Received: 26 November 2019 | Revised: 11 October 2020 | Accepted: 20 October 2020

DOI: 10.1002/acp.3755

RESEARCH ARTICLE

WILEY

Are criminals better lie detectors? Investigating offenders' abilities in the context of deception detection

Simon Schindler[1] | Laura K. Wagner[1] | Marc-André Reinhard[1] | Nico Ruhara[2] | Stefan Pfattheicher[3] | Joachim Nitschke[2]

[1]Department of Psychology, University of Kassel, Kassel, Germany

[2]Ansbach District Hospital, Ansbach, Germany

[3]Department of Psychology and Behavioural Sciences, Aarhus University, Aarhus, Denmark

Correspondence
Simon Schindler, Department of Psychology, University of Kassel, Holländische Str. 36–38, 34127 Kassel, Germany.
Email: schindler@uni-kassel.de

Summary

The present research examined lie detection abilities of a rarely investigated group, namely offenders. Results of the studies conducted thus far indicated a better performance of offenders compared to non-offenders when discriminating between true and false messages. With two new studies, we aimed at replicating offenders' superior abilities in the context of deception detection. Results of Study 1 (N = 76 males), in contrast, revealed that offenders were significantly worse at accurately classifying true and false messages compared to non-offenders (students). Results of Study 2 (N = 175 males) revealed that offenders' discrimination performance was not significantly different compared to non-offenders (clinic staff). An internal meta-analysis yielded no significant difference between offenders and non-offenders, questioning the generalizability of previous findings.

KEYWORDS

beliefs about deception, dark triad, deception detection, offenders

1 | INTRODUCTION

Lies and deception affect many areas of life, and the ability to accurately detect lies is important in many areas of society (Ekman, 1992). However, to accurately differentiate lies from true statements proves difficult (Hartwig & Bond, 2014). Meta-analyses reported an overall accuracy rate slightly above chance level (e.g., C. F. Bond & DePaulo, 2006). Even experts like police officers, detectives, psychologists and judges, all expected to bring a certain amount of experience in deception detection, often prove no more accurate than laypersons—at least when personal interaction and strategic intervention is not possible (e.g., Aamodt & Custer, 2006; C. F. Bond & DePaulo, 2006; Reinhard, Scharmach, & Müller, 2013; Vrij, 2008). The present research investigated lie detection abilities of another proposed expert group, namely offenders. However, few empirical studies have directly addressed this topic, perhaps because data from offenders are difficult to obtain. Results of these past studies suggest that offenders possess better lie detection skills than non-offenders (e.g., Hartwig, Granhag, Strömwall, & Andersson, 2004). Given that discriminating between lies and truths is a difficult task, it is important to obtain more evidence on the idea that accurate detection is less difficult for offenders compared to non-offenders.

The literature presents several reasons for the low accuracy rate in deception detection. For example, laypersons and experts do not pay attention to actual relevant cues when trying to detect deception, and they hold stereotypical beliefs about cues of deception (DePaulo et al., 2003; Hartwig & Bond, 2011). Liars also often remained undetected because they show few actual cues of deception (Hartwig & Bond, 2011; Sporer & Schwandt, 2006). For example, DePaulo et al. (2003) analyzed more than 150 different studies about cues in deception and found only a few cues to be correlated with deception.

One first empirical hint that offenders might be better than non-offenders in accurately classifying true and false messages was

provided by Vrij and Semin (1996) who found that offenders hold less stereotypical beliefs about cues of deception than do students and experts. In addition, Granhag, Andersson, Strömwall, and Hartwig (2004) found that prisoners' beliefs about cues of deception are less stereotypical than prison staff and students. Given that stereotypical beliefs are typically inaccurate, more accurate beliefs may increase detection deception ability.

Hartwig et al. (2004) were the first to directly test the idea that offenders are better at accurately classifying true and deceptive messages. Results of their study showed no significant difference between the two groups regarding overall accuracy, that is, the 52 prisoners did not outperform students (n = 52) in a direct comparison. Prisoners, however, performed significantly better than chance level (>50%), while students did not, pointing to a better ability.

Another general finding in deception detection research is that laypersons typically show a truth bias, meaning that messages are generally more often judged as truthful than deceptive (C. F. Bond & DePaulo, 2006). This logically results in higher accuracy rates for actually true messages (so called veracity effect; Levine, Park, & McCornack, 1999; Schindler & Reinhard, 2015). Such a truth bias may not be that strong among criminals who can be assumed to be suspicious toward others, possibly even resulting in a lie bias and thus improved detection rates when judging actual lies (reversed veracity effect; G. D. Bond, Malloy, Arias, Nunn, & Thompson, 2005). Such effects were found in the study of Hartwig et al. (2004): Prisoners showed a stronger lie bias and were significantly better than students at detecting actual lies. There was, however, no difference between the groups in detecting actually true messages.

In a more recent study, Jupe, Akehurst, Vernham, and Allen (2016) investigated deception abilities of 16 teenage offenders and 36 non-offenders. Their findings showed that teenage offenders' accuracy rate was significantly higher compared to non-offenders. However, using more valid cues of deception could not account for the better performance. Moreover, unlike teenage non-offenders, teenage offenders did not show a truth bias.

Two further studies investigated judgmental biases among prisoners and non-prisoners in face-to-face dyads (G. D. Bond, Malloy, et al., 2005; G. D. Bond, Thompson, & Malloy, 2005). While both studies revealed a lie bias among prisoners, results of the authors' signal detection analyses also included d', measuring discrimination ability between lies and truths (Green & Swets, 1966), with higher d' meaning better discrimination ability. In the study of G. D. Bond, Thompson, and Malloy (2005), d' was significantly higher among prisoners (n = 56) than among non-prisoners. In the study of G. D. Bond, Malloy, et al. (2005), no inferential statistical analysis on discrimination ability was reported, suggesting prisoners (n = 15) to be better only on a descriptive level. The authors did not, however, discuss further these findings on discrimination ability since this was not the topic of interest in these studies. Nevertheless, these two studies further support the claim that offenders are better lie detectors.

To explain offenders' better performance, Hartwig et al. (2004) and Jupe et al. (2016) referred to the so-called feedback hypothesis, meaning that specific feedback about accuracy in detecting lies and truth is needed to promote deception detection abilities. In most situations, however, where detecting lies is relevant, there is no systematic honest feedback. For example, customs officials do not receive any feedback about dutiable goods from those who have not been inspected (Granhag et al., 2004). While pure confrontation with situations in which people lie has been shown to be insufficient in improving an ability to detect lies (DePaulo & Pfeifer, 1986; Ekman & O'Sullivan, 1991; Vrij, 2008; Vrij & Semin, 1996), Hartwig et al. (2004) speculated that operating in a criminal environment might offer more opportunities for valid feedback regarding the success of detecting deception in others and in turn lead to more accurate beliefs about cues to deception and finally to better lie detection ability. However, direct evidence for the claim that offenders receive more valid feedback is lacking. It is therefore doubtful whether the feedback hypothesis provides a valid theoretical basis for the claim that offenders are actually better at lie detection.

One could presume that offenders and non-offenders differ in certain personality traits. Although individual differences have been generally found to play a minute role regarding lie detection abilities (Aamodt & Custer, 2006; C. F. Bond Jr. & DePaulo, 2008), more recent studies addressed the idea that Dark Triad traits (narcissism, Machiavellianism, and psychopathy; Paulhus & Williams, 2002) might be beneficial for lie detection. This idea basically refers to the notion that deception production ability is linked to lie detection ability and that deception production ability is a central component of the Dark Triad (Wright, Berry, Catmur, & Bird, 2015). While some studies provide initial evidence (Lyons, Croft, Fairhurst, Varley, & Wilson, 2017; Lyons, Healy, & Bruno, 2013), most of the studies conducted thus far fail to provide support for the beneficial role of the Dark Triad (e.g., Peace & Sinclair, 2012; Wissing & Reinhard, 2017; Wright et al., 2015; Zuckerman, DePaulo, & Rosenthal, 1981). These previous studies including the Dark Triad relied on psychologically rather healthy convenience samples, such as students, limiting the occurrence of extreme cases regarding the Dark Triad traits. Offenders might differ from the general population in this respect and might show more extreme values on the Dark Triad traits (e.g., Coid et al., 2009), potentially revealing new insights about the link to lie detection ability.

2 | THE PRESENT RESEARCH

In lie detection research, previous studies suggested that offenders have better skills than non-offenders in discriminating between false and true messages: Hartwig et al. (2004) showed that offenders (but not non-offenders) significantly performed better than chance, and findings of Jupe et al. (2016) revealed a significantly better performance among teenage offenders compared to non-offenders. One study by G. D. Bond, Thompson, and Malloy (2005) also reported significantly higher discrimination abilities of prisoners in a face-to-face setting, while a similar study by G. D. Bond, Malloy, et al. (2005) did not reveal a significant difference. Taking all studies together, the total number of investigated offenders amounts to only 139, thus the

available evidence is far from conclusive. By comparing deception detection rates between offenders and non-offenders, the present two studies aimed to provide more data on this matter. Furthermore, given that operating in a criminal environment possibly causes general distrust and suspicion toward others (G. D. Bond, Malloy, et al., 2005; Bond, Thompson, & Malloy, 2005), offenders were expected to show a lower percentage of messages judged true compared to non-offenders. In two studies, as reported below, we tested these ideas.

Recent research addressed the idea that Dark Triad traits are beneficial for lie detection ability, but revealed little to no support (e.g., Wissing & Reinhard, 2017). Given that offenders and non-offenders might differ in their Dark Triad traits (e.g., Coid et al., 2009), with offenders showing more extreme values, we included Dark Triad measures in Study 1 and in Study 2 to explore this idea.

In Study 1, we additionally addressed offenders' beliefs about deception and deception detection. First, we addressed the claim that having high confidence in the accuracy of one's veracity judgments is positively related to actual accuracy (Smith & Leach, 2019). Second, according to Granhag et al. (2004), we expected offenders to report fewer stereotypical beliefs about deception than non-offenders.

3 | STUDY 1

3.1 | Method

3.1.1 | Ethics statement

The study was conducted in full accordance with the Ethical Guidelines of the German Association of Psychologists (DGPs) and the American Psychological Association (APA) and was approved by the local ethics committee. All participants gave their written informed consent.

3.1.2 | Transparency

Data and the material of this study (except the videos) are available on the Open Science Framework (see osf.io/wf8jy).

3.1.3 | Participants

The total sample included 76 male participants, consisting of 33 offenders of the forensic psychiatry Marsberg in Germany and 43 non-offenders (students). The data were collected from October to November 2017. Inmates whom the local staff assessed as being potentially aggressive and agitating were not considered as participants. All other inmates were informed by local staff that a study on lie detection would be conducted and that they were welcome free to participate. Offenders' mean age was 29.76 years (range from 20 to 50 years). Nine offenders had not yet graduated school, 14 offenders mentioned secondary education, five named O-level and two named A-level as their highest educational qualification. The self-reported offenses ranged from theft and drug trafficking to human trafficking and attempted manslaughter. Addictions were reported by all offenders. The most commonly used drugs are cocaine (20), marijuana (19), heroin (11) and alcohol (8). Male non-offenders were recruited via social media or by asking them on the campus of a German university to participate. Mean age was 25.25 (ranging from 19 to 35). Twenty-eight students mentioned A-level as their highest educational qualification, while 16 mentioned the Bachelor's degree. According to their self-reports, crimes were not committed, and all self-reported addictions related exclusively to legal drugs, where nicotine (6) was most frequently reported.

3.1.4 | Power analysis

Using G*Power (Faul, Erdfelder, Buchner, & Lang, 2009), a *sensitivity* power analysis for an analysis of variance (ANOVA) (fixed effects, omnibus, one-way; number of groups = 2; alpha = .05) indicated that with this sample size (N = 76) we have 80% power to detect a significant effect of f = 0.33.

3.1.5 | Procedure and material

The study was divided in three parts and conducted equally in both groups. Participation took about 60 min.

3.1.6 | Dark triad

The first part assessed the personality traits of the Dark Triad. Machiavellianism was assessed by using the *Mach-IV* (Christie & Geis, 1970), consisting of 20 items (α = .72). Psychopathy was assessed by using the *Self-Report Psychopathy Scale III* (SRPS-III; Paulhus, Neumann, & Hare, 2009), consisting of 64 items (α = .90). Narcissism was assessed by using the *Narcissistic Admiration and Rivalry Questionnaire* (NARQ; Back et al., 2013), consisting of 18 items (α = .87). All answers were measured by using a 5-point Likert scale from “strongly disagree” to “strongly agree.”

3.1.7 | Stimulus material

The stimulus material was taken from Reinhard, Sporer, Scharmach, and Marksteiner (2011, Study 1). To create this stimulus material, 20 male students were recruited as target persons with the possibility of earning up to 30 Euros; and were randomly assigned to the truth or lie condition. The 10 participants in the truth condition had to play backgammon with another person while watching a confederate entering the room looking for her wallet after noticing that 20 Euros had gone missing from it. After this scenario, participants were then asked to go to the adjacent room where they would be interviewed

about the missing money. They were so far given 10 Euros for their participation but were told that they could earn an additional 20 Euros if they were successfully able to convince the interviewer that they had not taken the money.

In contrast, the 10 participants in the lie condition did not participate in the (backgammon) game. Instead, they were immediately given 10 Euros for participation and asked to take the 20 Euros from the wallet in the room then hide the money somewhere on their person. Then they were informed that they would have to go to the adjacent room where they would be interviewed about the missing money. They were instructed to deny having taken the money and were provided with a written alibi containing a story parallel to the situation in the truth condition (playing backgammon).

After that, both liars and truth tellers were sent to the adjacent room, where another experimenter posed as the interviewer. All 20 participants were asked the same questions. The interviewer began by asking, "You are suspected of having taken 20 Euros from the woman's wallet. Have you taken the money from the wallet?" Other statements and questions included, "What were you doing all the time in the first room?"; "Did anybody else enter the room while you were playing?"; "Please tell me about the game of backgammon you were playing"; "I have noticed that your game partner left the room to take a call. What did you do in the room while he was outside?"; and "So you would have had the opportunity to take the money while your game partner was outside?" After the recording, all participants were fully debriefed and received the additional 20 Euros (or could keep the 20 Euros they had hidden on their bodies, respectively).

In sum, the material used for the present study consisted of two video sets, each containing 10 videos (with five actually true and five deceptive messages). Participants were randomly assigned to one of the two video sets. Each video lasted about 2–3 min.

3.1.8 | Veracity judgments

In part two of the study, participants in our study were then told that they would watch a series of videos in which several persons were interrogated and suspected of having stolen 20 Euros from another person's wallet. They were further told that some of the persons they were about to see were telling the truth, as they had not stolen the money, and some of the persons were lying, as they had actually taken the money from the wallet. Importantly, participants were not informed about the balanced distribution. Participants were assigned randomly to one of the two sets. After each video, they indicated whether the statement in the video was false or true and indicated their confidence in that judgment, ranging from 0% (not confident at all) to 100% (very confident). Afterwards, participants indicated how many of the 10 videos they supposedly classified correctly. In order to increase motivation during the study participation, participants were told that they would receive feedback about their detection performance.

3.1.9 | Beliefs about deception

While the feedback on participants lie detection performance was prepared, we asked participants about their beliefs on how people behave when they are lying and in what ways people's behavior differs from when they are telling the truth. Inspired by Reinhard et al. (2013), we used 20 pervasive cues from the literature on subjective and objective indicators of deception (DePaulo et al., 2003), 3 referring to verbal, 4 to para-verbal, and 13 to non-verbal aspects. These items provided information on whether people believe that there are recognizable cues, how strong those beliefs about cues are and how stereotypical these beliefs actually are. Following the method of Akehurst, Köhnken, Vrij, and Bull (1996), we used a 7-point scale ranging from −3 (e.g., "Deceptive statements are less detailed than truthful statements") via 0 ("No difference") to +3 (e.g., "Deceptive statements are more detailed than truthful statements"). A list of the items (and their "correctness") is displayed in the Appendix (Table A1).

3.1.10 | Demographics

Then the demographic data were collected including offenses, addictions, mental problems, gender, age, education, and occupation. These data were exclusively for descriptive statistics and were stored separately for reasons of anonymity.

3.2 | Results

For descriptive statistics of overall classification accuracy, accuracy for actually true messages, accuracy for actually false messages, judgment confidence and percentage of messages judged true are displayed in Table 1. Four participants did not judge all of the videos (the number of missing values ranged from 1 to 4). Although we are unable to retrace the reasons for the missing values, in favor of higher power, we did not exclude these participants. Instead, calculation of percentages of classification accuracy and messages judged true were adapted to the number of veracity judgments.[1] Table 2 shows results of zero-order correlations between the most relevant variables.

3.2.1 | Messages judged true

Overall, 59.93% (SD = 16.47) of the presented messages were judged as true, which differed significantly from 50%, $t(75) = 5.26$, $p < .001$, indicating an overall truth bias. Moreover, the percentage of messages judged true of offenders (M = 61.80, SD = 16.59), $t(32) = 4.09$, $p < .001$, as well as of non-offenders (M = 58.50, SD = 16.41), $t(42) = 3.40$, $p = .002$, differed significantly from 50%, indicating a truth bias in both groups. There was, however, no significant difference between the groups, $F(1, 74) = 0.75$, $p = .390$, $\eta^2_p = .01$.

TABLE 1 Descriptive statistics for classification accuracy, confidence and messages judged true in offenders and students (Study 1)

	Offenders (n = 33)				Non-offenders (n = 43)			
Variable	***M***	***SD***	***Min***	***Max***	***M***	***SD***	***Min***	***Max***
Accuracy in %								
Overall	46.05	17.40	10.0	90.0	54.57	16.60	20.0	100.0
True messages	58.18	22.56	0.0	100.0	63.14	22.57	0.0	100.0
False messages	33.79	24.40	0.0	80.0	46.05	24.12	0.0	100.0
Confidence in %	74.67	13.22	47.0	100.0	65.94	13.59	21.4	85.5
Judged true in %	61.80	16.59	16.7	100.0	58.50	16.41	20.0	100.0

TABLE 2 Zero-order correlations between the main variables (Study 1)

		(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1)	Classification accuracy	1.00						
(2)	Confidence	**−.23**	1.00					
(3)	Messages judged true	−.06	**.24**	1.00				
(4)	Correct beliefs in cues	−.01	**−.29**	−.03	1.00			
(5)	Strength of belief in cues	−.05	.21	.15	**−.51**	1.00		
(6)	Psychopathy	−.14	.21	.05	−.11	.00	1.00	
(7)	Narcissism	.03	−.18	−.19	−.08	.02	**.55**	1.00
(8)	Machiavellianism	.01	−.09	−.07	.−04	−.10	**.39**	**.58**

Note: Bold marked correlations are significant with $p < .05$.

3.2.2 | Classification accuracy

Overall, 50.87% (SD = 17.37) of the presented messages were classified correctly, which differed not significantly from 50%, p = .662. To analyze differences in classification accuracy, we used a 2 × 2 mixed-model ANOVA with message type (true vs. false) as within-participants factor and group (offenders vs. non-offenders) as between-participants factor. Results showed a significant main effect of message type, $F(1, 74) = 30.35$, $p < .001$, $\eta^2_p = .29$, indicating higher accuracy in classifying true messages (M = 61.58, SD = 22.33) than false messages (M = 41.12, SD = 25.65). Furthermore, a significant main effect of group occurred, $F(1, 74) = 4.88$, p = .030, η^2_p = .06, indicating that non-offenders (M = 54.57, SD = 16.60) were better at distinguishing between true and false messages than offenders (M = 46.05, SD = 17.40). There was no significant interaction effect, p = .335. Analyzing each group separately, offenders' classification accuracy did not significantly differ from chance level, p = .078; this was also the case for non-offenders, p = .202.

3.2.3 | Confidence in own judgment

Confidence ratings differ significantly between offenders (M = 74.67, SD = 13.22) and non-offenders (M = 65.94, SD = 13.59), $t(71) = 2.73$, p = .008. Overall, confidence ratings were negatively correlated with classification accuracy, r = −.23, p = .049. and positively correlated with messages judged true, r = .24, p = .046.

3.2.4 | Beliefs about deception

To test whether offenders and non-offenders differ in beliefs regarding the cues about deception, in a first step, a multivariate analysis was calculated with group as independent factor and the 20 belief ratings as dependent variables. Results show a significant multivariate effect of group, $F(1, 54) = 1.99$, p = .023. In Figure 1, group differences between offenders and non-offenders, as well as observed effect sizes for cues, are depicted. Only the difference on "cooperative" was significant, whereas the differences on "pitch of voice" and "posture shifts" almost reached the conventional significance level of p = .05.

Next, we calculated the correctness of beliefs according to results of DePaulo et al. (2003); see Figure 1): For cues with significant effect sizes, participants received one point when the direction of their answer was in line with the direction of the observed cue's effect size by DePaulo et al. (see Figure 1). When observed effect sizes were nonsignificant, participants received one point when their answer was "no difference." The overall mean level of correct beliefs was 35.04% (SD = 14.80). Mean levels did not differ between offenders (M = 34.65, SD = 15.66) and non-offenders (M = 35.35, SD = 14.28), p = .969. Furthermore, there was no significant correlation between correct beliefs and overall classification accuracy, p = .917.

The data further allowed for analyzing participants' strength of beliefs about cues of deception; we recoded all negative values to positive values (e.g., −3 = 3, −2 = 2, −1 = 1) to get the amount of all assessed values. No differences in strength of beliefs between offenders (M = 1.34, SD = 0.67) and non-offenders (M = 1.26, SD = 0.51) were found, p = .558.

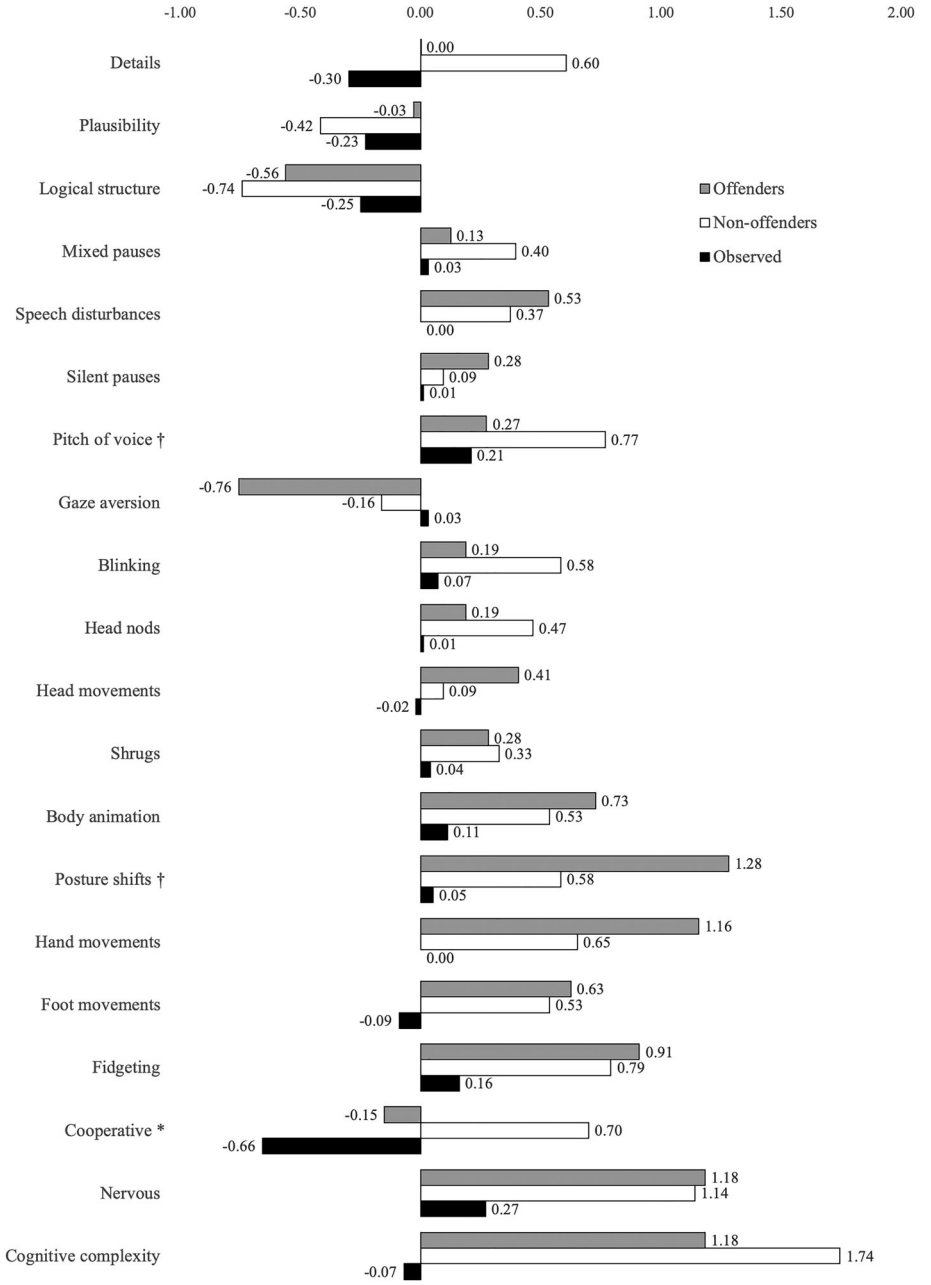

FIGURE 1 Average beliefs about cues of deception in offenders and non-offenders, $N = 76$ (Study 1). Significance marks refer to group differences between offenders and non-offenders. Observed values refer to the effect sizes *d* reported by DePaulo and colleagues [8], $^{*}p < .05$, $^{\dagger}p < .10$

3.2.5 | Dark triad

Testing for group differences revealed significantly higher levels of psychopathy for offenders ($M = 2.92$, $SD = 0.49$) than for non-offenders ($M = 2.37$, $SD = 0.36$), $t(74) = 5.56$, $p < .001$. No significant differences occurred regarding Machiavellianism and narcissism, both *p*s > .064. Furthermore, no significant correlations occurred between the Dark Triad traits and classification accuracy, all *p*s > .239, judgment confidence, all *p*s > .069, and correct beliefs about deception, all *p*s > .358.

3.3 | Discussion

Results of the first study revealed that offenders performed significantly worse than non-offenders regarding correct classification of true and false messages. Moreover, offenders as well as non-offenders showed an equally strong significant truth bias, resulting in significant better performance in classifying actually true statements (veracity effect). Furthermore, offenders and non-offenders did not significantly differ in their beliefs about cues (neither strength nor correctness). In fact, percentage of correct beliefs was quite low in both groups. This is in line with previous findings (Elaad, 2009; Granhag et al., 2004; Jupe et al., 2016), showing that both students and prison inmates hold stereotypical and inaccurate beliefs about cues indicative of lying. Further, there was no significant correlation between beliefs about cues in deception and classification accuracy. This could be due to inconsistent results and low effect sizes of cues in previous studies (Sporer & Schwandt, 2006) or to the lack of statistical power. Confidence was significantly negatively related to judgment accuracy. Offenders (vs. non-offenders) were, however, significantly more

confident regarding the accuracy of their judgments. The analyses of the Dark Triad personality traits showed that offenders have significantly higher levels of psychopathy than do non-offenders. No differences were found for Machiavellianism and narcissism. None of the three traits was significantly correlated with classification accuracy.

According to findings of Hauch, Sporer, Michael, and Meissner (2016), the effect size of feedback is small with $d = 0.19$. An a priori power analysis using this effect size revealed 222 participants are needed to have 80% power to detect a significant effect of this effect size. From this perspective, our study is underpowered. We thus collected more data to gain a more precise estimate.

4 | STUDY 2

4.1 | Method

4.1.1 | Ethics statement

The study was conducted in full accordance with the Ethical Guidelines of the German Association of Psychologists (DGPs) and the APA and was approved by the ethics committee of the German Association of Psychologists. All participants gave their written informed consent.

4.1.2 | Transparency

Material (except the videos) and the fully anonymized data of the constructs reported in this study are available on the Open Science Framework (see osf.io/wf8jy).

4.1.3 | Participants

The study was run in a district hospital in Ansbach, Germany; there are several different departments in the clinic, including the clinic for forensic psychiatry (divided into a psychiatric hospital and addiction treatment facility). We collected data from three groups of male participants. The first group consisted of 55 forensic patients ($M = 34.98$; range from 22 to 55 years), all of whom had been diagnosed with dissocial personality disorder according to ICD-10 criteria. The second group consisted of 62 forensic patients ($M = 34.37$; range from 21 to 57 years) not diagnosed with dissocial personality disorder, including diagnoses such as "mental and behavioral disorders due to use of alcohol" or other drug use. Across both groups, 23 patients had not yet graduated school, 69 offenders mentioned secondary education, 19 named O-level and three named A-level as their highest educational qualification; three patients did not answer the question. For individuals of both groups, the average number of registered entries in the German central registry amounts to 7.83 cases ($SD = 5.92$; all but 10 individuals had at least one entry in the past) against the law. As such, these two groups reflect the main target group of offenders. The third group ($M = 40.97$, range from 20 to 62 years) consisted of 58 people from the general population working at the clinic for forensic psychiatry. The total sample included 175 male participants (117 offenders and 58 non-offenders).[2] Data were collected from August 2018 to February 2019.

4.1.4 | Power analysis

A *sensitivity* power analysis for an ANOVA (fixed effects, omnibus, one-way; number of groups = 2; alpha = .05) indicated that with this sample size ($N = 175$) we have sufficient power (80%) to detect an effect of $f = 0.21$.

4.1.5 | Procedure and material

The study was divided in three parts and conducted equally in all three groups. Participation took about 90 minutes. We also assessed other behavioral paradigms not reported in this paper; the full list of paradigms is available on request from the fifth author of this paper.

4.1.6 | Dark triad

The 27-item Short Dark Triad questionnaire (Jones & Paulhus, 2014; nine items per dimension) was used to assess narcissism (e.g., "I have been compared to famous people"; $\alpha = .67$), Machiavellianism (e.g., "It's not wise to tell your secrets"; $\alpha = .81$), and psychopathy (e.g., "Payback needs to be quick and nasty"; $\alpha = .77$). Responses were given on a 5-point Likert scale ranging from "strongly disagree" to "strongly agree."

4.1.7 | Stimulus material

The stimulus material was taken from Dickhäuser, Reinhard, and Marksteiner (2012). Sixteen students (50% female) were recruited as target persons, with the possibility of earning up to 30 Euros, and the students were randomly assigned to the truth or lie condition. In both conditions, students worked in groups of four people in total, who were all blind to the experimental conditions; none was a confederate. Before working in groups, all participants were introduced to a young man named Markus (a confederate) and were told they would soon work with him on a task. After the introduction, Markus left the room but left his bag behind. Students were then given instructions and each student received 10 Euros for participation. In the truth condition, stimulus persons were asked to talk for 15 min. About the topics of school hobbies and current movies. In the lie condition, they were instructed to write a letter within 15 min to Markus with offending content. They were assured there would be no negative consequences for doing so, that they would stay anonymous, and that the material was only used for scientific purposes. Afterwards,

participants from both conditions were led to a different room where they were interviewed. The interview was videotaped. The interviewer was blind to the conditions. Participants were seated on a chair in front of a camera with 3 m distance. The interviewer accused all participants of having been involved in a harmful activity (i.e., having written a threatening letter). Immediately after entering the interview room, all students received another 10 Euros and were told they could only keep these additional 10 Euros if they were able to convince the interviewer during the interview that they had not been involved in the relational aggression. They were told they would receive 10 Euros reward additionally if they were able to convince the interviewer that they did not act as aggressors. Thus, convincing students expected a total reward of 30 Euros while unconvincing ones expected to only receive 10 Euros in sum (in fact, all participants received 30 Euros in the end).

During the interview, the interviewer asked all students the same questions and instructed them to answer each question within 30 s. The questions were, for example: "Have you been involved in writing this letter?" and, "If you haven't been involved, what exactly did you do in the first room?" In sum, the material consisted of two video sets, each containing eight videos (with four actually true and four deceptive messages). Each video lasted about 2 min.

4.1.8 | Veracity judgments

Participants were told that they would watch a series of videos in which several persons were interrogated and suspected of having written a threatening letter to a person named Markus during an unattended break and having then hidden the letter in Markus' bag. Parallel to Study 1, they were further told that some of the persons they were about to see were telling the truth, as they had not written the letter, and some of the persons were lying, as they had actually written the letter. Importantly, participants were not informed about the balanced distribution. Participants were assigned randomly to one of the two sets. After each video, they indicated whether the statement in the video was false or true.

4.1.9 | Demographics

Participants were asked for gender and age.

4.2 | Results

The descriptive statistics of all lie detection parameters are displayed in Table 3. Table 4, in turn, shows results of zero-order correlations.

4.2.1 | Messages judged true

Overall, 54.57% (*SD* = 24.69) of the presented messages were judged as true, which differed significantly from 50%, $t(174) = 2.45$, $p = .015$, indicating an overall truth bias. Moreover, the percentage of messages judged true among offenders ($M = 54.70$, $SD = 25.62$) significantly differed from 50%, $t(116) = 1.98$, $p = .050$, indicating a significant truth bias in this group. There was a truth bias among non-clinical non-offenders ($M = 54.31$, $SD = 22.89$), which failed, however, to reach conventional significance levels, $t(57) = 1.43$, $p = .157$. An ANOVA revealed no significant difference between the groups, $F(1, 173) = 0.10$, $p = .922$.

4.2.2 | Classification accuracy

Overall, 50.86% (*SD* = 15.08) of the presented messages were classified correctly, which differed non-significantly from 50%, $p = .453$. To analyze differences in classification accuracy, we used a 2 × 2 mixed-model ANOVA with message type (true vs. false) as within-participants factor and group (offenders vs. non-offenders) as between-participants factor. Results showed a significant main effect of message type, $F(1, 173) = 5.14$, $p = .025$, $\eta^2_p = .03$, indicating higher accuracy in classifying true messages ($M = 55.43$, $SD = 31.93$) than

TABLE 4 Zero-order correlations between the main variables (Study 2)

		(1)	(2)	(3)	(4)
(1)	Classification accuracy	1.00			
(2)	Messages judged true	**.25**	1.00		
(3)	Psychopathy	–.00	–.13	1.00	
(4)	Narcissism	.04	–.08	**.48**	1.00
(5)	Machiavellianism	.07	–.08	**.56**	**.43**

Note: Bold marked correlations are significant with $p < .01$.

TABLE 3 Descriptive statistics for classification accuracy, confidence and messages judged true in offenders and students (Study 2)

	Offenders (*n* = 117)				Non-offenders (*n* = 58)			
Variable	*M*	*SD*	*Min*	*Max*	*M*	*SD*	*Min*	*Max*
Accuracy in %								
Overall	51.07	15.01	25.0	87.5	50.43	15.35	25.0	87.5
True messages	55.77	32.23	0.0	100.0	54.74	31.67	0.0	100.0
False messages	46.37	26.92	0.0	100.0	46.12	22.85	0.0	100.0
Judged true in %	54.70	25.62	0.0	100.0	54.31	22.89	0.0	100

false messages (M = 46.29, SD = 25.58). There was no significant main effect of group, $F(1, 173) = 0.07$, $p = .793$, $\eta^2_p = .00$, and no significant interaction effect, $F(1, 173) = 0.10$, $p = .922$, $\eta^2_p = .00$. Analyzing each group separately, offenders' classification accuracy did not significantly differ from chance level, $p = .443$; this was also the case for non-offenders, $p = .831$.

4.2.3 | Dark triad

Separate ANOVAs revealed significant group differences for all three Dark Triad traits. First, there were significantly higher levels of psychopathy among offenders (M = 2.51, SD = 0.67) compared to non-offenders (M = 1.77, SD = 0.55), $F(1, 172) = 53.05$, $p < .001$, $\eta^2_p = .24$. Second, there were significantly high levels of narcissism among offenders (M = 2.65, SD = 0.65) compared to non-offenders (M = 2.38, SD = 0.61), $F(1, 173) = 7.21$, $p = .008$, $\eta^2_p = .04$. Third, there were significantly higher levels of Machiavellianism among offenders (M = 2.85, SD = 0.83) compared to non-offenders (M = 2.53, SD = 0.57), $F(1, 173) = 6.96$, $p = .009$, $\eta^2_p = .04$. However, no significant correlations occurred between the Dark Triad traits and classification accuracy, all ps > .350.

4.3 | Discussion

Results of this study revealed that offenders' detection performance was not significantly different compared to non-offenders. Moreover, there was no significant difference between the groups regarding the number of messages judged true; both groups showed an equally strong truth bias, although only the number of messages judged true differed significantly from 50% among offenders. Across both groups, there was a significant truth bias, resulting in better performance in classifying actually true statements. The analyses of the Dark Triad personality traits showed that offenders have significantly higher levels of psychopathy, Machiavellianism, and narcissism. However, none of the three traits was significantly correlated with classification accuracy.

4.4 | Internal meta-analysis

The present two studies have 80% power only to detect significant effects with f = 0.33 (Study 1) and f = 0.21 (Study 2), respectively. To gain a more precise estimate of the overall effect size, we performed an internal meta-analysis on the effect of group on classification accuracy using *Hedges'g* as effect size (Lipsey & Wilson, 2001). The total number of participants across the two studies was N = 251 (offenders = 150; non-offenders = 101). The meta-analysis was conducted using the *metafor* package for R (Viechtbauer, 2010). Results of a random effect model yielded a nonsignificant effect size, $g = -0.20$, $SE = 0.27$, $p = .455$, 95% CI: [−0.73, 0.33].

5 | GENERAL DISCUSSION

The present studies examined lie detection abilities of offenders and whether they differ in comparison to non-offenders. Results of Study 1 revealed that offenders were significantly worse than non-offenders in accurately classifying true and deceptive messages. In Study 2, there was no significant difference between offenders and non-offenders. A meta-analysis across both studies yielded no significant effect. These findings stand in contrast to previous findings (G. D. Bond, Thompson, & Malloy, 2005; Hartwig et al., 2004; Jupe et al., 2016), having suggested an improved classification ability of offenders.

According to the Hartwig et al. (2004), improved skills among offenders were expected according to the assumptions that (a) valid feedback on the accuracy of one's veracity judgments improve lie detection skills and that (b) operating in a criminal environment might offer more opportunities for feedback regarding the success of detecting deception of others. These assumptions may be invalid: First, recent meta-analytical findings across four studies revealed only a small effect size of trainings where systematic valid feedback was given (Hauch et al., 2016). Effect sizes may even be smaller regarding (unsystematic) feedback in a criminal environment. Furthermore, it seems likely that offenders do not receive valid and systematic feedback in all situations. In light of these arguments, the feedback hypothesis provides a weak theoretical basis for the claim that offenders are better lie detectors.

Lie detection in general has been shown to be a difficult task (C. F. Bond & DePaulo, 2006). To be successful, the stimulus material requires objective differences between liars and truth tellers. In Study 1, we used the stimulus material of Reinhard et al. (2011, Study 1) who found higher accuracy levels when situational familiarity with the judgmental context was high. In another study using the same material, Reinhard, Dahm, and Scharmach (2012) found higher detection accuracy among police officers after inducing a feeling of being experienced in lie detection. In Study 2, we used the stimulus material of Dickhäuser et al. (2012) who found higher accuracy levels when participants were instructed not to use stereotypical non-verbal cues for their judgment. These previously found moderating conditions point to objective differences between liars and truth tellers in our used stimulus materials, rendering substantial increases in accuracy rate over chance level possible.

While diagnostic cues seem often to be faint (e.g., Hartwig & Bond, 2011), DePaulo et al. (2003) found that most diagnostic cues refer to content aspects (e.g., fewer details, more repetitions, less logical structure, less plausibility). Correspondingly, C. F. Bond and DePaulo (2006) revealed that a focus on paraverbal and content cues (in audiovisual presentations, audio presentations, or transcripts) leads to higher accuracy rates compared to visual cues (in video-only presentations). As a consequence, using content related cues for veracity judgments should in particular increase the likelihood of improved classification performance. In line with this reasoning, Reinhard et al. (2011, Study 1) further found self-reported use of verbal content information was significantly positively correlated with detection

accuracy, whereas this was not the case for self-reported use of non-verbal information. Given that using content related cues has been assumed to require more cognitive effort and elaborate thinking than using stereotypical non-verbal cues (e.g., gaze aversion), cognitive abilities have been suggested to play an important role in successful lie detection (Reinhard, 2010; Reinhard & Sporer, 2008). From this perspective, offenders' lower education and higher drug use speak against the claim that they are better lie detectors. Even when assuming offenders actually receive valid and systematic feedback on their veracity judgments, their compromised cognitive abilities might inhibit them from benefiting from this feedback. Thus, future studies should include cognitive ability tests when comparing non-offenders and offenders.

In Studies 1 and 2, offenders and non-offenders showed an equally strong truth bias, resulting in better performance when classifying actually true statements (veracity effect). While Jupe et al. (2016) found no bias at all for either group, Hartwig et al. (2004) and G. D Bond, Malloy, et al. (2005); Bond, Thompson, and Malloy (2005) found a lie bias for offenders. This inconsistent picture may be due to the stimulus material and the heterogeneity of the samples. In our studies, offenders have been recruited from forensic clinics. In Study 2 in particular, about half of the offenders had been diagnosed with dissocial personality disorder. Beyond mental health, the so-far investigated offender samples are also likely to differ in their "level of criminality." To explain the found lie bias among offenders, Hartwig et al. (2004) reasoned that deception is probably more frequent in the criminal environment and that consequences of being duped, or conversely failing to dupe, are probably more severe in the criminal environment. Following this argument, level of criminality should predict the judgmental bias, that is, a lie bias should especially occur among offenders with high criminal experience. Hartwig et al. (2004) provide a further explanation for a lie bias, namely that offenders themselves probably have an extensive experience of lying, creating a false consensus bias that other people lie as often as themselves. Whereas no information on participants' everyday deception frequency had been collected in our studies, in Study 2 we were provided with the number of offenders' entries in the German central registry; there were, however, no significant correlations with messages judged true, $r = -0.03$, $p = .725$, or classification accuracy, $r = .02$, $p = .819$. In the so-far conducted studies on offenders and lie detection, little to no information was provided on the criminality background of the investigated offenders; as an exception, G. D. Bond, Thompson, and Malloy (2005) provided information about prisoners' sentence length. For better comparability of the samples, future research should provide more details on this matter.

The analyses of the Dark Triad personality traits showed that offenders have significantly higher levels of psychopathy (Studies 1 and 2), narcissism, and Machiavellianism (Study 2) than do non-offenders. Differences between the studies may be due to using different Dark Triad measures. However, none of the three traits correlated with classification accuracy, corroborating recent findings (e.g., Wissing & Reinhard, 2017). It remains doubtful, therefore, that strong Dark Triad traits go along with better deception detection skills.

The ecological validity in laboratory studies is a controversial subject. Video material in lie detection research often refers to people inventing a story or just imagining a situation. In contrast, in the present studies, the used video material referred to very realistic situations: The liar concealed an act that had already been committed, as is often the case with criminal offenses. According to recent results of Hartwig and Bond (2014), high-stakes lies were not found to be any easier to detect than low-stakes lies—that is, the absence in our stimulus material of harsh consequences when getting caught in the act of lying, and to follow the potentially low motivation for lying as credibly as possible, is less of a problem. Nevertheless, as already noted, the experimental laboratory situations differ from everyday situations where usually the possibility of interaction and reciprocal communication is given.

The present research aimed to provide additional data on the question regarding whether people with a criminal background have better or worse lie detection abilities. So far, only few studies have directly addressed this topic, and results suggest better abilities among offenders. In contrast, offenders in the present studies did not overall perform better than non-offenders, supporting the claim that lie detection is a difficult task, even for offenders.

ACKNOWLEDGEMENT

Open access funding enabled and organized by Projekt DEAL. [Correction added on 18 November 2020, after first online publication: Projekt Deal funding statement has been added.]

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Data and the material of this study (except the videos) are available on the Open Science Framework (see osf.io/wf8jy). Material (except the videos) and the fully anonymized data of the constructs reported in this study are available on the Open Science Framework (see osf.io/wf8jy).

ORCID


Simon Schindler https://orcid.org/0000-0003-1764-7241

ENDNOTES

[1] Excluding these four participants revealed that the difference between offenders and non-offenders regarding classification accuracy was not significant, $F(1, 70) = 3.24$, $p = .076$, $\eta^2_p = .04$. There was no significant effect between the groups regarding the number of messages judged true, $F(1, 70) = 0.08$, $p = .325$, $\eta^2_p = .01$.

[2] We report on the OSF parallel analyses with two groups of offenders (offenders with dissocial personality disorder and other offenders). Differences among the three groups regarding classification accuracy were nonsignificant, $p = .947$.

REFERENCES

Aamodt, M. G., & Custer, H. (2006). Who can best catch a liar? A meta-analysis of individual differences in detecting deception. *The Forensic Examiner*, *15*, 6–11.

Akehurst, L., Köhnken, G., Vrij, A., & Bull, R. (1996). Laypersons' and police officers' beliefs regarding deceptive behaviour. *Applied Cognitive Psychology*, *10*, 461–471.

Back, M. D., Küfner, A. C., Dufner, M., Gerlach, T. M., Rauthmann, J. F., & Denissen, J. J. (2013). Narcissistic admiration and rivalry: Disentangling the bright and dark sides of narcissism. *Journal of Personality and Social Psychology*, *105*, 1013–1037.

Bond, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review*, *10*, 214–234.

Bond, G. D., Malloy, D. M., Arias, E. A., Nunn, S. N., & Thompson, L. A. (2005). Lie-biased decision making in prison. *Communication Reports*, *18*, 9–19.

Bond, G. D., Thompson, L. A., & Malloy, D. M. (2005). Vulnerability of older adults to deception in prison and nonprison contexts. *Psychology and Aging*, *20*, 60–70.

Bond, C. F., Jr., & DePaulo, B. M. (2008). Individual differences in judging deception: Accuracy and bias. *Psychological Bulletin*, *134*, 477–492.

Christie, R., & Geis, F. L. (1970). *Studies in Machiavellianism*. New York, NY: Academic Press.

Coid, J., Yang, M., Ullrich, S., Roberts, A., Moran, P., Bebbington, P., … Singleton, N. (2009). Psychopathy among prisoners in England and Wales. *International Journal of Law and Psychiatry*, *32*, 134–141.

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin*, *129*, 74–118.

DePaulo, B. M., & Pfeifer, R. L. (1986). On-the-job experience and skill at detecting deception. *Journal of Applied Social Psychology*, *16*, 249–267.

Dickhäuser, O., Reinhard, M. A., & Marksteiner, T. (2012). Accurately detecting students' lies regarding relational aggression by correctional instructions. *Educational Psychology*, *32*, 257–271.

Ekman, P. (1992). *Telling lies: Cues to deceit in the marketplace, politics, and marriage*. New York, NY: Norton.

Ekman, P., & O'Sullivan, M. (1991). Who can catch a liar? *American Psychologist*, *46*, 913–920.

Elaad, E. (2009). Lie-detection biases among male police interrogators, prisoners, and laypersons. *Psychological Reports*, *105*, 1047–1056.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149–1160.

Granhag, P. A., Andersson, L. O., Strömwall, L. A., & Hartwig, M. (2004). Imprisoned knowledge: Criminals' beliefs about deception. *Legal and Criminological Psychology*, *9*, 103–119.

Green, D. A., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. NewYork, NY: Wiley.

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. (2014). Lie detection from multiple cues: A meta-analysis. *Applied Cognitive Psychology*, *28*, 661–676.

Hartwig, M., Granhag, P. A., Strömwall, L. A., & Andersson, L. O. (2004). Suspicious mind: Criminals' ability to detect deception. *Psychology, Crime & Law*, *10*, 83–95.

Hauch, V., Sporer, S. L., Michael, S. W., & Meissner, C. A. (2016). Does training improve the detection of deception? A meta-analysis. *Communication Research*, *43*, 283–343.

Jones, D. N., & Paulhus, D. L. (2014). Introducing the short dark triad (SD3): A brief measure of dark personality traits. *Assessment*, *21*, 28–41.

Jupe, L., Akehurst, L., Vernham, Z., & Allen, J. (2016). Teenage offenders' ability to detect deception in their peers. *Applied Cognitive Psychology*, *30*, 401–408.

Levine, T. R., Park, H. S., & McCornack, S. A. (1999). Accuracy in detecting truths and lies: Documenting the "veracity effect". *Communication Monographs*, *66*, 125–144.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Lyons, M., Croft, A., Fairhurst, S., Varley, K., & Wilson, C. (2017). Seeing through crocodile tears? Sex-specific associations between the dark triad traits and lie detection accuracy. *Personality and Individual Differences*, *113*, 1–4.

Lyons, M., Healy, N., & Bruno, D. (2013). It takes one to know one–psychopathy and deception detection ability. *Personality and Individual Differences*, *55*, 676–679.

Paulhus, D. L., Neumann, C. S., & Hare, R. D. (2009). *Manual for the self-report psychopathy scale*. Toronto, ON: Multi-health Systems.

Paulhus, D. L., & Williams, K. M. (2002). The dark triad of personality: Narcissism, Machiavellianism, and psychopathy. *Journal of Research in Personality*, *36*, 556–563.

Peace, K. A., & Sinclair, S. M. (2012). Cold-blooded lie catchers? An investigation of psychopathy, emotional processing, and deception detection. *Legal and Criminological Psychology*, *17*, 177–191.

Reinhard, M.-A. (2010). Need for cognition and the process of lie detection. *Journal of Experimental Social Psychology*, *46*, 961–971.

Reinhard, M. A., Dahm, J., & Scharmach, M. (2012). Perceived experience and police officers' ability to detect deception. *Policing: An International Journal of Police Strategies & Management*, *35*, 822–834.

Reinhard, M.-A., Scharmach, M., & Müller, P. (2013). It's not what you are, it's what you know: Experience, beliefs, and the detection of deception in employment interviews. *Journal of Applied Social Psychology*, *43*, 467–479.

Reinhard, M.-A., & Sporer, S. L. (2008). Verbal and nonverbal behaviour as a basis for credibility attribution: The impact of task involvement and cognitive capacity. *Journal of Experimental Social Psychology*, *44*, 477–488.

Reinhard, M.-A., Sporer, S. L., Scharmach, M., & Marksteiner, T. (2011). Listening, not watching: Situational familiarity and the ability to detect deception. *Journal of Personality and Social Psychology*, *101*, 467–484.

Schindler, S., & Reinhard, M.-A. (2015). Increasing skepticism toward potential liars: Effects of existential threat on veracity judgements and the moderating role of honesty norm activation. *Frontiers in Psychology*, *6*, 1312.

Smith, A. M., & Leach, A. M. (2019). Confidence can reliably discriminate between accurate and inaccurate lie decisions. *Perspectives on Psychological Science*, *14*, 1062–1071.

Sporer, S. L., & Schwandt, B. (2006). Paraverbal indicators of deception: A meta-analytic synthesis. *Applied Cognitive Psychology*, *20*, 421–446.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*, 1–48.

Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities*. Chichester, England: Wiley & Sons.

Vrij, A., & Semin, G. R. (1996). Lie experts' beliefs about nonverbal indicators of deception. *Journal of Nonverbal Behavior*, *20*, 65–80.

Wissing, B. G., & Reinhard, M. A. (2017). The dark triad and the PID-5 maladaptive personality traits: Accuracy, confidence and response bias in judgments of veracity. *Frontiers in Psychology*, *8*, 1549.

Wright, G. R., Berry, C. J., Catmur, C., & Bird, G. (2015). Good liars are neither 'dark' nor self-deceptive. *PLoS One*, *10*, e0127315.

Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. *Advances in Experimental Social Psychology*, *14*, 1–59.

How to cite this article: Schindler S, Wagner LK, Reinhard M-A, Ruhara N, Pfattheicher S, Nitschke J. Are criminals better lie detectors? Investigating offenders' abilities in the context of deception detection. *Appl Cognit Psychol*. 2021;35:203–214. https://doi.org/10.1002/acp.3755

APPENDIX

TABLE A1 Items about cues included in the questionnaire, given answer alternatives and correct answers (Study 1)

Item	Negative alternative (<)	Positive alternative (>)	Correct[a]
Details	Deceptive statements are less detailed than truthful statements.	Deceptive statements are more detailed than truthful statements.	<
Plausibility	Deceptive statements are less plausible than truthful ones.	Deceptive statements are more plausible than truthful ones.	<
Logical structure	Deceptive statements are less consistent than truthful ones.	Deceptive statements are less consistent than truthful ones.	<
Mixed pauses	Liars hesitate less than truth tellers.	Liars hesitate more than truth tellers.	—
Speech disturbances	Liars make fewer mistakes while speaking than truth tellers.	Liars make more mistakes while speaking than truth tellers.	—
Silent pauses	Liars halt less than truth tellers.	Liars halt more than truth tellers.	—
Pitch of voice	Liars' pitch is lower than truth tellers.	Liars' pitch is higher than truth tellers.	>
Gaze aversion	Liars look their conversational partner in the eyes less than do truth tellers.	Liars look their conversational partner in the eyes more than do truth tellers.	—
Blinking	Liars blink less than truth tellers.	Liars blink more than truth tellers.	—
Head nods	Liars nod less than truth tellers.	Liars nod more than truth tellers.	—
Head movements	Liars move their heads less than truth tellers.	Liars move their heads more than truth tellers.	—
Shrugs	Liars shrug their shoulders less than truth tellers.	Liars shrug their shoulders more than truth tellers.	—
Body animation	Liars move less than truth tellers.	Liars move more than truth tellers.	—
Posture shifts	Liars show fewer changes in their physical pose than truth tellers.	Liars show more changes in their physical pose than truth tellers.	—
Hand movements	Liar move their hands/fingers less than truth tellers.	Liar move their hands/fingers more than truth tellers.	—
Foot movements	Liar move their feet/legs less than truth tellers.	Liar move their feet/legs more than truth tellers.	—
Fidgeting	Liars fidget less than truth tellers.	Liars fidget more than truth tellers.	>
Cooperative	Liars are less cooperative than truth tellers.	Liars are more cooperative than truth tellers.	<
Nervous	Liars are less nervous than truth tellers.	Liars are more nervous than truth tellers.	>
Cognitive complexity	It takes less mental effort to lie than to tell the truth.	It takes more mental effort to lie than to tell the truth.	—

Note: For each item there was also a neutral (—) answer alternative named "no difference."
[a]Correct answers according to DePaulo et al. (2003).