicators are more similar. This is not the case when people rely on nonverbal cues, as has been reported by Hartwig and Granhag (2014).

The finding that the years of experience of the officers did not influence their beliefs about useful verbal cues is in line with numerous studies that reported that both lay people and professionals, such as police and custom officers, prison guards, and judges, hold many similar (often incorrect) beliefs (Akehurst et al., 1996; Strömwall et al., 2004; Strömwall & Granhag, 2003; Vrij, 2008a; Vrij et al., 2006). On the other hand, experience was related to an increase in accuracy for truthful judgments but a decrease in accuracy for fabricated statements. This truth bias contrasts with the finding that officers show a lie bias, that is, a higher accuracy for detecting lies than truths (e.g., Meissner & Kassin, 2002). However, in their meta-analysis, investigating 206 statements and almost 25 000 judgments, Bond and DePaulo (2006) also found no support for an increased lie bias with experience. This pattern of our findings might be explained by the fact that—as the majority of the studies in Bond and DePaulo (2006)—we did not use high-stake lies, whereas the lie bias seems especially pronounced when officers are asked to judge the veracity of high-stake lies showed on video (Masip, Alonso, Herrero, & Garrido, 2016; Wright Whelan, Wagstaff, & Wheatcroft, 2015). Yet meta-analytic findings of Hartwig and Bond (2014) show that motivation, context, and type of lie do not influence detectability.

The classification accuracy found in our study was moderate. It is important to note that it was not our goal to investigate how well the lies could be detected. In fact, based on the pilot, we selected truthful and fabricated statements with the same credibility ratings, meaning the overall lie-detection performance was expected to be poor. Yet we created room for improvement, depending on whether the statements were presented as video or transcript.

We also investigated to what extent presentation mode would influence credibility judgments. Importantly, the findings discussed here might be confounded because of the skewed distribution of videos and transcripts. Overall, more police officers than students viewed videos, which complicates the interpretation of our findings. Nonetheless, we will shortly discuss the main results. Contrary to our hypothesis, police officers and undergraduates overall performed no better when they were offered transcripts instead of videos, which contrasts earlier research (Burgoo

et al., 2008; Davis, Markus, & Walters, 2006; Leach et al., 2016; Maier & Thurber, 1968; Mann et al., 2008). One possible explanation may be that offering transcripts also removes access to diagnostic vocal signals (e.g., pitch and speech latency). On the other hand, Bond and DePaulo (2006) reported that lie-detection accuracy of videos is inferior to that of written transcripts, but that lie-detection accuracy of transcripts, audio–visual, and audio presentations does not significantly differ. We have used verbatim transcripts of the audio fragments, meaning we transcribed the utterances of the interviewee, and did not ask them to write their statement down instead of written statements, which might account for these findings.

When we specifically investigate the accuracy for truthful and fabricated statements separately for each presentation mode, our results showed that police officers were more accurate for true statements presented as transcripts but better for fabricated statements presented as videos. For students, this difference was less pronounced. It might be that officers are more distracted by the presence of invalid stereotypical nonverbal cues when judging videos, and better focused on more valid truth criteria when judging transcripts compared with students. Both students and police officers also rated transcripts more credible than videos, regardless of the veracity of the statement, which might be due to a reduced degree of interactivity for texts compared with videos (Burgoo et al., 2008).

Some limitations in this study deserve some attention. As pointed out earlier, more police officers judged the credibility of the video clips than the transcripts. The resulting unequal cell sizes might have influenced our results. Yet given that there were no interaction effects for questionnaire format, veracity, and presentation mode, influence on our reported data is minimal. Second, there was a limited degree of interactivity between observers and the interviewee in our study. An interview situation is obviously much more interactive than the videos and transcripts we included. However, this was deliberate. Allowing our participants to interact with the interviewees raises doubt about whether lie-detection performance is the result of using correct cues, or of the quality of the interview. By asking people to give a statement about a negative autobiographical experience, we attempted to obtain sufficient ecological validity, without losing control over the ground truth of the statements. Third, we aimed at using statements with high ecological validity without having to trade off certainty of ground truth. Nonetheless, the statements we used are not high stake. Based on the meta-analytic findings of Hartwig and Bond (2014) showing that level of detectability is stable across settings, it is unlikely that this influences the generalizability of our results. Lastly, the effect of our findings is small, and by no means we intend to claim that people who have better insight in verbal cues are so-called lie-detection wizards. Yet in the light of the findings of Hartwig and Bond (2011), it is interesting that lie-detection accuracy is at least partly explained by relying on valid verbal cues.

In sum, our findings show that people, including police officers, still have incorrect beliefs about nonverbal cues but have better insight into verbal cues. Moreover, a better insight in verbal cues was related to an increased accuracy

for identifying truthful statements, showing that verbal cues do drive credibility judgments to some extent. This finding might explain why some people are slightly better than others in identifying liars. Moreover, this finding emphasizes the importance of advising police officers about moving their focus from nondiagnostic nonverbal cues to more diagnostic verbal cues.

## ACKNOWLEDGEMENT

This research was funded by Politie en Wetenschap (PW/OC/2014/02).

## REFERENCES

- Akehurst, L., Köhnken, G., Vrij, A., & Bull, R. (1996). Lay persons' and police officers' beliefs regarding deceptive behaviour. *Applied Cognitive Psychology, 10*, 461–471.
- Amado, B. G., Arce, R., & Fariña, F. (2015). Undeutsch hypothesis and criteria based content analysis: A meta-analytic review. *The European Journal of Psychology Applied to Legal Context, 7*, 1–10.
- Amado, B. G., Arce, R., Fariña, F., & Vilariño, M. (2016). Criteria-based content analysis (CBCA) reality criteria in adults: A meta-analytic review. *International Journal of Clinical and Health Psychology, 16*, 201–210.
- Bogaard, G., Meijer, E., Vrij, A., & Merckelbach, H. (2016). Strong but wrong: Beliefs about verbal and non-verbal cues to deception. *PLoS One, 11*, e0156615. doi:https://doi.org/10.1371/journal.pone.0156615.
- Bond, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Individual Differences, 10*, 214–234.
- Bond, C. F., & DePaulo, B. M. (2008). Individual differences in judging deception: Accuracy and bias. *Psychological Bulletin, 134*, 477–492.
- Burgoon, J. K., Blair, P. J., & Strom, R. E. (2008). Cognitive biases and nonverbal cue availability in detecting deception. *Human Communication Research, 34*, 552–599.
- Davis, M., Markus, K. A., & Walters, S. B. (2006). Judging the credibility of criminal suspect statements: Does mode of presentation matter? *Journal of Nonverbal Behavior, 30*, 181–198.
- DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129*, 74–118.
- Granhag, P. A., Andersson, L. O., & Strömwall, L. A. (2004). Imprisoned knowledge: Criminals' beliefs about deception. *Legal and Criminological Psychology, 9*, 103–119.
- Hartwig, M., & Bond, C. F. (2011). Why do lie-catchers fail? A lens model meta-analysis of human lie judgments. *Psychological Bulletin, 137*, 643–659.
- Hartwig, M., & Bond, C. F. J. (2014). Lie detection from multiple cues: A meta-analysis. *Applied Cognitive Psychology, 28*, 661–676. doi:https://doi.org/10.1002/acp.3052.
- Hartwig, M., & Granhag, P. A. (2014). Exploring the nature and origin of beliefs about deception: Implicit and explicit knowledge among lay people and presumed experts. In P. A. Granhag, A. Vrij, & B. Verschuere (Eds.), *Detecting deception: Current challenges and cognitive approaches*, (pp. 125 – 154). Chichester: John Wiley & Sons.
- Hartwig, M., Granhag, P. A., Strömwall, L. A., & Andersson, L. O. (2004). Suspicious minds: Criminals' ability to detect deception. *Psychology, Crime and Law, 10*, 83–95.
- Hauch, V., Sporer, S. L., Michael, S. W., & Meissner, C. A. (2014). Does training improve the detection of deception? A meta-analysis. *Communication Research, 1*–61.
- Jarosch, A. F., & Wiley, J. (2014). What are the odds? A practical guide to computing and reporting Bayes factors. *The Journal of Problem Solving, 7*, 2–9.
- Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review, 88*, 67–85.
- Leach, A. M., Ammar, N., England, D. N., Remigio, L. M., Kleinberg, B., & Verschuere, B. J. (2016). Less is more? Detecting lies in veiled witnesses. *Law and Human Behavior, 1*–10. https://doi.org/10.1037/lhb0000189.
- Lee, M. D., & Wagenmakers, E. J. (2013). *Bayesian cognitive modeling: A practical course*. Cambridge university press.
- Levine, T. R., & McCormack, S. A. (2014). Theorizing about deception. *Journal of Language and Social Psychology, 33*, 431–440.
- Maier, N. R. F., & Thurber, J. A. (1968). Accuracy of judgments of deception when an interview is watched, heard, and read. *Personnel Psychology, 21*, 23–30.
- Mann, S., Vrij, A., & Bull, R. (2004). Detecting true lies: Police officers' ability to detect suspects' lies. *Journal of Applied Psychology, 89*, 137–149.
- Mann, S., Vrij, A., Fisher, R. P., & Robinson, M. (2008). See no lies, hear no lies: Differences in discrimination accuracy and response bias when watching or listening to police suspect interviews. *Applied Cognitive Psychology, 22*, 1062–1071.
- Masip, J., Alonso, H., Herrero, C., & Garrido, E. (2016). Experienced and novice officers' generalized communication suspicion and veracity judgments. *Law and Human Behavior, 40*, 169–181.
- Masip, J., & Herrero, C. (2015). Police detection of deception: Beliefs about behavioral cues to deception are strong even though contextual evidence is more useful. *Journal of Communication, 65*, 125–145.
- Masip, J., Sporer, A. L., Garrido, E., & Herrero, C. (2005). The detection of deception with the reality monitoring approach: A review of the empirical evidence. *Psychology, Crime and Law, 11*, 99–122.
- Meissner, C. A., & Kassin, S. M. (2002). "He's guilty!": Investigator bias in judgments of truth and deception. *Law and Human Behavior, 26*, 469–480.
- Oberlander, V. A., Naefgen, C., Koppehele-Gossel, J., Quinten, L., Banse, R., & Schmidt, A. F. (2016). Validity of content-based techniques to distinguish true and fabricated statements: A meta-analysis. *Law and Human Behavior, 1*–63.
- 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.
- Steller, M., & Köhnken, G. (1989). Criteria based statement analysis. In D. C. Raskin (Ed.), *Psychological methods in criminal investigation and evidence*, (pp. 217 – 245). New York: Springer.
- Strömwall, L. A., & Granhag, P. A. (2003). How to detect deception? Arresting the beliefs of police officers, prosecutors and judges. *Psychology, Crime and Law, 9*, 19–36.
- Strömwall, L. A., Granhag, P. A., & Hartwig, M. (2004). Practitioners' beliefs about deception. In L. A. Strömwall, & P. A. Granhag (Eds.), *The detection of deception in forensic contexts*. Cambridge: Cambridge University Press.
- Taylor, R., & Hick, R. F. (2007). Believed cues to deception: Judgments in self-generated trivial and serious situations. *Legal and Criminological Psychology, 12*, 321–331.
- The Global Deception Research Team (2006). A world of lies. *Journal of Cross-Cultural Psychology, 37*, 60–74.
- Vargha, A., & Delaney, H. D. (2000). A critique and improvement of the CL common language effect size statistics of McGraw and Wong. *Journal of Educational and Behavioral Statistics, 25*, 101–132.
- Vrij, A. (2005). Criteria based content analysis: A qualitative review of the first 37 studies. *Psychology, Public Policy, and Law, 11*, 3–41.
- Vrij, A. (2008a). Beliefs about nonverbal and verbal cues to deception. In A. Vrij (Ed.), *Detecting lies and deceit*, (pp. 115 – 140). Chichester: Wiley.
- Vrij, A. (2008b). *Detecting lies and deceit: Pitfalls and opportunities*. Chichester: Wiley.
- Vrij, A. (2008c). Nonverbal dominance versus verbal accuracy in lie detection: A plea to change police practice. *Criminal Justice and Behavior, 35*, 1323–1336.
- Vrij, A., Akehurst, L., & Knight, S. (2006). Police officers', social workers', teachers' and the general public's beliefs about deception in children, adolescents and adults. *Legal and Criminological Psychology, 11*, 297–312.
- Vrij, A., & Semin, G. (1996). Lie experts' beliefs about nonverbal indicators of deception. *Journal of Nonverbal Behavior, 20*, 65–80.
- Wagenmakers, E. J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., . . . Morey, R. D. (2016). Bayesian inference for psychology. Part II: Example applications with JASP, 1–26. Retrieved from http://maartenmarmans.com/wp-content/uploads/2017/04/WagenmakersEtAlPartII.pdf website:
- Wright, C., & Wheatcroft, J. M. (2017). Police officers' beliefs about, and use of, cues to deception. *Journal of Investigative Psychology and Offender Profiling*. https://doi.org/10.1002/jip.1478.
- Wright Whelan, C., Wagstaff, G., & Wheatcroft, J. M. (2015). High stakes lies: Police and non-police accuracy in detecting deception. *Psychology, Crime & Law, 21*, 127–138.

## APPENDIX A

1. Coherence—CBCA
2. Clarity—RM
3. Spontaneous corrections—CBCA and SCAN
4. Contradictions—CBCA
5. Perceptual information—RM
6. Emotions—CBCA, RM and SCAN
7. Quantity of details—CBCA
8. Spatial information—RM and CBCA
9. Unstructured production—CBCA
10. Description of interaction—CBCA
11. Temporal information—RM and CBCA
12. Superfluous details—CBCA
13. Reproduction of conversation—CBCA
14. Reconstructability—RM
15. Unusual details—CBCA
16. Plausibility—RM
17. Cognitive operations—RM

## APPENDIX B

1. A possible cue to deception is coherence. This cue refers to the extent to which all elements of the statement are logical and make up a unified whole. How do you believe the statements of liars relate to those of truth tellers?
  - Less coherence for liars
  - No differences in coherence
  - More coherence for liars
  - I don't know
2. A possible cue to deception is clarity. This cue refers to the extent to which a statement is clear, sharp, and vivid (instead of dim and vague). How do you believe the statements of liars relate to those of truth tellers?
  - Less clarity for liars
  - No differences in clarity
  - More clarity for liars
  - I don't know
3. A possible cue to deception is number of spontaneous corrections. For written statements, this cue refers to the information that has been crossed out. For spoken statements, this cue refers to corrections that are made or information that is added to material previously provided in the statement without having been prompted by the interviewer. For example: 'It was already 10 o'clock, oh no wait 11 o'clock.' How do you believe the statements of liars relate to those of truth tellers?
  - Less spontaneous corrections for liars
  - No differences in spontaneous corrections
  - More spontaneous corrections for liars
  - I don't know
4. A possible cue to deception is contradictions. This cue refers to the extent to which a statement contains elements that contradict each other. For example, at the beginning of a statement someone reported that the

perpetrator entered the house around midnight, while saying later that the perpetrator broke in at 2.00 AM. How do you believe the statements of liars relate to those of truth tellers?

- Less contradictions for liars
  - No differences in contradictions
  - More contradictions for liars
  - I don't know
5. A possible cue to deception is degree of perceptual information. This cue refers to the presence of sensory information in a statement. Does the statement include sensorial experiences such as sounds ('He shouted at me'), smells ('It had a smell of rotten fish'), tastes ('The chips were very salty'), physical sensations ('It really hurt'), and visual details ('I saw him entering the house'). How do you believe the statements of liars relate to those of truth tellers?
    - Less perceptual information for liars
    - No differences in perceptual information
    - More perceptual information for liars
    - I don't know
  6. A possible cue to deception is number of described emotions. This cue refers to information that describes how the participant felt during an event and how these feelings develop and change throughout an event ('I was very scared, but relieved when it was all over'). How do you believe the statements of liars relate to those of truth tellers?
    - Less emotions for liars
    - No differences in emotions
    - More emotions for liars
    - I don't know
  7. A possible cue to deception is quantity of details. This cue refers to whether a statement is rich in detail and includes specific descriptions of place, time, persons, objects, and events. How do you believe the statements of liars relate to those of truth tellers?
    - Less details for liars
    - No differences in details
    - More details for liars
    - I don't know
  8. A possible cue to deception is spatial details. This cue refers to information about locations ('It happened in a park') or the spatial arrangement of people and/or objects ('The man was sitting to the left of his wife'). How do you believe the statements of liars relate to those of truth tellers?
    - Less spatial details for liars
    - No differences in spatial details
    - More spatial details for liars
    - I don't know
  9. A possible cue to deception is unstructured production. This cue refers to the chronological order in which the statements is told. Are they told in the same order in which they occurred (chronological), or not

- (unstructured)? How do you believe the statements of liars relate to those of truth tellers?
- Less unstructured production for liars
  - No differences in unstructured production
  - More unstructured production for liars
  - I don't know
10. A possible cue to deception is the number of described interactions. This cue refers to how action and reaction of the different actors presented in the statement are linked. For example, 'I said go away but he didn't, he just smiled, and then I started crying'. How do you believe the statements of liars relate to those of truth tellers?
- Less description of interactions for liars
  - No differences in description of interactions
  - More description of interactions for liars
  - I don't know
11. A possible cue to deception is temporal details. This cue refers to information about when the event happened ('It was early in the morning', 'an hour after I left') or explicitly describes a sequence of events ('When the visitor heard all that noise, he became nervous and left'). How do you believe the statements of liars relate to those of truth tellers?
- Less temporal details for liars
  - No differences in temporal details
  - More temporal details for liars
  - I don't know
12. A possible cue to deception is superfluous details. This cue refers to irrelevant details that are related to the event, but not necessary to understand the event. For example, a witness who says that the perpetrator tried to get rid of the cat that entered the bedroom because the perpetrator is allergic to cats. How do you believe the statements of liars relate to those of truth tellers?
- Less superfluous details for liars
  - No differences in superfluous details
  - More superfluous details for liars
  - I don't know
13. A possible cue to deception is reproduction of conversation. This cue refers to whether parts of the conversation are reported in original form or if the different speakers are recognizable in the reproduced dialogs. For example, he asked: 'Are you okay?' How do you believe the statements of liars relate to those of truth tellers?
- Less reproduction of conversation for liars
  - No differences in reproduction of conversation
  - More reproduction of conversation for liars
  - I don't know
14. A possible cue to deception is reconstructability of the story. This cue refers to whether it is possible to reconstruct the event on the basis of the information given. How do you believe the statements of liars relate to those of truth tellers?
- Less reconstructability for liars
  - No differences in reconstructability
  - More reconstructability for liars
  - I don't know
15. A possible cue to deception is unusual details. This cue refers to details of people, objects, or events that are unique, unexpected, or surprising but meaningful in the context. Examples are a witness who gives a description of a tattoo on the perpetrator's arm, or a witness who says that the perpetrator had a stutter. How do you believe the statements of liars relate to those of truth tellers?
- Less unusual details for liars
  - No differences in unusual details
  - More unusual details for liars
  - I don't know
16. A possible cue to deception is plausibility. This cue refers whether the story is plausible and realistic and makes sense. How do you believe the statements of liars relate to those of truth tellers?
- Less plausible for liars
  - No differences in plausible
  - More plausible for liars
  - I don't know
17. A possible cue to deception is cognitive operations. These cue descriptions of inferences are made by the participant based on existing knowledge. For example, 'It must have been cold because I had my winter coat' or 'It seemed to me that he/she did not recognize this building' is a conclusion based on observed behavior. How do you believe the statements of liars relate to those of truth tellers?
- Less cognitive operations for liars
  - No differences in plausible
  - More cognitive operations for liars
  - I don't know