Perspectives on Psychological Science
2019, Vol. 14(6) 1062–1071

DOI: 10.1177/1745691619863431
www.psychologicalscience.org/PPS
SAGE

Confidence Can Be Used to Discriminate Between Accurate and Inaccurate Lie Decisions

Andrew M. Smith[1,2] and Amy-May Leach[3]
[1]Department of Psychology, Iowa State University; [2]Department of Psychology, Carleton University; and [3]Faculty of Social Science and Humanities, University of Ontario Institute of Technology

Abstract
There is a long-standing belief that confidence is not useful at discriminating between accurate and inaccurate deception decisions. Historically, this position made sense because people showed little ability to discriminate lie-tellers from truth-tellers. But, it is now widely accepted that, under certain conditions, people can discriminate between lie-tellers and truth-tellers. Nevertheless, belief that confidence does not discriminate between accurate and inaccurate responses persists. This belief is somewhat paradoxical because, to the extent that people can discriminate between lie-tellers and truth-tellers, signal detection theory naturally predicts a positive relationship between confidence and accuracy. In line with our signal-detection-based predictions, we show that, among decisions about whether someone is lying, those made with high confidence are more accurate than those made with low confidence. This important relationship has gone unnoticed in past work because of a reliance on inappropriate measures. Past research examining the confidence–accuracy relationship in deception research relied on correlating average confidence with proportion of correctly identified lies. These correlations provide information on whether more confident judges tend to be more accurate but remain silent on the arguably more important question of whether higher confidence decisions are more accurate than lower confidence decisions. We show that confidence–accuracy characteristic analyses are uniquely suited to measuring the confidence–accuracy relationship in deception research.

Keywords
deception detection, signal detection theory, confidence–accuracy characteristic analysis, metacognition

There is long-standing evidence that the majority of people have little innate ability to detect deception. A seminal meta-analysis, including more than 200 manuscripts and nearly 25,000 participants, concluded that people were accurate at deception detection only 54% of the time (Bond & DePaulo, 2006). Moreover, it has been argued that confidence cannot be used to help discriminate between accurate and inaccurate judgments (see meta-analysis by DePaulo, Charlton, Cooper, Lindsay, & Muhlenbruck, 1997). Given that people can detect deception at a level barely greater than chance, perhaps there is little reason to expect that their confidence would be related to the accuracy of their judgments. Yet more recent research paints a somewhat more optimistic view of deception detection. Indeed, over the past 15 years, researchers have uncovered several procedures that can help people to discriminate between lie-tellers and truth-tellers (e.g., Hartwig, Granhag, Strömwall, & Vrij, 2005; Vrij et al., 2008). Beyond being important developments in their own right, to the extent that people can discriminate between lie-tellers and truth-tellers, signal detection theory predicts that confidence should help to distinguish between judgments that are likely to be accurate and those that are likely to be inaccurate.

In the current study, we demonstrated that—contrary to the conventional stance, which holds that confidence does not predict accuracy (Vrij, 2011)—decisions about whether someone is lying (i.e., *lie decisions*) made with high confidence are more accurate than lie decisions made with lower levels of confidence. The important relationship between confidence and accuracy might have previously gone unnoticed because of researchers' reliance on using between-subjects Pearson product–moment

Corresponding Author:
Andrew M. Smith, Department of Psychology, W112 Lagomarcino Hall, Iowa State University, Ames, IA 50011
E-mail: amsmith@iastate.edu

correlations. In the current study, we used a more sophisticated confidence–accuracy characteristic (CAC) approach (Mickes, 2015). The CAC approach is part of a family of calibration approaches that has revolutionized understanding of the confidence–accuracy relationship in eyewitness memory research (e.g., Brewer & Wells, 2006; Juslin, Olsson, & Winman, 1996; Wixted & Wells, 2017); we believe that it has the potential to do the same for the deception-detection literature. After introducing the CAC approach, we show that signal detection theory predicts a strong confidence–accuracy relationship. Finally, we reanalyze the data from three recent experiments conducted in the laboratory of author A.-M. Leach, and we show that high-confidence lie decisions are more accurate than low-confidence lie decisions.

Using Confidence to Predict Accuracy

In a typical deception-detection experiment, participants complete a number of trials. Some of the targets on these trials are lying, and others are being truthful. On each trial, the participant decides whether the target is lying or telling the truth, and many experiments also have the participant qualify each decision by indicating how confident they are that they have made a correct decision (for an overview, see Vrij, 2011). These confidence statements are often—but not always—made in terms of percentage confidence (e.g., 0%, 10%, . . . , 100%). But what does confidence actually reflect? One intuitively appealing possibility is that confidence is an indicator of the probability that the participant has made an accurate decision. To the extent that this is true, the participant is well calibrated. Perfect calibration occurs when 100% of decisions made with 100% confidence are accurate, 90% of the decisions made with 90% confidence are accurate, and so on. Figure 1 provides a graphical display of this perfect calibration.

We suspect that few readers would disagree if we were to characterize the perfect calibration in Figure 1 as a strong relationship between confidence and accuracy. Unfortunately, even if this strong confidence–accuracy relationship did exist in the deception-detection literature, researchers would be unlikely to discover it. Indeed, standard practice for measuring the confidence–accuracy relationship in deception research involves examining whether participants who make more accurate decisions tend to be more confident overall (DePaulo et al., 1997). Specifically, researchers calculate the percentage of trials on which participants made correct decisions and participants' average confidence across all trials. Then, they compute a Pearson product–moment correlation coefficient to determine whether participants with higher accuracy rates tend to be more confident in their decisions.

Correlating average confidence with the percentage of correct decisions provides information about whether

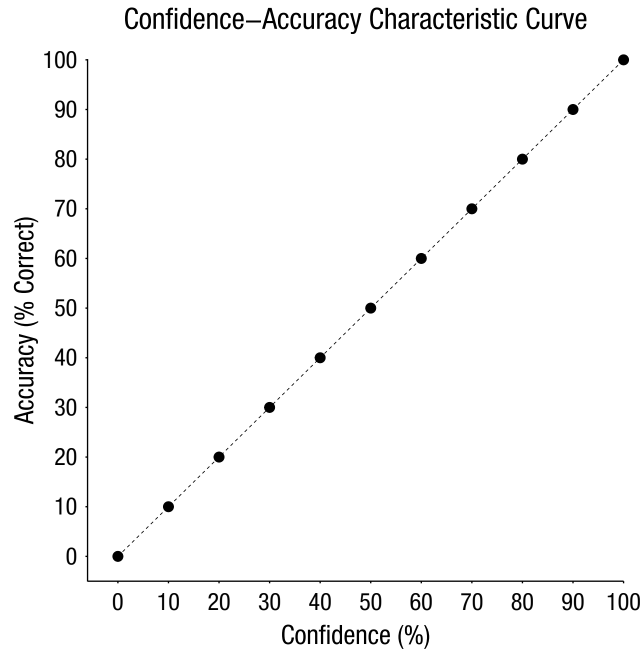

Fig. 1. Hypothetical confidence–accuracy calibration curve demonstrating perfect calibration between confidence and accuracy.

better judges tend to be more confident in their decisions but ignores the arguably more important question of whether higher confidence decisions are more accurate than lower confidence decisions. In the aggregate, people have difficulty detecting deception. But presumably there are instances in which a given participant will find the evidence for one response option (e.g., "lie") to greatly exceed the evidence for the alternative response option (e.g., "truth"), and in such instances, both the probability of making a correct decision and the participant's confidence in that decision should increase (Horry & Brewer, 2016; Smith & Vickers, 1988; Van Zandt, 2000). Simply correlating average confidence with the percentage of correct decisions does not provide any information on whether high-confidence decisions are more accurate than low-confidence decisions.

Fortunately, a family of calibration approaches can be used to shed light on the extent to which confidence might be used to discriminate between accurate and inaccurate decisions (e.g., Baranski & Petrusic, 1994; Brewer & Wells, 2006; Juslin et al., 1996; Mickes, 2015; Wixted & Wells, 2017). Each of these approaches involves plotting the percentage of accurate decisions against expressed level of confidence. As can be seen in Figure 1, for each confidence bin, researchers plot the percentage of accurate decisions that were made at that level of confidence. To the extent that confidence is positively associated with accuracy, the slope of the calibration curve will be monotonically increasing, and bins reflecting higher levels of confidence will be associated with higher levels of accuracy. Traditional calibration approaches are somewhat restrictive in that they require confidence to be measured on a percentage scale (e.g., 0%, 10%, . . . , 100%). The newly developed CAC approach is more flexible, however: It can be used to examine the confidence–accuracy relationship no matter how confidence is scaled (Mickes, 2015). Calibration analyses are uniquely situated to address a critical question that has remained unanswered in the deception detection literature to date: Are deception decisions made with high confidence more accurate than deception decisions made with low confidence?

Signal Detection Theory Predicts a confidence–Accuracy Relationship

Before using CAC analyses to measure the confidence–accuracy relationship from three recent deception-detection experiments, we now explain why, in theory, we should predict a strong association between confidence and accuracy. Our assumption here is that people have some ability to discriminate between lie-tellers and truth-tellers. Under some conditions, this assumption holds true. Under other conditions, it does not. When participants have no ability to discriminate between lie-tellers and truth-tellers (i.e., $d' \leq 0$), we see little reason to predict a confidence–accuracy relationship. But under conditions in which participants can discriminate between lie-tellers and truth-tellers (i.e., $d' > 0$), there are strong a priori reasons to believe that confidence should distinguish between likely accurate and likely inaccurate decisions (e.g., Mickes, Hwe, Wais, & Wixted, 2011; Wixted & Wells, 2017).

Figure 2 shows the signal-detection-based prediction for why high-confidence lie decisions should be more accurate than low-confidence lie decisions (for a

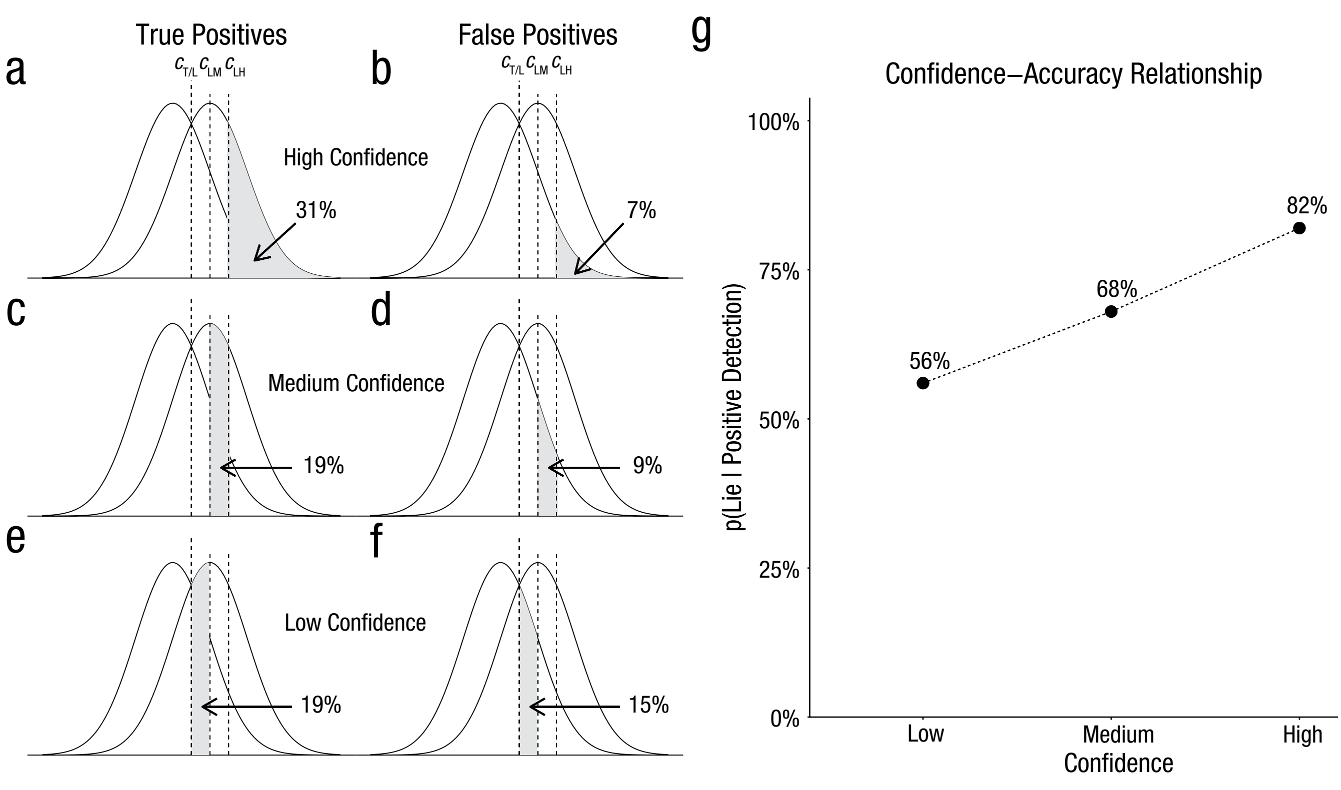

Fig. 2. True positives and false positives. The same signal detection model is displayed in panels (a) through (f). In each panel, the distribution on the left represents the signal strength (i.e., evidence of lying) emitting from truth-tellers, and the distribution on the right represents the signal strength emitting from lie-tellers. Because lie-tellers will tend to emit a stronger signal sensation than will truth-tellers, the lie-teller distribution is shifted to the right of the truth-teller distribution. In the present model, the lie-teller distribution is 1.00 standard units to the right of the truth-teller distribution ($d' = 1.00$). The three dashed vertical lines in each panel represent the decision criteria. The leftmost criterion (i.e., the longest vertical line in each panel) is the criterion for making a lie decision. If the evidence that the target person is lying exceeds this criterion, the participant makes a lie decision; otherwise, the participant makes a truth decision. The two criteria to the right of this represent the criteria for medium- and high-confidence lie decisions, respectively. If the evidence that the target person is lying exceeds the low-confidence lie criterion ($C_{T/L}$) but falls short of the medium-confidence lie criterion (c_{LM}), then the participant makes a lie decision with low confidence. If the evidence that the target person is lying exceeds the c_{LM} criterion but falls short of the high-confidence lie criterion (c_{LH}), then the participant makes a lie decision with medium confidence. If the evidence that the target person is lying exceeds the c_{LH} criterion, then the participant makes a lie decision with high confidence. Panels (a) through (f) illustrate one example of percentages of hits and false alarms made with high, medium, and low levels of confidence. Panel (g) is a confidence–accuracy characteristic curve showing the percentage of correct decisions that the model predicts for low, medium, and high levels of confidence. The positive confidence–accuracy slope demonstrates that signal detection theory naturally predicts a confidence–accuracy relationship.

similar, although slightly more complex, model in the eyewitness identification context, see Fig. 11 in Wixted and Wells, 2017). Note from the outset that the graphs in Figures 2a through 2f are identical in all respects, except that we have shaded different regions of the figure to highlight the percentages of true-positive decisions (hits) and false-positive decisions (false alarms) made at low, medium, and high levels of confidence, respectively. For simplicity, we assume that the strength of the lie signals (i.e., the evidence or cues that someone is lying) emanating from both truth-tellers and lie-tellers is normally distributed, with equal variances of 1.00. In each panel, the distribution on the left represents the strength of the lie signal emanating from truth-tellers, and the distribution on the right represents the strength of the lie signal emanating from lie-tellers. It makes sense that the lie-teller distribution is to the right of the truth-teller distribution (i.e., has a higher mean) because, on average, we would expect that lie-tellers would emit a stronger lie signal than would truth-tellers. The standardized difference between the means of these two distributions is d'. In Figure 2, d' is equal to 1.00 because the mean of the lie-teller distribution is 1.00 standard units higher than the mean of the truth-teller distribution.

Why does signal detection theory predict that high-confidence lie decisions will be more accurate than low-confidence lie decisions? Consider Figures 2a and 2b. These panels show the predicted percentages of hits and false alarms occurring with high confidence. It is relatively unlikely that someone who is telling the truth would emit a very strong lie signal, and so only a small percentage (7%) of the truth-telling distribution exceeds the high-confidence criterion. This is the false-positive rate. Someone who actually is lying would be far more likely to emit a strong lie signal; thus, a comparatively larger portion of the lie-telling distribution exceeds the high-confidence criterion (31%). This is the true-positive rate. Accordingly, under the previously specified parameters, the model predicts that 82% (31%/[31% + 07%]) of high-confidence lie detections will be accurate. Figures 2c and 2d show the percentages of correct and false detections that we would expect to occur with a medium level of confidence. Nineteen percent of the lie-telling distribution falls between the medium-confidence criterion and high-confidence criterion, and 9% of the truth-telling distribution falls between these two criteria. Hence, the model predicts a more modest 68% (19%/[19% + 09%]) accuracy rate at the medium-confidence level. Finally, Figures 2e and 2f show nearly equal percentages of the lie-telling (19%) and truth-telling (15%) distributions fall between the low-confidence and medium-confidence criteria. Hence, at low confidence, the model predicts accuracy to be only slightly greater than chance at 56% (19%/[19% + 05%]).

Signal detection theory also predicts that high-confidence truth decisions will be more accurate than low-confidence truth decisions. Figure 3 shows the signal-detection-based predictions for why high-confidence truth decisions will be more accurate than low-confidence truth decisions. Note that the model depicted in Figure 3 is identical to the model depicted in Figure 2, except that we are now showing criteria for making truth decisions with low, medium, and high levels of confidence, respectively.

As was the case with lie decisions, it is evident from examining Figures 3a through 3f that signal detection theory also predicts a strong confidence–accuracy relationship for truth decisions. For this particular model, signal detection theory predicts low-, medium-, and high-confidence truth decisions to be correct 56%, 68%, and 82% of the time, respectively (see Fig. 3g).

Confidence Can Discriminate Between Accurate and Inaccurate Lie Decisions

The signal-detection-based prediction laid out in the previous section provides strong a priori reason to expect that high-confidence decisions will be more accurate than low-confidence decisions. To test this model-generated prediction, we examined the confidence–accuracy relationship from three recent experiments carried out in A.-M. Leach's laboratory. Note that we measured the confidence–accuracy relationship separately for lie decisions and truth decisions. As we note above, truth judgments occupy different criteria than do lie judgments. Moreover, there is considerable evidence that confidence does far better at distinguishing between likely accurate and likely inaccurate *signal-present* decisions than it does at distinguishing between likely accurate and likely inaccurate *signal-absent* decisions (e.g., Brewer & Wells, 2006; Tekin & Roediger, 2017). Thus, we consider it most appropriate to analyze the confidence–accuracy relationship separately for lie judgments and truth judgments.

Figure 4 shows the three CAC curves—one for each experiment—for lie decisions. Because CAC curves bin accuracy by confidence, a large number of responses are required to produce stable estimates. Accordingly, in appropriate cases, we collapsed across conditions from the original experiment. Moreover, in line with previous research implementing CAC analysis, we divided confidence into three categories: < 70% confidence, 70% to 89% confidence, and 90% to 100% confidence. Because we are binning confidence in this way, note that we would not necessarily predict a linear relationship between confidence and accuracy. More generally, we would simply predict that accuracy is monotonically increasing with confidence.

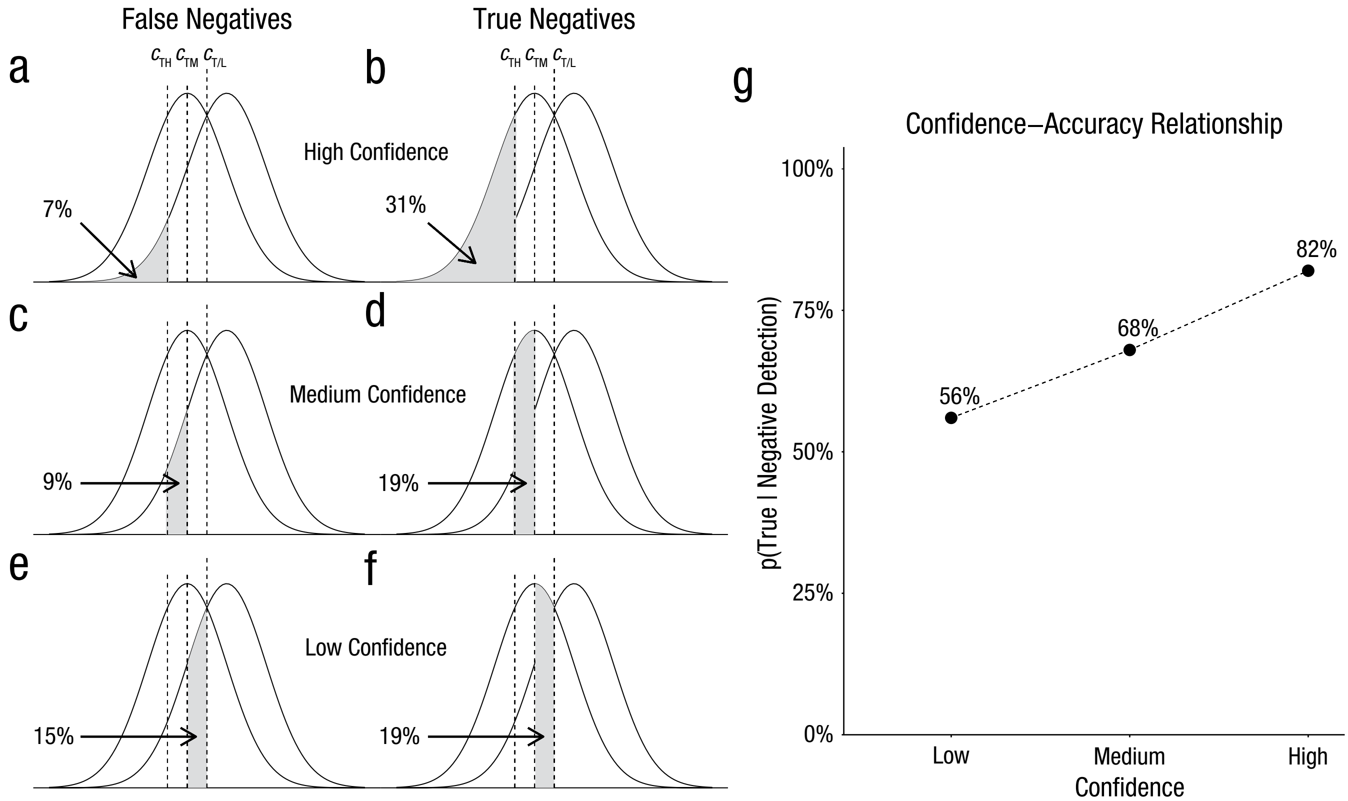

Fig. 3. True negatives and false negatives. The same signal detection model is displayed through each of panels (a) through (f). In each panel, the distribution on the left represents the signal strength (i.e., evidence of lying) emitting from truth-tellers, and the distribution on the right represents the signal strength emitting from lie-tellers. Because lie-tellers will tend to emit a stronger signal sensation than will truth-tellers, the lie-teller distribution is shifted to the right of the truth-teller distribution. In the present model, the lie-teller distribution is 1.00 standard units to the right of the truth-teller distribution (d' = 1.00). The three vertical lines in panels (a) through (f) represent the decision criteria. The rightmost criterion (i.e., the longest vertical line in each panel) is the criterion for making a truth decision. If the evidence that the target person is telling the truth falls short of this criterion, the participant makes a truth decision; otherwise, the participant makes a lie decision. The two criteria to the left of this criterion represent the criteria for medium-confidence and high-confidence truth decisions, respectively. If the evidence that the target person is telling the truth falls between the low-confidence truth criterion ($c_{T/L}$) and the medium-confidence truth criterion (c_{TM}), then the participant makes a truth decision with low confidence. If the evidence that the target person is telling the truth falls between the c_{TM} criterion and high-confidence truth criterion (c_{TH}), then the participant makes a truth decision with medium confidence. If the evidence that the target person is lying falls below the c_{TH} criterion, then the participant makes a truth decision with high confidence. Panels (a) through (f) show the percentages of true-negative and false-negative decisions made with high, medium, and low levels of confidence, respectively. Panel (g) is a confidence–accuracy characteristic curve showing the percentage of correct decisions the model predicts for low, medium, and high confidence truth decisions. The positive confidence–accuracy slope demonstrates that signal detection theory naturally predicts a confidence–accuracy relationship.

Before interpreting the data, it is important to provide context. The deception detection task was identical across experiments: First, participants viewed an equal number of lie- and truth-tellers. After each target was shown, participants indicated whether the target was lying or telling the truth; participants then indicated how confident they were in their decision on a scale from 0% to 100%. Figure 4a consists of data from Woolridge and Leach's (2019) work on misattributions—that is, inducing targets to believe that lighting could increase arousal or relaxation (vs. control)—in which 102 participants watched 20 targets being interviewed about evidence of a bomb plot. The data in Figure 4b also are from an experiment that used this bomb-plot paradigm (Elliott & Leach, 2016): Here, the experimenters studied 132 participants' ability to detect deception in 14 targets who spoke English at the basic, intermediate, advanced, or native level. Finally, the data in Figure 4c are from an experiment in which 232 participants viewed 20 women providing testimony about a mock theft while wearing a hijab or niqab or remaining unveiled (Study 1; Leach et al., 2016).[1]

We provide the overall accuracy rates (and d' values) along with the accuracy rates for lies and truths in Table 1. As noted at the outset of this article, a large meta-analysis concluded that the overall accuracy rate of deception decisions was 54% (Bond & DePaulo, 2006). This same large-scale meta-analysis estimated that the separate accuracy rates for lie and truth decisions were 47% and 61%, respectively (Bond & DePaulo, 2006).

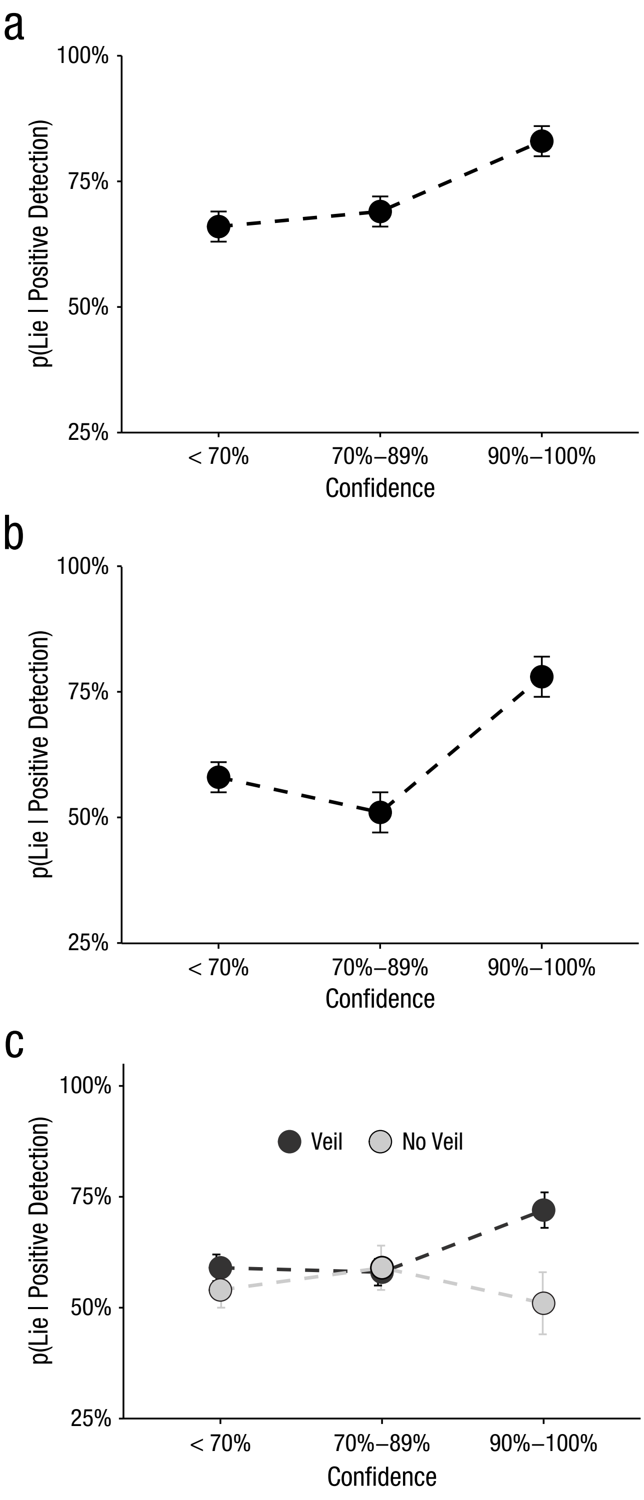

Fig. 4. Confidence–accuracy characteristic curves for lie decisions. The data depict the average accuracies of affirmative lie decisions made with low (< 70%), medium (70%–89%), and high (90%–100%) levels of confidence. The data from (a) Woolridge and Leach (2019) are collapsed over all conditions. The data from (b) Elliott and Leach (2016) are again collapsed over all conditions. Finally, the data from (c) Leach et al. (2016, Study 1) are collapsed over the two veiled conditions (nijab and hiqab), and we compare the veil and no-veil conditions.

Table 1. Accuracy Rates for the Three Experiments Reanalyzed in the Present Work

	Decision type			
Experiment	Overall	*d*′	Lies	Truths
Woolridge and Leach (2019)	66%	1.03	60%	73%
Elliott and Leach (2016)	57%	0.43	50%	64%
Leach et al. (2016, Experiment 1)				
Veiled participants	59%	0.40	62%	56%
Nonveiled participants	53%	0.11	55%	51%

The accuracy rates for the experiments we review in the present work were only slightly higher than these estimates. Indeed, the overall accuracy rate ranged from 53% to 66%, the lie-decision accuracy rate ranged from 50% to 62%, and the truth-decision accuracy rate ranged from 51% to 73%.

A brief note on how confidence was measured in these experiments is in order. Participants were not asked to indicate how confident they were that a target person was lying or telling the truth but rather to indicate how confident they were in their deception decision. This is an important distinction. On many signal detection tasks, participants are not asked to indicate how confident they are that they have made a correct decision but rather to indicate how confident they are that the signal is present (see Macmillan & Creelman, 2005). When a participant is asked to indicate how confident he or she is that the signal is present, a confidence rating of less than 50% is akin to saying that the participant believes the signal is absent. For example, if a participant says there is only a 20% chance that a target person is lying, he or she is also saying that there is an 80% chance that the target person is telling the truth. But that is not the confidence question that was asked in the experiments we review here, nor is it the question asked in most of the deception-detection literature. Participants in these experiments were first asked to indicate whether they believed the target person was lying or telling the truth and then to indicate how confident they were in that decision. In this context, a lie decision made with less than 50% confidence does not imply that the participant believes the target person is telling the truth. Indeed, if the participant believed that the target person was telling the truth, then the participant would have indicated so before providing his or her confidence statement. Technically speaking, the deception literature does not typically use a pure rating task in which binary classifications and confidence are made in a single decision with a Likert-type scale (e.g., 1 = *certain truth*, 6 = *certain lie*); rather, participants typically make a binary classification (e.g., the target is lying or telling the truth)

Table 2. Response Frequencies of Affirmative Lie Decisions for Each Confidence Bin Presented in Figure 4

	Confidence level			
Experiment and response	< 70%	70%–89%	90%–100%	Overall
Woolridge and Leach (2019)				
True positives	215	239	151	605
False positives	109	121	42	272
Elliott and Leach (2016)				
True positives	190	158	117	465
False positives	145	144	41	330
Leach et al. (2016, Experiment 1)				
Veiled participants				
True positives	251	228	153	632
False positives	187	180	62	429
Nonveiled participants				
True positives	110	91	45	246
False positives	102	76	44	222

Note: The frequencies in this table were aggregated across participants.

followed by a separate confidence statement (e.g., 0%, 10%, . . . , 100%).

With this clarification in mind, we now address the question of whether high-confidence lie decisions are more accurate than low-confidence lie decisions. In the experiments depicted in Figures 4a and 4b, the original conditions—reported lighting effects and targets' language proficiencies—have no conceptual importance for the current analyses and tended to yield similar levels of discriminability. Thus, we examined the confidence–accuracy relationship collapsed over all conditions. In Figure 4a, we can see that decisions made with low and moderate levels of confidence were accurate 66% and 69% of the time, respectively. However, decisions made with high confidence were much more accurate: Indeed, they were correct an impressive 83% of the time. The data in Figure 4b tell a similar story. Lie decisions made with low and medium levels of confidence were correct 58% and 51% of the time, respectively, with accuracy increasing to 78% for high-confidence decisions. Hence, the data in Figures 4a and 4b suggest that confidence might be an important tool for discriminating between accurate and inaccurate lie decisions.

To illustrate how the confidence–accuracy relationship is dependent on discriminability, we present the experimental data from two conditions in Figure 4c. Leach et al. (2016) found that participants were better able to discriminate between lie-tellers and truth-tellers who wore a veil (either a hijab or niqab) compared with those who were not veiled. Moreover, when target persons were not veiled, participants showed chance levels of discrimination. Because we predict a confidence–accuracy relationship to the extent that participants can discriminate between lie-tellers and truth-tellers, it should emerge in the veiled condition only. Indeed, that is precisely what we found. As shown in Figures 4a and 4b, when targets wore veils, high-confidence lie decisions (72%) were more accurate than low- or medium-confidence lie decisions (59% and 58%, respectively). When target persons were not veiled, confidence was unrelated to accuracy. These analyses provide compelling support for our signal-detection-based prediction that confidence in lie decisions would be related to the accuracy of lie decisions. The frequencies for each point on these CAC curves are provided in Table 2.

Next, we examined whether confidence could distinguish between likely accurate and likely inaccurate truth decisions. The CAC curves for truth decisions are depicted in Figure 5, and the frequencies associated with each point on those CAC curves are provided in Table 3. Here, confidence proved much less useful. For the data in Figure 5a, the relationship was as expected: Low-confidence decisions were correct 53% of the time, medium-confidence decisions were correct 67% of the time, and high-confidence decisions were correct 75% of the time. However, in Figure 5b, medium-confidence decisions were more accurate (68%) than either low-confidence (47%) or high-confidence (49%) decisions. In Figure 5c, there was virtually no relationship between confidence and accuracy. Speculation as to why confidence appears to have only a sporadic relationship with the accuracy of truth decisions is beyond the scope of the current article. But, it should be noted that this finding is consistent with results in both the eyewitness memory and basic recognition literatures (e.g., Brewer & Wells, 2006; Tekin & Roediger, 2017).

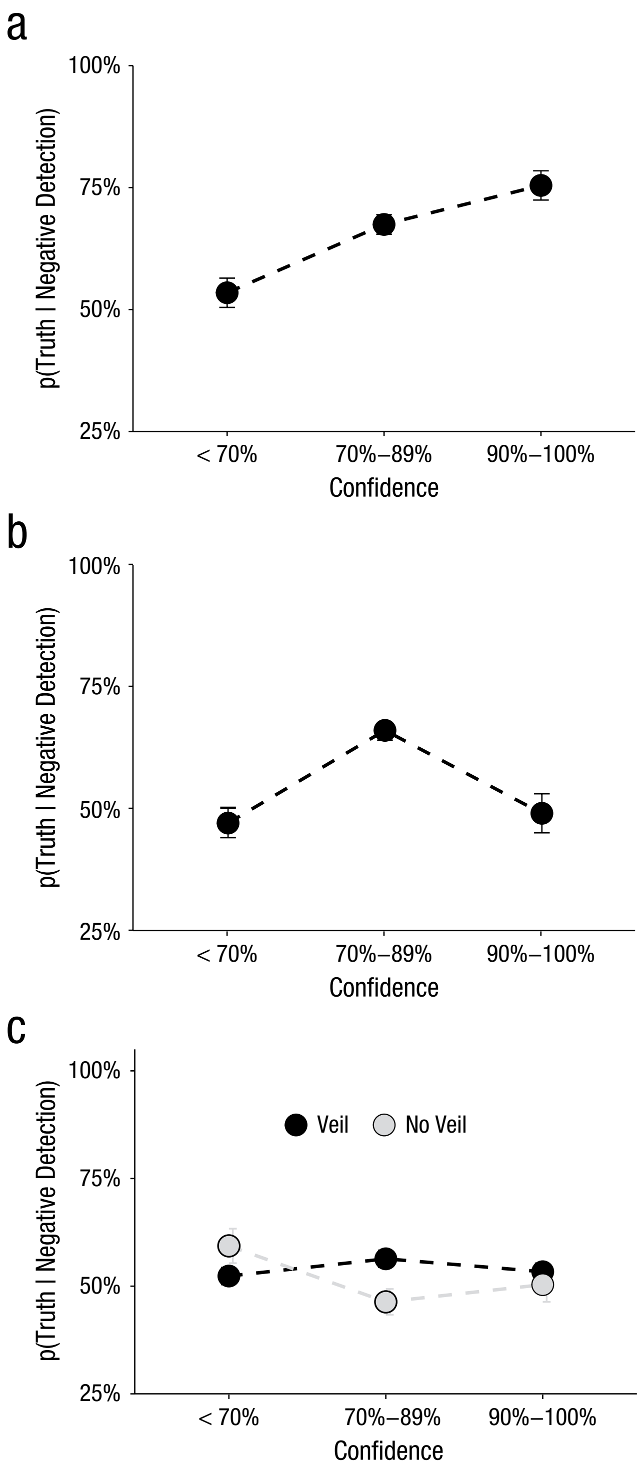

Fig. 5. Confidence–accuracy characteristic curves for truth decisions. The data depict the average accuracies of truth decisions made with low (< 70%), medium (70%–89%), and high (90%–100%) levels of confidence. The data from (a) Woolridge and Leach (2019) are collapsed over all conditions. The data from (b) Elliott and Leach (2016) are again collapsed over all conditions. Finally, the data from (c) Leach et al. (2016, Study 1) are collapsed over the two veiled conditions (nijab and hiqab), and we compare the veil and no-veil conditions.

Correlating Proportion Correct With Average Confidence

We conclude our analysis by demonstrating that the common practice of correlating percentage correct with average accuracy between participants does not lead one to the conclusion that high-confidence lie decisions are more accurate than low-confidence lie decisions. For this analysis, we calculated the percentage of correct decisions and average confidence for each participant and then correlated these two variables. These correlations hovered around zero in every single experiment. We also correlated confidence and accuracy separately for lie decisions and truth decisions. Again, all of the correlations hovered around zero. All correlations are presented in Table 4. The failure to find a correlation between percentage correct and average confidence is despite the fact that CAC analysis clearly demonstrates that high-confidence lie decisions are more accurate than low-confidence lie decisions.

But what if we had found a correlation between percentage correct and average confidence; would this imply that all is well in the deception-detection literature and that CAC analysis is not needed to examine the confidence–accuracy relationship? We think not. The rationale for collecting confidence statements is that confidence might be useful in discriminating between accurate and inaccurate detection decisions. But, correlating percentage correct with average confidence does not inform on this issue. Rather, it merely addresses the question of whether more confident participants also tend to be more accurate. Although that might be important in terms of personnel selection, it cannot tell us whether higher-confidence decisions are more accurate than lower-confidence decisions. That question is precisely the one that CAC analysis answers. CAC analysis is uniquely suited to measuring the confidence–accuracy relationship and has the potential to revolutionize understanding of the confidence–accuracy relationship in the deception detection literature.

A Word of Caution

We are not saying that high-confidence lie decisions should universally be interpreted as implying high accuracy. Such a statement is beyond the scope of what we can reasonably conclude from our analyses. We examined the confidence–accuracy relationship from three deception-detection experiments conducted in a laboratory setting. Whether this pattern of results generalizes beyond this context is an empirical question. What we can assert, however, is that correlating percentage correct and average confidence does not provide information regarding whether confidence can discriminate between

Table 3. Response Frequencies of Truth Decisions for Each Confidence Bin Presented in Figure 5

	Confidence level			
Experiment and response	< 70%	70%–89%	90%–100%	Overall
Woolridge and Leach (2019)				
True negatives	178	325	235	738
False negatives	131	168	106	405
Elliott and Leach (2016)				
True negatives	148	338	108	594
False negatives	158	189	112	459
Leach et al. (2016, Experiment 1)				
Veiled participants				
True negatives	311	454	349	1,114
False negatives	247	368	296	911
Nonveiled participants				
True negatives	139	202	205	546
False negatives	146	217	160	523

Note: The frequencies in this table were aggregated across participants.

accurate and inaccurate deception decisions, but CAC analysis does. Contrary to long-standing belief that confidence does not discriminate between accurate and inaccurate decisions, our CAC analyses clearly show that high-confidence lie decisions are more accurate than low-confidence lie decisions.

Conclusions

The present results correct a long-standing belief in the deception-detection literature that confidence is unrelated to accuracy (e.g., DePaulo et al., 1997; Vrij, 2011). Consistent with the predictions of signal detection theory, our primary conclusion is that confidence *can* discriminate between accurate and inaccurate lie detections. Our analyses suggest that there are three boundary conditions to this conclusion. First, confidence was better at discriminating between accurate and inaccurate signal-present (lie) decisions than it was at discriminating between accurate and inaccurate signal-absent (truth) decisions. This is consistent results from other literatures. Second, confidence can discriminate between accurate and inaccurate decisions only to the extent that people can discriminate between lie-tellers and truth-tellers. Third, and perhaps most important, confidence can discriminate between accurate and inaccurate lie detections only to the extent that researchers use an appropriate analytical tool. Correlating average confidence with the percentage of correct decisions between participants does not inform on whether decisions made with higher confidence are more accurate than decisions made with lower confidence. The CAC approach that has taken center stage in the eyewitness literature is uniquely situated to evaluate whether high-confidence deception decisions tend to be more accurate than low-confidence deception decisions. Our data are consistent with findings in the eyewitness literature, the basic-recognition literature, and many other domains of human judgment and decision making: Confidence can discriminate between accurate and inaccurate deception decisions.

Table 4. Pearson Product–Moment Correlations Between Proportion of Correct Decisions and Average Confidence

	Decision type		
Experiment	All	Lies	Truths
Woolridge and Leach (2019)	−.15	−.08	−.24
Elliott and Leach (2016)	.04	−.11	−.07
Leach et al. (2016, Experiment 1)			
Veiled participants	.00	−.11	−.02
Nonveiled participants	.12	−.12	.15

Action Editor

Laura A. King served as action editor for this article.

ORCID iD

Andrew M. Smith https://orcid.org/0000-0002-4184-9364

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Note

1. A brief note on this veiling effect: Leach et al. (2016) found that participants were able to discriminate between lie-tellers and truth-tellers when targets were veiled but not when targets were nonveiled. This may come as a surprise to the generalist reader who is unfamiliar with the deception literature; however, this finding is consistent with a strong theoretical foundation. Indeed, one reason why people have difficulty discriminating between lie-tellers and truth-tellers in the first place is because they rely on unreliable facial cues (e.g., DePaulo et al., 2003). By veiling target persons, the researchers prevented participants from using these potentially erroneous cues, and the result was better discrimination. Moreover, Leach et al. (Experiment 2) directly replicated this experiment with a European sample (the original sample was Canadian) and reached the same conclusion.

References

Baranski, J. V., & Petrusic, W. M. (1994). The calibration and resolution of confidence in perceptual judgments. *Perception & Psychophysics*, *55*, 412–428. doi:10.3758/BF03205299

Bond, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review*, *10*, 214–234. doi:10.1207/s15327957pspr1003_2

Brewer, N., & Wells, G. L. (2006). The confidence-accuracy relationship in eyewitness identification: Effects of lineup instructions, foil similarity, and target-absent base rates. *Journal of Experimental Psychology: Applied*, *12*, 11–30. doi:10.1037/1076-898X.12.1.11

DePaulo, B. M., Charlton, K., Cooper, H., Lindsay, J. J., & Muhlenbruck, L. (1997). The accuracy-confidence correlation in detection of deception. *Personality and Social Psychology Review*, *1*, 346–357. doi:10.1207/s15327957pspr0104_5

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin*, *129*, 74–118. doi:10.1037/0033-2909.129.1.74

Elliott, E., & Leach, A.-M. (2016). You must be lying because I don't understand you: Language proficiency and lie detection. *Journal of Experimental Psychology: Applied*, *22*, 488–499. doi:10.1037/xap0000102

Hartwig, M., Granhag, P. A., Strömwall, L. A., & Vrij, A. (2005). Detecting deception via strategic disclosure of evidence. *Law and Human Behavior*, *29*, 469–484. doi:10.1007/s10979-005-5521-x

Horry, R., & Brewer, N. (2016). How target-lure similarity shapes confidence judgments in multiple-alternative decision tasks. *Journal of Experimental Psychology: General*, *145*, 1615–1634. doi:10.1037/xge0000227

Juslin, P., Olsson, N., & Winman, A. (1996). Calibration and diagnosticity of confidence in eyewitness identification: Comments on what can be inferred from the low confidence-accuracy correlation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 1304–1316. doi:10.1037/0278-7393.22.5.1304

Leach, A.-M., Ammar, N., Kleinberg, B., Verschuere, B. J., England, D. N., & Remigio, L. M. (2016). Less is more? Detecting lies in veiled witnesses. *Law and Human Behavior*, *40*, 401–410. doi:10.1037/lhb0000189

Macmillan, N. A., & Creelman, D. C. (2005). *Detection theory: A user's guide* (2nd ed.). Mahwah, NJ: Erlbaum.

Mickes, L. (2015). Receiver operating characteristic analysis and confidence-accuracy characteristic analysis in investigations of system variables and estimator variables that affect eyewitness memory. *Journal of Applied Research in Memory and Cognition*, *4*, 93–102. doi:10.1016.j.jarmac.2015.01.003

Mickes, L., Hwe, V., Wais, P. E., & Wixted, J. T. (2011). Strong memories are hard to scale. *Journal of Experimental Psychology: General*, *140*, 239–257. doi:10.1037/a0023007

Smith, P. L., & Vickers, D. (1988). The accumulator model of two-choice discrimination. *Journal of Mathematical Psychology*, *32*, 135–168. doi:10.1016/0022-2496(88)90043-0

Tekin, E., & Roediger, H. (2017). The range of confidence scales does not affect the relationship between confidence and accuracy in recognition memory. *Cognitive Research: Principles and Implications*, *2*, 49–62. doi:10.1186/s41235-017-0086-z

Van Zandt, T. (2000). ROC curves and confidence judgments in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 582–600. doi:10.1037/0278-7393.26.3.582

Vrij, A. (2011). *Detecting lies and deceit: Pitfalls and opportunities* (2nd ed.). West Sussex, England: Wiley.

Vrij, A., Mann, S. A., Fisher, R. P., Leal, S., Milne, R., & Bull, R. (2008). Increasing cognitive load to facilitate lie detection: The benefit of recalling an event in reverse order. *Law and Human Behavior*, *32*, 253–265. doi:10.1007/s10979-007-9103-y

Wixted, J. T., & Wells, G. L. (2017). The relationship between eyewitness confidence and identification accuracy: A new synthesis. *Psychological Science in the Public Interest*, *18*, 10–65. doi:10.1177/1529100616686966

Woolridge, L., & Leach, A.-M. (2019, June). *Shedding light on deception: The effect of manipulating attributions of arousal on deception detection*. Paper Presented at the Society for Research in Memory and Cognition's Biennial Conference, Cape Cod, MA.