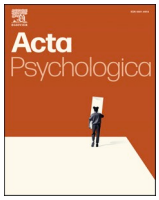
Acta Psychologica 263 (2026) 106278

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy



Assessing lying about opinions through extending the devil's advocate interview protocol with a same example statement

Sharon Leal [a], Aldert Vrij [a,*], Haneen Deeb [a], Dora Giorgianni [a], Lara Geermann [b], Ronald P. Fisher [c]

[a] School of Psychology, Sport & Health Sciences, University of Portsmouth, UK
[b] Behavior and Social Sciences, IE University, Spain
[c] Department of Psychology, Florida International University, USA

ARTICLE INFO

Keywords:
Same example statement
Deception
Interviewing
Opinions
Eloquence

ABSTRACT

The devil's advocate interview protocol is a method to detect lies about opinions. Interviewees are asked to provide reasons that support their opinion (eliciting opinion request) and reasons that oppose their opinion (devil's advocate request). Truth tellers are expected to be more eloquent, passionate and less predictable than lie tellers in responding to the eliciting opinion request. The difference in eloquence, passion and predictability between the two requests is expected to be more pronounced in truth tellers than in lie tellers. In the present experiment we examined whether adding a same example statement to the protocol would increase the ability to distinguish between truth tellers and lie tellers. In the same example statement condition, participants watched an example in which someone eloquently expressed their opinion about the same topic and opinion that the interviewee was going to discuss in the interview. We predicted the same example model statement to enhance the efficacy of the devil's advocate interview protocol.

A total of 171 truth tellers and lie tellers expressed their true opinions or false opinions about societal issues (e.g. covid vaccinations). Half of them watched a same opinion statement. The hypotheses were supported for eloquence and passion.

1. The devil's advocate interview protocol

Detecting lies about opinions can be important. For example, authorities want to know whether their government employees falsely claim to support them, or whether they are considering terror or espionage related acts.

The Devil's Advocate Approach is an interview protocol designed to detect lies about opinions (Leal, Vrij, Mann, & Fisher, 2010). It contains three stages (Vrij, Leal, & Fisher, 2023). The first stage is an *expression of attitude* request where interviewees are asked whether they are in favour or against a specific attitude-topic, for example: "Do you believe that the Conservative government did what was best for the country?" The second stage is an *eliciting-opinion request* where interviewees are asked to provide all the reasons they can think of that led them to having this opinion: "OK, you say that you do believe that the Conservative government did what was best for the country". Please can you tell me all the reasons you can think of that led you to having this opinion?". The third stage is a *devil's advocate request* where interviewees are asked to play devil's advocate -to imagine that they do not have this view at all- and to give all the reasons they can think of that go against the view they have and just expressed: "OK, now try to play the devil's advocate and imagine that you are opposed to the view that the Conservative government did what was best for the country. Please tell me all the reasons you can think of why the Conservative government did not do what was best for the country?"

Truth tellers (those who in the example believe the Conservative government did what was best for the country) are telling the truth in both replies. They express their real opinion in the eliciting-opinion request. They express a view opposite to their opinion in the devil's advocate request but since they do not pretend to have that view, they are telling the truth. Lie tellers (those who in the example believe that the Conservative government did not do what was best for the country) are lying in both replies. They present a view they do not have in the eliciting-opinion request. They express their real views in the devil's

* Corresponding author at: School of Psychology, Sport & Health Sciences, University of Portsmouth, King Henry Building, King Henry 1 Street, PO1 2DY, Hants, Portsmouth, UK.
E-mail address: aldert.vrij@port.ac.uk (A. Vrij).

https://doi.org/10.1016/j.actpsy.2026.106278
Received 12 August 2025; Received in revised form 9 December 2025; Accepted 16 January 2026
Available online 21 January 2026

advocate request, but since they pretend not to have these views, they are still lying.

A recent Devil's Advocate Approach experiment (Leal, Vrij, Deeb, & Fisher, 2024) found that such lies can be better detected if, at the beginning of the interview, interviewees listened to an example in which someone eloquently expressed their opinion about a topic unrelated to the topic of investigation (a so-called model statement). In the present experiment we extended Leal et al. (2024). We used an identical procedure with two differences, both in the model statement. First, in the present experiment examinees watched (rather than listened to) an example. Second, participants watched an example in which someone eloquently expressed their opinion about the same (rather than a different) topic and opinion that the interviewee was going to discuss in the interview (so-called same example statement). We compared this same example statement condition with a no example condition.

1.1. Theoretical underpinning of the expected veracity differences in a devil's advocate interview

Truth tellers and lie tellers are expected to respond differently to the eliciting-opinion and devil's advocate requests. Truth tellers should find it easier to respond to the eliciting-opinion request than to the devil's advocate request. Arguments that support someone's attitude (eliciting-opinion request) are typically more readily available than arguments that oppose someone's attitude (devil's advocate request) (Ajzen, 2001). This is the result of a confirmation bias, people's tendency to seek information that confirms rather than disconfirms their views (Darley & Gross, 1983; Hart et al., 2009; Nickerson, 1998). When people have a strong belief, information contradicting this belief often evokes a strong negative emotion (Mochon & Schwartz, 2024). Avoiding exposure to this belief-inconsistent information is one way to avoid or reduce experiencing negative emotions. Those who want to form an entirely balanced view must search all the available information. This is impossible because there is too much information available about an attitude-topic (Sommer, Musolino, & Hemmer, 2024). People must therefore decide which information to search and tend to search for information that support their views. To explain this, Sommer et al. (2024) use as example someone who does not have their car keys but believes that they are in the kitchen. It makes then sense for this person to start looking for their keys in the kitchen rather than in the lounge. The confirmation bias makes people more practiced in thinking about arguments that confirm their views than arguments that oppose their views.

Another explanation why truth tellers should find it easier to respond to the eliciting-opinion request than to the devil's advocate request is that arguments confirming people's views are more compatible with their other related views and beliefs than arguments disconfirming their views (Knobloch-Westerwick, 2015; Knobloch-Westerwick, Mothes, & Polavin, 2020). Truth tellers may refer to those compatible views and beliefs when responding to the eliciting-opinion request.

Since pro-arguments are more readily available in people's minds and processed at a deeper level than anti-arguments, truth tellers should be able to express their pro-arguments more eloquently (the collective term for plausibility, immediacy, and clarity) than their anti-arguments. The lack of deep thinking about their anti-arguments should make these arguments sound more predictable than their pro-arguments. Strong beliefs about a topic could result in one more effect. Strongly held beliefs often become a fundamental part of a person's identity and worldview (Vallerand, 2015). People therefore often feel intense passion about these strong beliefs which becomes evident when talking about them (Vallerand, 2015).

The differences between the eliciting-opinion and devil's advocate requests are expected to be less pronounced in lie tellers. Since they express their true views in the devil's advocate request, they could, in theory, be more eloquent, more passionate and less predictable when responding to the devil's advocate request than to the eliciting-opinion request. Yet this is unlikely to happen, because lie tellers may fear that this will give their lies away. Lie tellers often try to be consistent to come across as sincere (Deeb et al., 2017), and consistency in a devil's advocate context would be to give responses to both requests that sound similar in terms of eloquence, passion and predictability (Deeb et al., 2018).

Lie tellers often prepare themselves for interviews by thinking of answers to questions they can possibly be asked (Clemens, Granhag, & Strömwall, 2013; Granhag & Hartwig, 2008). In an interview about opinions a request to explain why they held a certain belief is likely to be anticipated and lie tellers will therefore in preparation of the interview think about arguments that support their alleged opinion. Although preparation typically enhances the eloquence of a response (DePaulo et al., 2003; Sporer & Schwandt, 2006), their response is still likely to fall short in terms of eloquence than the answers of those who have deeply thought about the topic (truth tellers). Since lie tellers do not 'own' the belief they express, their responses to the eliciting-opinion request should also sound more predictable and less passionate than truth tellers' responses.

1.2. Previous devil's advocate interview research

Devil's advocate deception research published to date (Leal et al., 2010; Leal et al., 2024; Leal, Vrij, Deeb, & Dabrowna, 2023; Mann, Vrij, Deeb, & Leal, 2022, 2023) supports these two predictions: (i) the difference in eloquence, passion and predictability in responses to the eliciting-opinion and devil's advocate requests is more pronounced in truth tellers than in lie tellers and (ii) Truth tellers are more eloquent, more passionate, and less predictable than lie tellers when replying to the eliciting-opinion request. See Vrij, Leal, and Fisher (2023) for an overview of devil's advocate deception research.

In their devil's advocate deception experiment, Leal et al. (2024) let half of the truth tellers and half of the lie tellers listen to a model statement, an example of an eloquent response to an eliciting-opinion request about a topic that was unrelated to the topics the interviewees were going to discuss (Leal, Vrij, Warmelink, Vernham, & Fisher, 2015). To date, a model statement is always about a topic other than the topic the interviewees discuss in the interview because the aim of the model statement is to give interviewees an idea about how much information they should provide rather than coaching them what to say (Vrij, Leal, & Fisher, 2018). A model statement raises expectations about how much information someone needs to provide (Ewens et al., 2016; Leal et al., 2024; Vrij et al., 2018). It works better than an instruction to be detailed because examples (e.g. model statement) are easier to follow than instructions. Examples provide concrete demonstrations of how to apply the rules, whereas instructions often require interpretation and can be less clear (Vrij et al., 2018). Since truth tellers' reasons to support their opinion are processed at a deeper level than the reasons that oppose their views, they should find it easier to elevate their eloquence levels to the eliciting-opinion request than to the devil's advocate request. Lie tellers should find it easier to raise the eloquence level in the devil's advocate than the eliciting-opinion request but may be reluctant to do so out of fear that it raises suspicion. In alignment with this reasoning, Leal et al. (2024) found that introducing the model statement enlarged the veracity effects in (i) responding to the eliciting-opinion request and (ii) differences in responses to the eliciting-opinion and devil's advocate requests.

1.2.1. The present experiment

In real life, when considering an opinions statement, it could be advantageous to align the example with the opinion and topic that the interviewee will discuss. When people with strong opinions are confronted with information that opposes their own view, the result is often a strong negative emotional response in the form of outrage (Mochon & Schwartz, 2024; Smith & Ellsworth, 1985). This is relevant for a devil's advocate interview. To facilitate lie detection the interviewer should

avoid presenting information that will outrage a truth teller, because a negative emotion has a negative effect on people's willingness to talk and on people's memory recall (Vrij et al., 2017). In contrast, outraging lie tellers (which could be the result of being exposed to their pretended opinion) may be beneficial. They should avoid showing negative emotions by controlling their demeanour. This could have a detrimental effect on their speech content because paying attention simultaneously to demeanour and speech is mentally taxing (Patterson, 1995, 2006).

In the present experiment half of the interviewees watched an actor talking about the same topic the interviewee was going to discuss and where the actor provides arguments in favour of the opinion the interviewee claims to have, the other half were not shown any example but just asked to provide all reasons that led them to having that opinion. To make it distinctive from a model statement, where actors always talk about different topics than those discussed by interviewees, we label this the 'same example statement.' In the present experiment the same example statement contains arguments truth tellers likely agree with. Since pro-arguments are processed relatively deeply, truth tellers should be able to remember them and elaborate on them. Hearing these pro-arguments could also trigger thoughts about pro-arguments not mentioned in the same example statement. The result is that truth tellers can use this same example statement to enrich their replies to the eliciting-opinion request. Since the same example statement is not related to the devil's-advocate request, it should not affect their responses to that request. The same example statement may also increase lie tellers' eloquence in answering the eliciting-opinion question, but probably not to the same extent as that of truth tellers because they have not processed those reasons as deeply as truth tellers. Neither will the same example statement enlarge in lie tellers the difference in eloquence, passion and predictability between the eliciting-opinion and devil's advocate question as they will strive to achieve consistency in responding to both requests. The same example statement was video recorded. People generally prefer visual to auditory channels when learning (Jawed, Amin, Malik, & Faye, 2019).

1.2.2. Hypotheses

The experiment was pre-registered at: https://osf.io/w4ke5/?view_only=6adc61fc36a94e08af058c9613ad4e5a. We tested the same four hypotheses as Leal et al. (2024):

- Truth tellers will be more eloquent (more plausible, more immediate and clearer), more passionate, and less predictable than lie tellers when replying to the eliciting-opinion request (Veracity main effect, Hypothesis 1).

- Truth tellers will be more eloquent, more passionate, and less predictable in response to the eliciting-opinion request than to the devil's advocate request. These differences will be less pronounced in lie tellers (Veracity x Type of Question interaction effect, Hypothesis 2).

- The differences in eloquence, passion, and predictability between truth tellers and lie tellers will be more pronounced in the same example statement-present condition than in same example statement-absent condition (Veracity x Same Example Statement interaction effect, Hypothesis 3).

- Participants in the same example statement-present condition will provide more arguments, and be more eloquent, more passionate, and less predictable than participants in the same example statement-absent condition (Same Example Statement main effect, Hypothesis 4).

1.2.3. Strategies

There are two differences between the pre-registration and the manuscript, both issues will be explored: First, strategies used by participants to sound convincing (this section); and second, video vs transcript coding (next section). We explored in the experiment the strategies participants said that they were going to use in the interview (in a pre-interview questionnaire) and the strategies they said to have used in the interview (in a post-interview questionnaire). Measuring strategies is important because it gives valuable insight into truth tellers' and lie tellers' thinking (DePaulo et al., 2003; Vrij & Granhag, 2012). If these thought processes differ from each other, interview protocols could be designed that highlight those differences and subsequently reveal or enlarge veracity effects (Vrij & Granhag, 2012). Truth tellers and lie tellers use different strategies when they report alleged activities (Granhag & Hartwig, 2008; Hartwig, Granhag, & Strömwall, 2007; Vrij et al., 2022). Truth tellers are forthcoming and tell it all whereas lie tellers avoid reporting incriminating evidence and prefer to keep their stories simple. Interview protocols that exploit these differences when discussing alleged activities include Assessment Criteria Indicative of Deception (ACID, Colwell, Hiscock-Anisman, Memon, Taylor, & Prewett, 2007); Cognitive Credibility Assessment (CCA, Vrij, Meissner, et al., 2017; Vrij, Mann, Leal, & Fisher, 2021), the Strategic Use of Evidence (Hartwig & Granhag, 2023; Oleszkiewicz & Watson, 2021) and the Verifiability Approach (Nahari, 2019; Palena, Caso, Vrij, & Nahari, 2021). See Vrij et al. (2022) for a review of these protocols. Although truth tellers may also use their being forthcoming strategy when discussing opinions, lie tellers' favourite strategies, particularly avoiding incriminating evidence, could be more difficult to generalise from discussing alleged activities to discussing opinions.

1.2.4. Video versus transcript coding

In all devil's advocate lie detection research published to date, eloquence and predictability were measured via coding the transcribed interviews. Although this is the standard practice in coding speech content (Vrij, 2008), it would be beneficial for real life application to code these variables in real time during the interviews. We therefore introduced video coding in addition to transcript coding: Raters coded eloquence and predictability while watching the videos. We report these video analyses in Appendix 4, but summarise the main results in the Results section and compare the text and video analyses in the Discussion.

2. Method

2.1. Design

We replicated the analyses presented by Leal et al. (2024). The present experiment employed a 2 (Veracity) x 2 (Type of Question) x 2 (Same Example Statement) mixed factorial design. Veracity and Same Example Statement were between-subjects factors and Type of Question was a within-subjects factor. The dependent variables were eloquence, predictability and passion. Significant Veracity x Type of Question interaction effects can be analysed in two different ways. The first way is to examine the differences in responses to the two types of question in truth tellers and lie tellers separately and decide in which Veracity condition (truth tellers or lie tellers) the effects are most pronounced (comparing *d*-scores). The second way is to examine the Veracity effects on residue scores (the difference in responses to the eliciting opinion and devil's advocate requests). Following Leal et al. (2024) we used this second option because it examines the interaction effects directly (a significant effect means that the residue score differed between truth tellers and lie tellers). Previous Devil's Advocate Approach research also reported residue scores (Leal et al., 2024; Vrij, Leal, & Fisher, 2023). To create residue scores, the scores for the devil's advocate request responses were deducted from the scores for the eliciting-opinion request responses.

Again, following Leal et al. (2024), we conducted a 2 (Veracity) x 2 (Same Example Statement) MANOVA with pro-alleged opinion and anti-alleged opinion arguments as dependent variables to test Hypothesis 3. Since Hypothesis 3 refers to a Same Example Statement main effect, including the Veracity factor would not be necessary. Following Leal et al. (2024), we included Veracity as a second factor because the results of this factor are relevant for this experiment that focuses on deception.

2.2. Participants

A G*Power analysis revealed that we needed at least 164 participants to ensure a statistical power of 0.90, an alpha level of 0.05, and a medium effect size ($f = 0.3$). The medium effect size choice was based on previous devil's advocate deception research (Vrij, Leal, & Fisher, 2023) and on real life application. To make a veracity assessment tool useful in real life, veracity differences should be clearly noticeable for which at least medium effect sizes are required (Cohen, 1988; Rice & Harris, 2005).

A total of 171 participants took part of whom 122 were female, 42 were male, one was non-binary and six did not say. The average age in the sample was $M = 23.27$ years ($SD = 8.20$). Most participants ($n = 64$) were from the United Kingdom; the other participants were other white ($n = 24$), other black ($n = 4$), Indian ($n = 21$), African ($n = 21$), Chinese ($n = 6$), Bangladeshi ($n = 5$), Pakistani ($n = 4$), other Asian ($n = 5$), mixed ($n = 6$), other ($n = 2$) or did not say ($n = 9$). For 125 participants English was the first language whereas this was not the case for 44 participants (two participants did not say). We examined the effect of language on the main dependent variables (eloquence, predictability, passion and number of arguments reported). The effects of language for eloquence, passion and number of arguments reported were not significant (all $Fs < 0.25$, all $ps > 0.623$) but the effect for predictability was significant, $F(1, 167) = 7.38$, $p = .007$, $d = 0.47(0.12,0.81)$, $BF_{10} = 5.164$. The replies of the native English-speaking participants sounded more predictable ($M = 4.41$, $SD = 0.94$, 95%CI [4.25,4.58]) than the answers from the other group ($M = 3.97$, $SD = 0.93$, 95%CI [3.69,4.24]). We return to this finding in the Discussion. The experiment conformed with the principles of the Declaration of Helsinki and ethics approval was granted by the University faculty's ethics committee (C-2023-087 A) and the sponsor's ethics committee (2024-11_713–23).

2.3. Procedure

2.3.1. Recruitment

We used the same procedure as Leal et al. (2024). We recruited participants via the university web portals for staff and students and also contacted those who had expressed a desire to be informed about upcoming studies. The recruitment advert -entitled 'In my opinion you're wrong! Tell us about your opinions on societal issues'- mentioned that participants would be interviewed online to discuss their opinions about a societal issue. Truth tellers would express their truthful opinions whereas lie tellers would express an opinion that opposes their true opinion. It was also mentioned that those who took part in Leal et al. (2024) could not take part again (none of the participants in the present experiment did).

2.3.2. Opinion questionnaire

An interview date was arranged for those who volunteered to take part. We emailed them at least 24 h before the interview date the participant information sheet, the consent form, and (via Gorilla) a 6-item opinion questionnaire. The questionnaire listed six controversial topics: COVID-19 vaccinations, CCTV-footage, abortion, immigrants, previous Conservative Government and the Israel / Palestine conflict. Participants were asked to give their true opinion on 7-point Likert scales ranging from 1 *(totally disagree)* to 7 *(totally agree)*. Five of the six topics were also used by Leal et al. (2024) but the item about the Israel / Palestine conflict was new. See Appendix 1 for the list of topics.

2.3.3. Experimental manipulations and instructions given to participants prior to the interview

The interviews took place via Zoom. The experimenter discussed with the participants the consent form and participant information sheet. All participants were happy to continue and signed the consent form. The experimenter then checked whether the participant had completed the 6-item opinion questionnaire. They were asked to complete it immediately if they had not done so.

Allocation to a truth teller or lie teller condition occurred randomly. The experimenter selected one item with the strongest support (6 or 7) or opposition (1or 2) on the opinion questionnaire for the interview. Truth tellers ($n = 87$) were asked to discuss their opinion about this item truthfully during the interview. Lie tellers ($n = 84$) were asked to pretend that they have the opposite opinion and to discuss that false opinion. Participants were then given as much time as they wished to prepare for the interview. We gave them this opportunity as it reflects real life: interviewees (particularly lie tellers) prepare themselves for possible interviews (Clemens et al., 2013). Participants were told about the importance to come across as convincing during the interview. If the interviewer believed their stories, they would be entered in a prize draw with vouchers worth £50, £100 and £150. If the interviewer did not believe them, they would have to prepare a written statement about their opinion they and would not be entered in the prize draw. In fact, all participants were entered in the draw, and nobody had to write a statement.

Whilst the participants prepared themselves for the interview, the experimenter informed the interviewer which opinion to discuss in the interview without mentioning the veracity status of the participant. The experimenter also informed the interviewer to conduct a same example statement-absent ($n = 85$) or a same example statement-present ($n = 86$) interview. Allocation to these two same example statement conditions occurred randomly.

2.3.4. Same example statements

The same example statements were video recordings. We prepared twelve video recordings where an actor expressed pro or anti arguments about each of the six topics listed on the opinion questionnaire (the same actor was used for all twelve video recordings). The participant always saw a video recording in which the actor discussed the topic and the position on the topic they were going to discuss. For example, if the participant was going to discuss why they were in favour of Covid vaccinations, they saw a same example statement video recording where the actor spoke in favour of Covid vaccinations. Appendix 2 provides the full texts of the pro and anti-Covid vaccination statements. The arguments are presented in italics and numbered (13 arguments in both statements). We prepared the pro- and anti- arguments based on news reports and on arguments frequently provided by participants in Leal et al. (2024). In other words, the arguments we presented were common (well-known) arguments. The twelve same example statements were all similar in length (about 230 words) and we included between 10 and 15 arguments in each of the twelve same example statements. In each same example statement the actor started a new argument before cutting off the video recording. This gives the impression that there is much more to say about the topic than the actor did say.

2.3.5. Pre-interview questionnaire

When the participants indicated to be ready for the interview, they completed a pre-interview questionnaire measuring demographics (gender, age, nationality) and preparation. Preparation thoroughness was measured with three items −1 (*shallow*) to 7 (*thorough*); 1 (*insufficient*) to 7 (*sufficient*); and 1 (*poor*) to 7 (*good*)- and the responses were averaged (Cronbach's alpha = 0.90). Preparation time was measured with a single item: 'Do you think the amount of time you were given to prepare was: 1 (*insufficient*) to 7 (*sufficient*).' We included preparation time because participants may have felt pressured not to take up too much preparation time in the experiment. The same items also appeared on Leal et al. (2024) but we added questions about the strategy the participants were considering using. Participants were asked whether they were going to use a strategy to be convincing in the interview (yes/no) followed by an open-ended question to describe their strategy or to explain why they are not going to use a strategy.

2.3.6. The interview

After completing the pre-interview questionnaire, participants were brought into contact with the interviewer who carried out the same interview as used by Leal et al. (2024). The interviewer started the interview as follows:

"I understood from the experimenter that you are strongly in favour of/opposed to (opinion)." In the same example statement-absent condition, the interviewer then asked the following:

Q1 Please can you tell me all the reasons you can think of that led you to having this opinion?" (eliciting-opinion request);

After responding to the eliciting-opinion request, participants were asked:

Q2 OK, now try to play the devil's advocate and imagine that you do not have this view at all. That is, you have the opposite view to the one you just expressed. Please tell me anything you can think of that is opposed to the opinion you just expressed? (devil's advocate request).

In the same example statement-present condition, Q1 was introduced as follows: "I would now like to ask you to tell me all the reasons you can think of that led you to having this opinion. But, before doing so, I would like to play you a video clip which serves as an example of how many details I would like you to include in your response. The example I will play is a so called 'same example statement' as it gives you an idea of a detailed response to a question. After watching the example, I will ask you to provide all the reasons you can think of that led you to having this opinion and I would like you to be that detailed in your response, OK?"

After watching the same example statement, the interviewer said: "Bearing in mind the amount of detail you heard in that clip, please provide all the reasons you can think of that led you to having this opinion?" (eliciting-opinion request). This was followed with the devil's advocate request (Q2).

2.3.7. Post-interview questionnaire and debrief

After the interview, participants completed a post-interview questionnaire via Gorilla. It measured motivation, rapport with the interviewer, percentage of lying/truth telling, confidence that they were believed by the interviewer, likelihood of having to write a statement and likelihood of being entered in prize draw. To measure motivation, participants answered the question "To what extent were you motivated to perform well during the interview?" on a 7-point scale ranging from 1 (*not at all motivated*) to 7 (*very motivated*). To measure rapport, participants completed the nine-item Interaction Questionnaire (Vallano & Schreiber Compo, 2011). They rated the interviewer on nine characteristics including smooth, engrossed, involved and active, using 7-point scales ranging from 1 (*not at all*) to 7 (*extremely*) (Cronbach's alpha = 0.78). Rapport is considered important for a productive interview (Brimbal, Dianiska, Swanner, & Meissner, 2019) Participants reported the estimated percentage of truth telling during the interview on an 11-point rating scale ranging from *0%* to *100%* and rated on 7-point Likert scales ranging from 1 (*not at all*) to 7 (*very much*) whether they thought (i) the interviewer believed them, (ii) they had to write a statement and (iii) they would be entered in a prize draw. All these questions were taken from Leal et al. (2024) but we added questions about strategies. Participants were finally asked whether they used a strategy to be convincing in the interview (yes/no) followed by an open-ended question to describe their strategy or to explain why they did not use a strategy.[1]

After completing the post-interview questionnaire, participants were sent the debrief form via Gorilla and were told how they could claim the £15 participant money (if they did not receive course credits). The experimenter informed participants that the interviewer believed them and that they would be entered into the prize draw worth £50, £100 and £150. The three prize winners were randomly selected after data collection was completed.

2.4. Coding

All interviews were video- and audio recorded. The audio recordings were transcribed and the transcripts were used to code the reported arguments, eloquence and predictability. The video recordings were used to code passion and (again) eloquence and predictability.

Two raters, blind to the veracity status of the participants, coded all the transcripts.

(100% coding by both raters). They counted the number of arguments supporting and opposing the alleged opinion, and rated plausibility, immediacy, clarity, and predictability on 7-point Likert scales (1 = *not at all* to 7 = *totally*).

An *argument* was defined as 'A reason provided in favour (pro) or against (anti) the alleged opinion.' The definitions of plausibility, immediacy and clarity were derived from Leal et al. (2023, 2024) and Mann et al. (2023). *Plausibility* was defined as 'Does the answer sound reasonable and genuine and was there enough of an answer to sound convincing'; *Immediacy* refers to a statement being 'personal and not distanced'; *Clarity* was defined as 'How clearly does the reader understand what the participant was saying by the end of the answer?'; *Predictability* was defined as: 'If you ask ten people the same question, how likely is it that you would get the same response?'

Below is the response of someone who said they strongly opposed to the statement that the (British) Conservative Government did what was best for the UK. The twelve arguments the participant gave in support of their alleged view are in italics.

Response: My view has to do with COVID. I think the way the whole thing was handled - it was handled so poorly. So I think firstly, *like we were too late to implement anything, because the cases were getting higher and higher*[1], and the way that *they first dealt with the very first people that got COVID, having to transport them on the coaches. I don't know why the government thought that was a good idea*[2]. And it also took *them quite a while to shut the schools, so many people felt quite unsafe going to school*[3]. It also took a *while for them to tell people to stay at home, so people were still traveling and kind of living in fear that they might catch this virus that's like unknown as well as this*[4]. I think the way that they *handled the lockdown was very, very bad, because they were all meeting up and having bring your own booze parties*[5], whilst *everyone else had to stay inside and they had to watch their family members suffer with COVID*[6], and some of *them unfortunately passed away whilst all of these MPs and our own Prime Minister was out partying and drinking and telling people to stay at home during the conferences every day*[7]. I just think the way that was handled was extremely poor and I think the after like the aftershock from that as well, *so trying to go back to normal after that was also handled very poorly*[8]. The fact that we had just got out of the *lockdown just to be put straight back into one with the tears around Christmas time, it made it tough for everybody, and it was as if they were just putting their own interests at the center, rather than that of the public and the community*[9]. So yeah, personally, I think the way they handled it was extremely poor, and I don't think they *really did anything to help anybody*[10]. They claim that they did, but I don't think they were doing enough. And as well other examples, *like with energy costs, and like the elderly struggling and just they just haven't handled it well*[11]. And like I said, it's as if they're *just putting their own interests first, not that of the UK*[12]. So, yeah, I just disagree that they are doing what's best well, did what's best when they were in power.

The response received high ratings on plausibility (7), clarity (6), and medium ratings on immediacy (4) and predictability (4). The response was considered very plausible because the response sounds reasonable

[1] Participants in the same example statement-present condition were also asked the extent to which they agreed with the person in the same example statement. Answers were given on a 7-point scale ranging from (1) *strongly disagree* to (7) *strongly agree*. Due to a technical error only 40 participants (22 truth tellers and 18 lie tellers) were asked this question. Truth tellers agreed with the person considerably more (M = 6.64, SD = 0.58, 95% CI[5.98,7.29]) than lie tellers (M = 2.44, SD = 2.18, 95% CI[1.72,3.17]), $F(1,38) = 75.52$, $p < .001$, $d = 2.76(1.85,357)$, $BF_{10} = 4.340 \times 10^7$.

and describes the participant's thought process succinctly and in detail. It was considered clear because it presents the arguments well. It was considered medium on immediacy because the participant relates the question to the population in general and not specific to themselves. It was seen as medium on predictability. All these arguments are presented in the news which would make the answer predictable, but the large number of arguments provided makes it less predictable.

Below is the single argument the participant mentioned against their alleged opinion in response to the devil's advocate request:

Response: Well, I think, they did somewhat provide a direction of what to do, going back to the pandemic example, because for me, that's kind of like the main bulk of what I can remember of them being in power. So they did provide *some kind of direction of what to do, because it was a very unprecedented time, so obviously people weren't sure what to do*[1]. So I can appreciate that they did somewhat provide guidance, like with the lockdowns and telling people to not go to school and to work.

The response received a high rating on predictability (6), a medium rating on immediacy (4) and low ratings on plausibility (3), immediacy (4), and clarity (3). The response was considered predictable because how the government handled COVID has been widely discussed in the UK press. It was considered medium on immediacy because, again, the participant relates the response to people in general rather than to themselves. The response was considered somewhat implausible because it did not give much of an answer. It was seen as somewhat unclear because the participant does not elaborate much on the single argument they provide.

Following Leal et al. (2024) we clustered the variables plausibility, immediacy and clarity in the cluster 'eloquence' (Cronbach's alpha = 0.77). We report the results for the three individual variables (plausibility, immediacy and clarity) in Appendix 4. Inter-rater reliability between the two raters was measured using the two-way random effects model measuring consistency. Hallgren (2012) reported that inter-rater reliability is poor for intraclass correlation coefficient (ICC) values less than 0.40, fair for values between 0.40 and 0.59, good for values between 0.60 and 0.74, and excellent for values between 0.75 and 1. The inter-rater reliability varied from fair to excellent: Pro-alleged opinion arguments (Average Measures ICC = 0.85), anti-alleged opinion arguments (Average Measures ICC = 0.89), eloquence (Average Measures ICC = 0.75), and predictability (Average Measures ICC = 0.57). We averaged the ratings of the two raters and used these averaged scores in the analyses. For each participant, we calculated two scores for eloquence and predictability: One for the elicit-opinion request and one for the devil's advocate request. For the pro-opinion arguments, we recorded whether the argument was mentioned in the same example statement.

Two raters coded passion, eloquence (plausibility, immediacy and clarity) and predictability by watching the videos. We used Leal et al.'s (2024) definition of passion: 'How passionate and enthusiastic was the participant when delivering their responses'. The inter-rater reliability for all three variables was good: passion (Average Measures ICC = 0.60); eloquence (Average Measures ICC = 0.74) and predictability (Average Measures ICC = 0.60) We averaged the ratings of the two raters and used these averaged scores in the analyses. For each participant, we calculated two scores for eloquence, predictability and passion: One for the elicit-opinion request and one for the devil's advocate request. Since the passion variable was included in the hypotheses, we present the results of that variable in the main text together with the text analyses of eloquence and predictability (Tables 2 to 4). We present the eloquence and predictability video analyses in Appendix 5. The text and video eloquence measures, $r(171) = 0.70$, $p < .001$ and the text and video predictability measures, $r(171) = 0.63$, $p < .001$ correlated highly with each other.

One rater read the open-question responses regarding the strategies the participants said they were going to use and have used and put them into categories. A second coder was given these categories and also asked to put each of the participant responses in these categories. The reliability between the two coders was good for the pre-interview questionnaire strategies (Kappa = 0.74) and for the post-interview questionnaire strategies (Kappa = 0.71). The disagreement were resolved by a third rater.

3. Results

3.1. Pre- and post-interview questionnaire variables

We carried out frequentist analyses (to test for significance) and Bayesian analyses of variance (to test the likelihood of the data under the alternative hypothesis [H1]). Bayes Factors (BF_{10}) between 1 and 3 indicate weak evidence for the alternative hypothesis (H1), between 3 and 20 indicate positive evidence, between 20 and 150 indicate strong evidence, and above 150 indicate very strong evidence (Jarosz & Wiley, 2014). In other words, only Bayes Factor results of 3 and higher provide sufficient evidence for the alternative hypothesis and we will only interpret significant results with a Bayes Factor larger than 3. We also report Cohen's d effect size to test the magnitude of effects (Cohen, 1988). An effect size of $d = 0.20$ is considered small, $d = 0.50$ is considered medium and $d = 0.80$ is considered large. Effect sizes can be transformed into estimated accuracy scores (correct classifications of truth tellers and lie tellers) as follows: $d = 0.20$ equals 58% estimated accuracy; $d = 0.50$ equals 69% estimated accuracy, $d = 0.80$ equals 79% estimated accuracy and $d = 1.40$ equals 92% estimated accuracy (McLeod, 2023).

A 2 (Veracity) x 2 (Same Example Statement) MANOVA was carried out with the eight pre- and post-interview questionnaire variables listed in Table 1 as dependent variables. The multivariate Veracity main effect was significant, $F(9, 160) = 25.03$, $p < .001$, $\eta_p^2 = 0.56$, but the Same Example Statement main effect, $F(9, 160) = 1.47$, $p = .172$, $\eta_p^2 = 0.07$, and the Veracity x Same Example Statement interaction effect, $F(9, 160) = 0.79$, $p = .610$, $\eta_p^2 = 0.04$, were not. Table 1 presents the univariate Veracity results.

Veracity differences emerged for truth telling, confidence to be believed by the interviewer, likelihood of writing a statement and likelihood to be entered in prize draw. The Bayes Factors analysis did not reveal sufficient evidence for the likelihood of writing a statement effect. Regarding the other variables, truth tellers reported having told the truth more than lie tellers (indicating a successful manipulation). Compared to lie tellers, truth tellers also expressed more confidence that the interviewer would believe them and thought it to be more likely to be entered in the prize draw. The mean scores further showed that the participants were very motivated to do well in the interview. They were somewhat satisfied with the thoroughness of their preparation and very satisfied with the time given to prepare themselves.

3.2. Pre-registered hypothesis testing

A mixed 2 (Veracity) x 2 (Type of Question) x 2 (Same Example Statement) MANOVA was conducted. Veracity and Same Example Statement were the between-subjects factors and Type of Question was the within-subjects factor. Eloquence, predictability and passion were the dependent variables. Multivariate significant effects occurred for Type of Question, $F(3, 165) = 19.96$, $p < .001$, $\eta_p^2 = 0.27$, Same Example Statement, $F(3, 165) = 3.84$, $p = .011$, $\eta_p^2 = 0.07$, and Veracity x Type of Question, $F(3, 165) = 13.81$, $p < .001$, $\eta_p^2 = 0.20$. All other multivariate effects were not significant, all $Fs < 2.42$, all $ps > 0.068$.

The Same Example Statement main effects (Hypothesis 4) and the Veracity x Type of Question interaction effects (Hypotheses 1 and 2) are relevant for the hypotheses-testing purposes, whereas the Type of Question main effects are not. See Appendix 3 for the Type of Question main effects.

3.2.1. Hypotheses 1 and 2

The univariate Veracity x Type of Question interaction effects were

Table 1
Pre- and post-interview questionnaire variables as a function of veracity.

Questionnaire variables	Truth			Lie			*F*	*p*	*d*	BF_{10}
	M	*(SD)*	95% CI	*M*	*(SD)*	95% CI				
Pre-interview questionnaire										
Motivation (1–7)	6.18	(0.90)	5.98,6.38	5.93	(1.02)	5.72,6.13	2.95	.088	0.26 (−0.04,0.56)	0.693
Preparation thoroughness (1–7)	4.78	(1.32)	4.50,5.04	4.42	(1.24)	4.14,4.69	3.33	.070	0.28 (−0.02,0.58)	0.830
Preparation time (1–7)	5.92	(1.20)	5.68,6.14	5.97	(1.04)	5.68,6.15	0.01	.974	0.04 (−0.26,0.34)	0.166
Post-interview questionnaire										
Rapport (1–7)	5.72	(0.86)	5.54,5.88	5.55	(0.79)	5.38,5.73	1.58	.211	0.21 (−0.10,0.50)	0.366
Truth telling (percentage)	93.97	(13.66)	0.88,0.99	40.36	(33.27)	0.35,0.46	189.06	<.001	2.12 (1.71,2.46)	1.476×10^{26}
Confidence to be believed (1–7)	5.32	(1.38)	5.00,5.63	3.99	(1.57)	3.69,4.30	34.65	<.001	0.90 (0.57,1.20)	580,692.878
Likelihood of writing a statement (1–7)	3.68	(1.60)	3.35,4.02	4.26	(1.57)	3.92,4.61	5.72	.018	0.37 (0.06,0.66)	2.311
Likelihood of being entered in prize draw (1–7)	4.33	(1.68)	3.98,4.68	3.19	(1.56)	2.99,3.70	15.13	<.001	0.70 (0.38,1.00)	157.869

significant for eloquence, $F(1, 167) = 22.34$, $p < .001$, $\eta_p^2 = 0.12$, and passion $F(1, 167) = 38.24$, $p < .001$, $\eta_p^2 = 0.19$ but not for predictability, $F(1, 167) = 2.98$, $p = .086$, $\eta_p^2 = 0.02$. To further examine the interaction effects, we followed Leal et al. (2024) and conducted ANOVAs with Veracity as the only factor and eloquence, predictability and passion as dependent variables. Separate analyses were carried out for the eliciting-opinion requests results (Hypothesis 1) and residue scores (Hypothesis 2). The findings are presented in Table 2. We included predictability in these analyses to make the analyses complete.

When responding to the eliciting-opinion request, truth tellers sounded more eloquent, less predictable and more passionate than lie tellers. The Bayes Factor results showed sufficient support for eloquence and passion, supporting Hypothesis 1 for eloquence and passion. The residue scores were larger for truth tellers than for lie tellers for eloquence and passion with sufficient Bayes Factors evidence for both variables. This supports Hypothesis 2 for eloquence and passion. Following Leal et al. (2024), we carried out one-sample *t*-tests comparing the residue scores with '0', the score that would be obtained if the participants' responses to the eliciting-opinion and devil's advocate requests would be identical. For truth tellers, the residue scores differed significantly from '0' for eloquence, $t(86) = 7.42$, $p < .001$, predictability, $t(86) = -5.72$, $p < .001$, and passion, $t(86) = 7.98$, $p < .001$. For lie tellers, the residue scores did differ significantly from '0' for predictability, $t(83) = 3.55$, $p < .001$ but not for eloquence, $t(83) = 1.14$, $p = .269$, and passion, $t(83) = 0.08$, $p = .470$.

3.2.2. Hypothesis 3

The multivariate interaction effects that included both Veracity and Same Example Statement were not significant: Veracity x Same Example Statement, $F(3, 165) = 1.95$, $p = .124$, $\eta_p^2 = 0.03$ and Veracity x Type of Question x Same Example Statement, $F(3, 165) = 2.41$, $p = .069$, $\eta_p^2 = 0.04$. The absence of a significant Veracity x Same Example Statement.

interaction effect suggests lack of support for Hypothesis 3. However, Hypothesis 3 referred to attenuated interaction effects which are particularly difficult to detect (Blake & Gangestad, 2020). Following Leal et al. (2024), to avoid missing possible interaction effects, we carried out ANOVAs to analyse Veracity effects for each of the same example statement conditions separately, see Table 3.

In responding to the eliciting-opinion request, larger effect sizes emerged in the same example statement-present for eloquence and predictability (*d*-scores were 1.01 and 0.40 respectively) than in the same example statement-absent condition (*d*-scores were 0.17 and 0.23 respectively), suggesting that the same example statement did enhance the veracity differences. The Veracity effect was significant with sufficient Bayes Factor evidence for eloquence in the same example statement-present condition only. This supports Hypothesis 3 for eloquence. An identical pattern of results emerged for the differences in responding to the eliciting-opinion and devil's advocate requests: The effect sizes for eloquence and predictability were the largest in the same example statement-present condition, and the Veracity effect was significant with sufficient Bayes Factor evidence for eloquence in the same example statement-present condition only. This supports Hypothesis 3 for eloquence.

For truth tellers in the same example statement-absent condition, the residue scores differed significantly from '0' for eloquence, $t(41) = 3.84$, $p < .001$, predictability $t(41) = 2.37$, $p = .011$ and passion, $t(41) = 5.46$, $p < .001$. For truth tellers in the same example statement-present condition, the residue scores also differed significantly from '0' for eloquence, $t(44) = 6.79$, $p < .001$, predictability, $t(44) = 5.69$, $p < .001$, and passion, $t(44) = 5.86$, $p < .001$.

For lie tellers in the same example statement-absent condition, the residue scores differed significantly from '0' for predictability, $t(42) = 2.08$ $p = .043$ but not for eloquence, $t(42) = 0.98$, $p = .331$ and passion, $t(42) = 0.52$, $p = .604$. For lie tellers in the same example statement-present condition the same pattern of results emerged. The residue scores differed significantly from '0' for predictability, $t(40) = 2.80$, $p = .008$ but not for eloquence, $t(40) = 0.60$, $p = .550$ and passion, $t(40) = 0.56$, $p = .578$.

Table 2
Interview responses as a function of veracity.

Interview responses	Truth			Lie			NHST			BF_{10}
	M	*(SD)*	95% CI	*M*	*(SD)*	95% CI	*F*	*p*	*d*	
Eliciting-opinion request										
Eloquence (text analyses)	4.84	(1.14)	4.62,5.06	4.22	(0.95)	3.99,4.45	14.82	<.001	0.59 (0.27,0.89)	131.26
Predictability (text analyses)	3.78	(1.17)	3.53,4.02	4.13	(1.14)	3.88,4.38	4.02	.047	0.30 (0.27,0.89)	0.95
Passion (video analyses)	5.01	(1.39)	4.06,4.36	4.08	(0.94)	3.92,4.23	26.34	<.001	0.78 (0.46,1.08)	17,621.705
Eliciting-opinion minus devil's advocate request										
Eloquence (text analyses)	1.39	(1.75)	1.04,1.75	0.19	(1.58)	−0.17,0.55	22.10	<.001	0.72 (0.40,1.02)	2981.35
Predictability (text analyses)	−0.86	(1.41)	−1.15,-0.57	−0.49	(1.31)	−0.78,-0.20	3.13	.078	0.27(−.03,0.57)	0.70
Passion (video analyses)	1.60	(1.87)	1.25,1.95	0.01	(1.42)	−0.35,0.37	38.82	<.001	0.96 (0.62,1.26)	2.701×10^6

Note. NHST = Null-hypothesis significance testing.

Table 3
Interview responses as a function of veracity for each same example statement condition.

Interview responses	Truth			Lie			NHST			BF_{10}
	M	*(SD)*	95% CI	*M*	*(SD)*	95% CI	*F*	*p*	*d*	
Eliciting-opinion request										
Same example statement-absent										
Eloquence (text analyses)	4.35	(0.96)	4.06,4.65	4.19	(0.96)	3.90,4.48	0.59	.446	0.17 (−0.26,0.59)	0.293
Predictability (text analyses)	4.07	(1.24)	3.68,4.47	4.37	(1.34)	3.98,4.76	1.16	.285	0.23 (−0.20,0.66)	0.375
Passion (video analyses)	4.70	(1.26)	4.37,5.03	3.87	(0.84)	3.55,4.20	12.79	<.001	0.78 (0.32,1.21)	47.67
Same example statement-present										
Eloquence (text analyses)	5.29	(1.12)	4.98,5.60	4.25	(0.94)	3.92,4.57	21.58	<.001	1.01 (0.54,1.44)	1415.390
Predictability (text analyses)	3.50	(1.04)	3.22,3.78	3.88	(0.84)	3.58,4.17	3.37	.070	0.40 (−0.03,0.83)	0.973
Passion (video analyses)	5.30	(1.45)	4.93,5.67	4.29	(1.00)	3.90,4.68	13.78	<.001	0.80 (0.35,1.23)	71.030
Eliciting-opinion minus devil's advocate request										
Same example statement-absent										
Eloquence (text analyses)	0.85	(1.43)	0.40,1.30	0.22	(1.50)	−0.22,0.67	3.84	.053	0.43 (−0.01,0.85)	1.194
Predictability (text analyses)	−0.43	(1.17)	−0.81,-0.05	−0.42	(1.32)	−0.80,-0.04	0.01	.971	0.01 (−0.42,0.43)	0.226
Passion (video analyses)	1.39	(1.65)	0.94,1.85	−0.10	(1.32)	−0.56,0.35	21.37	<.001	1.00 (0.53,1.43)	1294.132
Same example statement-present										
Eloquence (text analyses)	1.90	(1.88)	1.37,2.43	0.16	(1.69)	−0.40,0.72	20.39	<.001	0.97 (0.51,1.40)	905.505
Predictability (text analyses)	−1.27	(1.49)	−1.69,-0.85	−0.57	(1.31)	−1.01,.0.14	5.19	.025	0.50 (0.06,0.92)	2.108
Passion (video analyses)	1.79	(2.05)	1.25,2.33	0.13	(1.53)	−0.43,0.70	17.73	<.001	0.91 (0.45,1.34)	331.190

Note. NHST = Null-hypothesis significance testing.

3.2.3. *Hypothesis 4*

The Same Example Statement main effects (Hypothesis 4) are presented in Table 4.

Participants in the same example statement-present condition were more eloquent, less predictable and more passionate than participants in the same example statement-absent condition. The Bayes Factor results showed sufficient support for eloquence and passion but not for predictability. Hypothesis 4 was therefore supported for eloquence and passion only.

To examine the pro- and anti-alleged opinion arguments component in Hypothesis 4, a 2 (Veracity) x 2 (Same Example Statement) MANOVA was carried out with the pro- and anti-alleged opinion arguments and the arguments mentioned in the same example statement as dependent variables. At a multivariate level the Same Example Statement main effect was significant, $F(3, 165) = 16.32$, $p < .001$ $\eta_p^2 = 0.23$, whereas the Veracity main effect, $F(3, 165) = 1.36$, $p = .256$, $\eta_p^2 = 0.02$, and the Veracity x Same Example Statement interaction effect, $F(3, 165) = 1.46$, $p = .227$, $\eta_p^2 = 0.03$, were not significant. At a univariate level, none of the Veracity main effects and Veracity x Same Example Statement interaction effects were significant either, all $Fs < 2.50$, all $ps > 0.115$.

The univariate results for the Same Example Statement main effect are presented in Table 4. Compared to the same example statement-absent condition, participants in the same example statement-present condition, reported more pro-alleged opinion arguments and mentioned more arguments that also occurred in the same example statement than in the same example statement-absent condition. The Bayes Factor results showed sufficient support for both effects. Hypothesis 4 was therefore supported for pro-alleged opinion arguments.

The finding that participants in the same example statement-present condition also mentioned more pro-alleged arguments they heard in the same example statement than participants in the same example statement-absent condition may have affected the pro-alleged arguments results. Since participants in the same example statement-present condition were exposed to these arguments whereas participants in the same example statement-absent condition were not, this may have given participants in the same example statement-present condition an advantage. We therefore computed a new pro-alleged arguments variable. It included only pro-alleged arguments not mentioned in the same example statement. A 2 (Veracity) x 2 (Same Example Statement) ANOVA with this variable as dependent variable showed a main effect for Same Example Statement (see Table 4), whereas the Veracity main effect, $F(1, 167) = 2.53$, $p = .113$, $\eta_p^2 = 0.02$ and the Veracity x Same Example Statement interaction effect, $F(1, 167) = 1.43$, $p = .233$, $\eta_p^2 = 0.01$ were not significant. As Table 4 shows, participants in the same

Table 4
Interview responses as a function of same example statement.

Interview responses	Same example statement-present			Same example statement-absent			NHST			BF_{10}
	M	*(SD)*	95% CI	*M*	*(SD)*	95% CI	*F*	*p*	*d*	
Main effect results										
Eloquence (text analyses)	4.26	(0.57)	4.13,4.39	4.01	(0.68)	3.87,4.14	6.96	.009	0.40 (0.09,0.70)	4.003
Predictability (text analyses)	4.15	(0.85)	3.95,4.35	4.44	(1.03)	4.23,4.63	3.96	.048	0.31 (0.00,0.60)	1.025
Passion (video analyses)	4.32	(0.70)	4.17,4.47	3.96	(0.72)	3.81,4.12	10.77	.001	0.51 (0.19,0.80)	21.997
Pro-alleged opinion arguments (text analyses)	9.94	(4.73)	9.01,10.88	6.54	(4.00)	5.60,7.47	25.80	<.001	0.78 (0.45,1.07)	14,063.660
Anti-alleged opinion arguments (text analyses)	4.77	(3.73)	4.08,5.45	4.13	(2.61)	3.44,4.82	1.68	.197	0.20 (−0.10,0.50)	0.360
Arguments mentioned in the same example statement (text analyses)	3.15	(1.80)	2.81,3.48	1.61	(1.27)	1.28,1.95	41.38	<.001	0.99 (0.66,1.29)	7.341×10^6
Pro-alleged opinion arguments minus arguments mentioned in the same example statement (text analyses)	6.80	(4.42)	5.93,7.67	4.92	(3.72)	4.05,5.80	8.69	.004	0.46 (0.15,0.76)	9.978

Note. NHST = Null-hypothesis significance testing.

example statement-present condition also reported more pro-alleged arguments not mentioned in the same example statement than participants in the same example statement-absent condition.

3.3. Strategies

There was no association between indicating in the pre-questionnaire to going to use a strategy and Veracity, $\chi^2(1, N = 171) = 2.06$, $p = .151$. The strategies reported by the 30 truth tellers (44.1% of the sample of truth tellers) and 38 lie tellers (55.9% of lie tellers) are presented in Table 5.

The strategies frequently mentioned by lie tellers (e.g. by at least ten lie tellers) in the pre-interview questionnaire were using background knowledge and facts, using the opposite view from their own perspective and using the prepared arguments and examples. Truth tellers frequently mentioned paying attention to their own demeanour and being truthful. The main reason for lie tellers not to plan a strategy was that they preferred to give a spontaneous answer. For truth tellers the main reasons not to plan a strategy was that they were going to tell the truth or preferred to give a spontaneous answer.

There was no association between indicating in the post-interview questionnaire to have used a strategy and Veracity $\chi^2(1, N = 171) = 2.35$, $p = .125$. The strategies used by the 57 truth tellers (66.5% of the sample of truth tellers) and 64 lie tellers (76.2% of lie tellers) are presented in Table 5. The most frequently reported strategies by lie tellers were using their background knowledge and facts, using their prepared arguments and examples, paying attention to their demeanour and using the opposite view to their own view. The most frequently reported strategies by truth tellers were being truthful, paying attention to their demeanour, using their background knowledge and facts and using their prepared arguments and examples.

Table 5
Strategies reported by the participants.

Strategy		
Pre-interview questionnaire: What strategies are you going to use?	**Lie tellers ($n = 38$)**	**Truth tellers ($n = 30$)**
Use my background knowledge and facts	14	9
Use the opposite from my own perspective	13	0
Demeanour	10	13
Prepared arguments and provide examples	9	4
Consistency	1	1
Be truthful	0	11
Keep it simple and to the point	3	4
Pre-interview questionnaire: Why didn't you prepare a strategy?	**Lie tellers ($n = 44$)**	**Truth tellers ($n = 57$)**
Prefer a spontaneous answer	22	16
I will be convincing enough without a strategy	7	5
Don't know what strategy to use	7	2
If I use a strategy the interviewer will figure it out	5	6
There is no need for a strategy	4	7
Don't know what will be asked	3	3
Will tell the truth	0	37
I am passionate about the topic	0	2
Post-interview questionnaire: Which strategy did you use?	**Lie tellers (n = 64)**	**Truth tellers (n = 57)**
Use my background knowledge and facts	22	17
Used my prepared arguments and gave examples	17	11
Paid attention to my demeanour	17	19
Used the view opposite to my own view	11	4
Pretended to be confused in the DA response	6	0
Used personal experiences	5	0
Gave more details in EO than in DA response	4	3
Copied responses from the same example statement	3	0
Tried to be detailed	3	5
Consistency in nonverbal behavior and conversational style	2	2
Tried not to be detailed	2	1
Be truthful	0	42
Could not do well in DA response	0	7

3.4. Eloquence and predictability results based on video analyses

Appendix 5 shows that the analyses of the videos followed a similar pattern to the text analyses. Table 2B shows that Veracity differences for eloquence emerged when responding to the eliciting-opinion request (truth tellers were more eloquent than lie tellers) and for the differences in responses to the eliciting-opinion minus devil's advocate requests (the difference was larger for truth tellers than for lie tellers). When examining the results for the two same example statement conditions separately (Table 3B), a significant Veracity effect in responding to the eliciting-opinion request for eloquence emerged in the same example statement-present condition only, but the Bayes Factor results failed to show sufficient evidence. A significant Veracity effect with sufficient Bayes Factor evidence emerged for eloquence regarding the residue score (difference in responding to the eliciting-opinion and devil's advocate requests) in both same example statement conditions and the effect was most substantial in the same example statement-present condition.

4. Discussion

4.1. Hypotheses-testing results

Table 6 provides a summary of the significant Veracity effects ($p < .05$) that received sufficient Bayes Factors evidence ($BF_{10} > 3$). Truth tellers were more eloquent and more passionate than lie tellers when responding to the eliciting-opinion request, supporting Hypothesis 1 for eloquence and passion. The difference in eloquence and passion between the eliciting-opinion and devil's advocate requests was more pronounced in truth tellers than in lie tellers, supporting Hypothesis 2 for these two variables. The same example statement enhanced the eloquence Veracity effects, supporting Hypothesis 4 for eloquence.

The hypotheses were not supported for predictability. This is particularly unfortunate because this is a cue to deceit (lie tellers report the cue more frequently than truth tellers) rather than a cue to truthfulness (truth tellers report the cue more frequently than lie tellers). Most verbal veracity cues are cues to truthfulness. For example, all 19 Criteria-Based Content Analysis cues (Amado, Arce, Fariña, & Vilarino, 2016) are cues to truthfulness, and so are most Reality Monitoring cues (Gancedo, Fariña, Seijo, Vilariño, & Arce, 2021). Verbal cues to deceit are scarce and except for the statement – evidence inconsistency cue (Oleszkiewicz & Watson, 2021), verbal cues to deceit are less diagnostic than cues to truthfulness (Gancedo et al., 2021; Vrij, Palena, Leal, & Caso, 2021). In real life it would be easier for practitioners to decide that someone is lying if the absence of a cue to truthfulness is associated with the presence of a cue to deceit (Vrij, Fisher, & Leal, 2023). Predictability may be a difficult cue to use when judging opinions. The interrater reliability between the two raters was lower for predictability than for the other cues, suggesting that it is a more subjective cue than the other cues. And the statements by non-native English speakers were considered less predictable than the statements from native English speakers. Predictability is perhaps related to someone's knowledge: If interviewees mention reasons that observers do not think of themselves, the statement comes across as less predictable. This hypothesis is in alignment with the false consensus effect, people's tendency to think that their own way of thinking is more common than it is (Gilovich, 1990; Mullen et al., 1985; Ross, Greene, & House, 1977).

This is the first devil's advocate experiment where eloquence and predictability were also coded by watching videotapes. We obtained high correlations between text and video analyses for both eloquence and predictability, but Table 6 shows that the veracity effects were more

Table 6
Summary of the significant veracity results.

Interview responses	Present experiment			Leal et al. (2024)		
	All participants	Same example statement-absent	Same example statement-present	All participants	Model statement-absent	Model statement-present
	d	*d*	*d*	*d*	*d*	*d*
Eliciting-opinion request						
Eloquence text analyses	0.59		1.01	0.55		0.80
Predictability text analyses						
Passion video analyses	0.78	0.78	0.80	0.55		0.80
Eloquence video analyses	0.41					
Predictability video analyses						
Eliciting-Opinion Minus Devil's Advocate Request						
Eloquence text analyses	0.72		0.97	0.76	0.55	0.98
Predictability text analyses				−0.44		−0.69
Passion video analyses	0.96	1.00	0.91	0.51		0.56
Eloquence video analyses	0.63	0.54	0.69			
Predictability video analyses						

Note. A positive *d*-score indicates a higher score for truth tellers than for lie tellers, whereas a negative *d*-score indicates a higher score for lie tellers than for truth tellers.

pronounced in text analyses than in video analyses. Many factors could explain this result (e.g. the video and text raters were different people) but if we would have formulated a hypothesis we would have predicted the obtained result. To code eloquence and predictability, raters should focus on what the person says. Raters can probably do that more thoroughly when reading transcripts than when watching videos. Veracity differences are therefore more likely to emerge when reading transcripts than when watching videos. Future research could test this hypothesis.

Listening to a same example statement encouraged participants to provide more reasons themselves, supporting Hypothesis 3. The absence of a Veracity x Same Example Statement interaction effect shows that a same example statement led to a similar increase in the number of arguments provided by both truth tellers and lie tellers. In other words, a same example statement did not result in a quantitative Veracity effect (e.g. number of arguments), but only in a qualitative Veracity effect (e.g. eloquence). The same pattern occurs when using a model statement when interviewees report their alleged activities. A model statement increases the number of details that truth tellers and lie tellers report to the same extent (absence of a quantitative Veracity effect), but the quality of the statements improves more in truth tellers than in lie tellers (Vrij et al., 2018).

4.2. Strategies

Only two differences emerged in the strategies truth tellers and lie tellers *planned to use* in the forthcoming interviews and both were obvious (lie tellers reported that they were going to present a view different to their own view and truth tellers reported that they were going to be truthful). The main strategy planned by both truth tellers and lie tellers is to base their answers on knowledge and facts. Both veracity groups also planned to control their demeanour.

Also, considerable overlap emerged in the strategies truth tellers and lie tellers *reported to have used* in the interviews. Two strategies appeared particularly popular: Both truth tellers and lie tellers reported to have used to a similar extent background knowledge, facts and their prepared arguments and examples. They also reported to have paid attention to their demeanour to a similar extent. These two strategies are in alignment with what participants indicated that they planned to do.

In other words, the typical finding that truth tellers plan to be forthcoming and lie tellers plan to keep it simple (Hartwig et al., 2007), did not emerge in the present experiment. The strategies they did use instead were more like each other than the strategies truth tellers and lie tellers use in interviews about alleged activities. This is the first experiment that examined truth tellers' and lie tellers' strategies in opinion interviews and more research is required. However, if the present finding is replicated in future research and truth tellers and lie tellers report similar strategies, it would make designing interview protocols that exploit veracity strategies in opinion interviews more challenging than designing interview protocols that exploit veracity strategies in activities interviews.

Veracity effects in eloquence and passion did emerge despite truth tellers and lie tellers reporting to have used similar strategies. The analyses showed that truth tellers were more eloquent and passionate in responding to the eliciting-opinion than to the devil's advocate request, whereas lie tellers showed similar level of eloquence and passion in responding to these two requests. Lie tellers therefore showed more consistency than truth tellers when comparing eliciting-opinion and devil's advocate responses. Lie tellers believe that inconsistency raises suspicion (Deeb et al., 2018) and they perhaps strived to be equally eloquent when responding to the eliciting-opinion and devil's advocate requests, something that gives their lies away in a devil's advocate interview. Consistency was mentioned by only one lie teller when they discussed the strategies they were planning to use during the interview. However, at that time they did not know that they would be interviewed with the devil's advocate interview protocol. They knew this when asked in the post-interview questionnaire to report which strategies they had used during the interview, but only two out of 84 lie tellers reported this strategy. It could be that they used consistency but did not report it (Hartwig & Bond, 2011; Nisbett & Wilson, 1977). Alternatively, strategies they frequently reported to have used, particularly using background knowledge, facts, prepared arguments and examples and using a view opposite to their own view, could result in consistency. That is, these strategies most likely referred to responses in support of their alleged opinion. This would increase the eloquence in their eliciting-opinion responses which could make those responses more equal in eloquence compared to their devil's advocate responses.

4.3. Passion: an atypical nonverbal veracity cue

Passion (a nonverbal cue) emerged as a strong veracity indicator. This seems to contradict the general trend in deception research that nonverbal cues are generally weak and unreliable (Denault, Patterson,

Zloteanu, & Talwar, 2025; Patterson, Fridlund, & Crivelli, 2023; Vrij & Fisher, 2019). However, passion differs from the more typical nonverbal cues that people associate with deception (such as gaze aversion and fidgeting) in two important ways. First, passion is a holistic concept whereas nonverbal behaviours such as gaze aversion and fidgeting are single cues. Holistic concepts made of multiple cues are typically more diagnostic veracity indicators than individual cues (DePaulo & Morris, 2004; Hartwig & Bond, 2014). Second, nonverbal behaviours are typically thought to reveal deception because they relate to nervousness. The dominant idea in nonverbal lie detection (Vrij & Fisher, 2020) is that lie tellers are nervous and show nervous behaviours (such as gaze aversion and fidgeting). Passion is not a cue to nervousness; it refers to enthusiasm instead. It is therefore not a typical nonverbal cue, which also becomes evident because it is a cue to truthfulness, whereas typical nonverbal cues are cues to deceit.

4.4. Practical implications

When independent evidence is absent, verbal lie detection is based on analysing the quality of speech content. This is a challenging but not impossible task (Mac Giolla & Luke, 2021; Vrij et al., 2022). Coding tools, such as Criteria-Based Content Analysis and Reality Monitoring, examine specific types of detail. These tools distinguish truth tellers from lie tellers well above chance level (Amado et al., 2016; Amado, Arce, & Fariña, 2015; Gancedo et al., 2021; Sporer & Masip, 2025) In addition, interview protocols such as Cognitive Credibility Assessment and Reality Interviewing further enhance the veracity effects of these details (Mac Giolla & Luke, 2021; Vrij et al., 2022). These coding tools and interview protocols address lying about alleged activities and are used in real life (Vrij et al., 2022). The Devil's Advocate Approach is not as widely examined as these coding tools and interview protocols and all the devil's advocate research to date comes from one research lab. Some caution is therefore warranted. However, the published research shows that the devil's advocate interview protocol is successful in discriminating between truth tellers and lie tellers whereas an alternative method to detect opinion lies is unavailable. We therefore leave it to practitioners to decide whether to use the Devil's Advocate Approach in real life.

4.5. Future research

The eloquence and passion results in Leal et al. (2024) where a Model Statement was used and the present experiment are quite similar, see Table 6. This suggests that it makes little difference whether someone uses a model statement (example unrelated to the topic of investigation) or a same example statement (example is the same as the topic of investigation). To properly examine this, a model statement and a same example statement should be included in the same experiment. Such experiment is worth pursuing.

The underlying rationale of the Devil's Advocate Approach is that arguments in favour of someone's opinion are more readily available than arguments against someone's opinion. This has two implications. First, the tool is perhaps only effective if someone has a strong opinion, because if the opinion is not strong, people will have a more balanced view about the arguments in favour or against that opinion. Second, individuals engaged in critical thinking (those who explore different sides of an argument intensively) should be less suitable for deception detection via the devil's advocate approach. Future research could test these hypotheses.

When lying about alleged activities, lie tellers often stay close to the truth and report embedded lies. That is, they report a largely truthful account in which they include deceptive elements (Leins, Fisher, & Ross, 2013; Markowitz, 2024). Telling an embedded lie has two advantages for lie tellers: They do not have fabricate many details that they then have to remember in case they are asked about it in later interviews; and the fewer incorrect details they provide the less opportunity they give to interviewers to verify that they are lying. Telling an embedded lie is not possible when someone lies about their opinion and someone is forced to present a totally false opinion instead. This may make lying about opinions more difficult than lying about activities which could result in larger veracity effects in opinion lies than in activity lies. This merits further investigation.

Appendix 1. Opinions questionnaire

"We are interested in your views about various contentious political issues. It is important that you give honest answers. If you don't feel you can answer a question freely, don't answer that question. We prefer you to leave an item blank than to give a dishonest answer"

CRediT authorship contribution statement

Sharon Leal: Investigation, Conceptualization, Writing – review & editing. **Aldert Vrij:** Funding acquisition, Formal analysis. **Haneen Deeb:** Supervision, Investigation, Data curation, Writing – review & editing. **Dora Giorgianni:** Validation, Investigation, Data curation, Writing – review & editing. **Lara Geermann:** Validation, Investigation, Data curation, Writing – review & editing. **Ronald P. Fisher:** Conceptualization, Writing – review & editing. **aldert vrij:** Writing – original draft.

Informed consent

Informed consent was obtained from all individual participants included in the study.

Declaration of competing interest

Sharon Leal has declared no conflicts of interest.
Aldert Vrij has declared no conflicts of interest.
Haneen Deeb has declared no conflicts of interest.
Dora Giorgianni has declared no conflicts of interest.
Lara Geermann has declared no conflicts of interest.
Ronald P Fisher has declared no conflicts of interest.

Acknowledgement

This work was supported by the High-Value Detainee Interrogation Group [Grant number 15F06723P0002337]. The funding body had no other involvement in the research than providing financial support. Any opinions, findings, conclusions, or recommendations expressed in this article are those of the authors and do not necessarily reflect the views of the U.S. Government.

Ethical standards statement

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional [University of Portsmouth Faculty of Science and Health Ethics Committee SHFEC C-2023-087A] and the funding (HIG) research committee (2024-11_713-23) and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

1) Capital punishment (i.e. death penalty) should be a legal option in judicial systems for very serious crimes
2) CCTV in streets and public areas is a good thing
3) Our previous Conservative government did what was best for us
4) I feel that minority groups are taking over our country
5) I support Covid vaccinations
6) I support Israel in the Israel/ Palestine conflict

Appendix 2. The same example statement

Pro vaccination:
"I understand from the experimenter that you are strongly agree with the statement that 'I support Covid vaccinations'. Could you please tell me in as much detail as possible everything that has led you to having that opinion, please include all details you remember no matter how insignificant they may seem".

Sure, no problem. Okay, so I strongly agree that covid vaccines are a good thing. I mean, I guess like in terms of context, I'm not a medical professional, and I don't have a science background, but I *grew up in a household where both my parents worked for the NHS – they're doctors. So, I think I probably did grow up with just quite a lot of faith in the medical system*[1] in general, *and of course I trusted doctors*[2] (laugh)

I also really trust the scientists[3] I'm *convinced that we only got out of the covid crisis was due to people having the vaccination*[4]. *All my family have had the vaccination*[5], in fact *we've always had any recommended vaccinations, they've stopped so many horrible diseases*[6]. *With the covid vaccine the worst you feel is that you have a mild flu afterwards*[7], *far better than having actual covid*[8]. *To me it was just so obvious to have it*[9], *we would all be still locked up indoors now if we didn't*[10]. *You also have to think that you are helping others, that due to certain conditions are unable to protect themselves*[11]. *To be honest, it makes me angry that some people refused to have it, it is so selfish when people were dying from catching covid*[12]. I *just think it's stupid and wrong not to have it*[13]. (226 words)

And then of course there's….

Anti vaccination
"I understand from the experimenter that you are strongly opposed to the statement that 'I support Covid vaccinations'. Could you please tell me in as much detail as possible everything that has led you to having that opinion, please include all details you remember no matter how insignificant they may seem"

Sure, no problem. Okay so I strongly disagree that covid vaccines are a good thing. I mean, I guess like in terms of context, I'm not a nutritionist but I did grow up in a healthy- vegan actually- household. *Both my parents were always very careful about not putting any unnecessary toxins in your body*[1]. *They also taught me to question everything! I mean, I just don't trust the scientists with regards to the covid vaccine*[2]; *science has been wrong before*[3] and *there is no way that this vaccine was properly tested*[4]. *They did it all far too quick*[5]. *None of my family have had the vaccination, and we have all been fine*[6]. *If you live healthily and put good things into your body, then your own immune system protects you, you don't need all those toxic chemicals*[7].

I also know people who did have the vaccine and then became quite ill afterwards[8]. *There have even been cases where people have died from having it*[9] although, *I think the main media were told not to report a lot about those cases*[10]. In *my mind, having the vaccine is far more dangerous than catching covid*[11]. *To be honest, it makes me angry that the government tried to push people to have it, when they couldn't possibly know the side effects*[12]. *I think it's just stupid and wrong*[13]. (226 words)

And then of course there's….

Appendix 3. The type of question main effect results

Interview Responses as a Function of Type of Question

Interview responses	Eliciting-opinion request			Devil's advocate request			NHST			BF_{10}
	M	*(SD)*	95% CI	*M*	*(SD)*	95% CI	*F*	*p*	*d*	
Eloquence (text analyses)	4.55	(1.09)	4.37,4.67	3.73	(1.09)	3.58,3.90	39.13	<.001	0.75 (0.52,0.96)	5.097×10^8
Plausibility (text analyses)	4.60	(1.29)	4.41,4.76	3.85	(1.37)	3.66,4.06	20.98	<.001	0.56 (0.34,0.77)	53,103.076

(*continued on next page*)

(*continued*)

Interview responses	Eliciting-opinion request			Devil's advocate request			NHST			BF_{10}
	M	*(SD)*	95% CI	*M*	*(SD)*	95% CI	*F*	*p*	*d*	
Immediacy (text analyses)	4.47	(1.34)	4.27,4.65	3.25	(1.13)	3.09,3.43	80.98	<.001	0.98 (0.75,1.19)	1.142×10^{16}
Clarity (text analyses)	4.54	(1.15)	4.36,4.69	4.09	(1.15)	3.93,4.26	12.14	<.001	0.39 (0.17,0.60)	108.273
Predictability (text analyses)	3.95	(1.17)	3.78,4.13	4.63	(1.18)	4.45,4.81	43.44	<.001	0.58 (0.35,0.79)	1.449×10^{7}
Passion (video analyses)	4.77	(0.93)	4.62,4.90	3.89	(0.79)	3.77,4.01	109.34	<.001	1.02 (0.78,1.23)	4.693×10^{18}

All effects were significant with sufficient Bayes Factor support. When responding to the eliciting-opinion request, participants were more eloquent (i.e. more plausible, immediate and clearer), less predictable and more passionate than when responding to the devil's advocate request.

Appendix 4. The results for the three separate eloquence variables (plausibility, immediacy and clarity)

Table 2A
Interview responses as a function of veracity.

Interview responses	Truth			Lie			NHST			BF_{10}
	M	*(SD)*	95% CI	*M*	*(SD)*	95% CI	*F*	*p*	*d*	
Eliciting-opinion request										
Plausibility (text analyses)	5.07	(1.24)	4.82,5.32	4.11	(1.16)	3.85,4.37	27.07	<.001	0.80 (0.48,1.10)	23,830.914
Immediacy (text analyses)	4.65	(1.44)	4.37,4.93	4.28	(1.20)	3.99,4.57	3.32	.070	0.28 (−0.03,0.58)	0.763
Clarity (text analyses)	4.80	(1.18)	4.56,5.04	4.27	(1.05)	4.03,4.51	9.65	.002	0.47 (0.16,0.77)	13.409
Eliciting-opinion minus devil's advocate request										
Plausibility (text analyses)	1.57	(2.14)	1.13,2.01	−0.11	(2.03)	−0.56,0.34	27.98	<0.001	0.81 (0.48,1.10)	34,600.334
Immediacy (text analyses)	1.16	(2.08)	0.75,1.56	0.05	(1.74)	−0.36,0.47	14.05	<0.001	0.58 (0.26,0.88)	93.741
Clarity (text analyses)	1.03	(1.63)	0.69,1.38	−0.16	(1.74)	−0.51,0.20	22.51	<.001	0.71 (0.39,1.00)	3541.286

Note. NHST = Null-hypothesis significance testing.

When responding to the eliciting-opinion request truth tellers sounded more plausible and clearer than lie tellers with sufficient Bayes Factors support for both effects. The difference in responses to the eliciting-opinion and devil's advocate requests were more pronounced for truth tellers than for lie tellers regarding all three variables, and the Bayes Factor analyses showed sufficient support for all three variables.

Table 3A
Interview responses as a function of veracity for each same example statement condition.

Interview responses	Truth			Lie			NHST			BF_{10}
	M	*(SD)*	95% CI	*M*	*(SD)*	95% CI	*F*	*p*	*d*	
Eliciting-opinion request										
Same example statement-absent										
Plausibility (text analyses)	4.58	(1.04)	4.24,4.93	4.08	(1.18)	3.74,4.42	4.30	.041	0.45 (0.01,0.87)	1.447
Immediacy (text analyses)	4.11	(1.30)	3.71,4.50	4.06	(1.26)	3.67,4.45	0.03	.860	0.04 (−0.39,0.46)	0.229
Clarity (text analyses)	4.37	(1.10)	4.05,4.69	4.44	(0.97)	4.13,4.76	0.10	.758	0.07 (−0.36,0.49)	0.237
Same example statement-present										
Plausibility (text analyses)	5.52	(1.25)	5.17,5.88	4.15	(1.14)	3.77,4.52	28.17	<.001	1.15 (0.67,1.59)	15,412.478
Immediacy (text analyses)	5.16	(1.39)	4.78,5.53	4.51	(1.10)	4.12,4.90	5.58	.020	0.52 (0.08,0.94)	2.489
Clarity (text analyses)	5.20	(1.12)	4.87,5.53	4.09	(1.10)	3.74,4.43	21.60	<.001	1.00 (0.54,1.43)	1426.267
Eliciting-opinion minus devil's advocate request										
Same example statement-absent										
Plausibility (text analyses)	0.95	(1.76)	0.39,1.52	−0.01	(1.90)	−0.57,0.54	5.88	.017	0.52 (0.08,0.95)	2.828
Immediacy (text analyses)	0.48	(1.81)	−0.05,1.00	−0.03	(1.62)	−0.56,0.49	1.88	.174	0.30 (−0.13,0.72)	0.514
Clarity (text analyses)	0.58	(1.28)	0.15,1.02	0.01	(1.56)	−0.42,0.44	3.42	.068	0.40 (−0.04,0.82)	0.994
Same example statement-present										
Plausibility (text analyses)	2.16	(2.31)	1.49,2.82	−0.22	(2.17)	−0.92,0.48	23.95	<.001	1.06 (0.59,1.50)	3389.276
Immediacy (text analyses)	1.79	(2.14)	1.19,2.39	0.15	(1.87)	−0.48,0.77	14.26	<.001	0.81 (0.36,1.24)	85.945
Clarity (text analyses)	1.46	(1.81)	0.93,1.98	−0.33	(1.74)	−0.88,0.22	21.56	<.001	1.01 (0.54,1.44)	1407.061

Note. NHST = Null-hypothesis significance testing.

In responses to the eliciting-opinion request, a significant effect for plausibility emerged in the same example statement-absent condition but there was not sufficient evidence for the effect. Significant differences emerged for all three variables in the same example statement-present condition with sufficient Bayes Factor support for plausibility and clarity: Truth tellers sounded more plausible and clearer than lie tellers.

Regarding the differences in responses to the eliciting-opinion and devil's advocate requests (residue effects), in the same example statement-absent condition a significant effect for plausibility emerged, but with insufficient Bayes Factor support. All three effects were significant in the same example statement-present condition with sufficient Bayes Factor support for all three effects. The three residue effects were larger in truth tellers than in lie tellers. In truth tellers they differed from '0' for plausibility, $t(44) = 6.25$, $p < .001$, immediacy, $t(44) = 5.61$, $p < .001$ and clarity, $t(44) = 5.38$, $p < .001$, whereas they did not differ from '0' for plausibility, $t(40) = 0.65$, $p = .522$, immediacy, $t(40) = 0.50$, $p = .619$ and clarity, $t(40) = 1.21$, $p = .233$ in lie tellers.

Table 4A
Interview responses as a function of same example statement.

Interview responses	Same example statement-present			Same example statement-absent			NHST			BF_{10}
	M	*(SD)*	95% CI	*M*	*(SD)*	95% CI	*F*	*p*	*d*	
Main effect results										
Plausibility (text analyses)	4.35	(0.61)	4.21,4.50	4.10	(0.78)	3.95,4.25	5.76	.017		2.329
Immediacy (text analyses)	4.06	(0.77)	3.88,4.23	3.66	(0.87)	3.59,3.84	9.84	.002		14.587
Clarity (text analyses)	4.37	(0.70)	4.21,4.53	4.26	(0.79)	4.10,4.42	0.89	.348		0.249

Note. NHST = Null-hypothesis significance testing.

The responses sounded more plausible and more immediate in the same example statement-present condition than in the same example statement-absent condition, with only sufficient Bayes Factors support for the immediacy variable.

Appendix 5. The results for eloquence, predictability and the three separate eloquence variables (plausibility, immediacy and clarity) for the VIDEO analyses

Table 2B
Interview responses as a function of veracity.

Interview responses	Truth			Lie			NHST			BF_{10}
	M	*(SD)*	95% CI	*M*	*(SD)*	95% CI	*F*	*p*	*d*	
Eliciting-opinion request										
Eloquence (video analyses)	5.05	(0.91)	4.88,5.21	4.72	(0.66)	4.55,4.90	7.02	.009	0.41 (0.10,0.71)	4.11
Predictability (video analyses)	2.82	(0.94)	2.63,3.00	3.03	(0.80)	2.84,3.22	2.54	.113	0.24 (−0.06,0.54)	1.87
Plausibility (video analyses)	5.03	(1.03)	4.82,5.23	4.60	(0.90)	4.39,4.81	8.27	.005	0.44 (0.13,0.74)	7.23
Immediacy (video analyses)	4.94	(1.12)	4.74,5.15	4.71	(0.79)	4.50,4.92	2.49	.117	0.24 (−0.07,0.53)	1.92
Clarity (video analyses)	5.17	(1.01)	4.98,5.35	4.86	(0.72)	4.67,5.05	5.08	.025	0.35 (0.05,0.65)	1.71
Eliciting-opinion minus devil's advocate request										
Eloquence (video analyses)	1.40	(1.04)	1.19,1.60	0.80	(0.85)	0.60,1.01	16.59	<.001	0.63 (0.31,0.93)	282.22
Predictability (video analyses)	−0.93	(1.18)	−1.16,-0.70	−0.85	(0.95)	−1.08,-0.61	0.27	.602	0.07 (−0.23,0.37)	0.19
Plausibility (video analyses)	1.52	(1.30)	1.25,1.78	0.79	(1.21)	0.52,1.06	14.09	<.001	0.58 (0.27,0.88)	95.27
Immediacy (video analyses)	1.24	(1.14)	1.01,1.46	0.83	(0.96)	0.60,1.06	6.36	.013	0.39 (0.08,0.69)	3.06
Clarity (video analyses)	1.44	(1.26)	1.20,1.69	0.79	(1.03)	0.54,1.03	13.89	<.001	0.56 (0.25,0.86)	87.18

Note. NHST = Null-hypothesis significance testing.

In response to the eliciting-opinion request significant Veracity effects emerged for eloquence, plausibility and clarity with those for eloquence and plausibility receiving sufficient Bayes Factor support. Truth tellers sounded more eloquent than lie tellers particularly because they sounded more plausible.

Regarding the differences in responses to the eliciting-opinion and devil's advocate requests (residue effects), significant Veracity effects emerged for eloquence, plausibility, immediacy and clarity, with all these effects receiving sufficient support. These residue scores were larger for truth tellers than for lie tellers. For truth tellers the residue scores differed from '0' for all five variables: eloquence, $t(86) = 12.48$, $p < .001$, predictability, $t(86) = 7.34$, $p < .001$, plausibility, $t(86) = 10.86$, $p < .001$, immediacy, $t(86) = 10.08$, $p < .001$, and clarity, $t(86) = 10.71$, $p < .001$. For lie tellers the residue scores also differed from '0' for all five variables: eloquence, $t(83) = 8.61$, $p < .001$, predictability, $t(83) = 8.15$, $p < .001$, plausibility, $t(83) = 6.02$, $p < .001$, immediacy, $t(83) = 7.89$, $p < .001$, and clarity, $t(83) = 6.97$, $p < .001$.

Table 3B
Interview responses as a function of veracity for each same example statement condition.

Interview responses	Truth			Lie			NHST			BF_{10}
	M	*(SD)*	95% CI	*M*	*(SD)*	95% CI	*F*	*p*	*d*	
Eliciting-opinion request										
Same example statement-absent										
Eloquence (video analyses)	4.77	(0.70)	4.57,4.98	4.58	(0.65)	4.38,4.79	1.74	.190	0.28 (−0.15,0.70)	0.484
Predictability (video analyses)	3.24	(0.91)	2.97,3.51	3.23	(0.83)	2.97,3.50	0.01	.977	0.01 (−0.41,0.44)	0.226
Plausibility (video analyses)	4.75	(0.87)	4.48,5.02	4.40	(0.88)	4.13,4.66	3.50	.065	0.40 (−0.04,0.82)	1.030
Immediacy (video analyses)	4.61	(0.98)	4.33,4.88	4.56	(0.80)	4.29,4.83	0.06	.802	0.06 (−0.37,0.48)	0.233
Clarity (video analyses)	4.96	(0.87)	4.72,5.21	4.79	(0.73)	4.55,5.03	1.01	.319	0.21 (−0.22,0.64)	0.351
Same example statement-present										
Eloquence (video analyses)	5.30	(1.01)	5.05,5.55	4.87	(0.64)	4.61,5.14	5.34	.023	0.50 (0.07,0.93)	2.244
Predictability (video analyses)	2.42	(0.79)	2.20,2.65	2.82	(0.72)	2.58,3.05	5.81	.018	0.53 (0.09,0.95)	2.741
Plausibility (video analyses)	5.29	(1.11)	4.99,5.69	4.82	(0.89)	4.50,5.13	4.66	.034	0.46 (0.03,0.89)	1.684
Immediacy (video analyses)	5.26	(1.15)	4.96,5.55	4.87	(0.76)	4.56,5.17	3.37	.070	0.40 (−0.04,0.82)	0.970
Clarity (video analyses)	5.36	(1.11)	5.08,5.64	4.93	(0.72)	4.65,5.23	4.20	.043	0.46 (0.02,0.88)	1.387
Eliciting-opinion minus devil's advocate request										
Same example statement-absent										
Eloquence (video analyses)	1.15	(0.80)	0.90,1.39	0.72	(0.79)	0.48,0.96	6.16	.015	0.54 (0.10,0.97)	3.177

(*continued on next page*)

Table 3B (*continued*)

Interview responses	Truth			Lie			NHST			BF_{10}
	M	*(SD)*	95% CI	*M*	*(SD)*	95% CI	*F*	*p*	*d*	
Predictability (video analyses)	−0.44	(0.86)	−0.70,-0.18	−0.71	(0.83)	−0.97,-0.45	2.13	.148	0.32 (−0.11,0.74)	0.573
Plausibility (video analyses)	1.20	(1.09)	0.85,1.55	0.66	(1.19)	0.32,1.01	4.76	.032	0.47 (0.03,0.90)	1.761
Immediacy (video analyses)	0.96	(0.95)	0.67,1.26	0.69	(0.96)	0.40,0.98	1.80	.183	0.28 (−0.15,0.71)	0.497
Clarity (video analyses)	1.27	(0.94)	0.99,1.56	0.81	(0.89)	0.54,1.09	5.36	.023	0.50 (0.06,0.93)	2.268
Same example statement-present										
Eloquence (video analyses)	1.63	(1.19)	1.31,1.95	0.89	(0.92)	0.55,1.22	10.31	.002	0.69 (0.25,1.12)	17.676
Predictability (video analyses)	−1.39	(1.26)	−1.74,-1.04	−0.99	(1.05)	−1.35,-0.63	2.54	.115	0.34 (−0.09,0.76)	0.680
Plausibility (video analyses)	1.80	(1.42)	1.41,2.19	0.93	(1.22)	0.51,1.34	9.25	.003	0.65 (0.21,1.08)	11.464
Immediacy (video analyses)	1.49	(1.25)	1.16,1.82	0.98	(0.95)	0.63,1.32	4.49	.037	0.46 (0.02,0.88)	1.569
Clarity (video analyses)	1.60	(1.49)	1.20,2.00	0.76	(1.17)	0.34,1.17	8.42	.005	0.62 (0.18,1.05	8.148

Note. NHST = Null-hypothesis significance testing.

In responding to the eliciting-opinion request no Veracity differences emerged in the same example statement-absent condition. All Veracity effects except for immediacy were significant in the same example statement-present condition but the Bayes Factor results showed insufficient support for all these significant effects.

Regarding the differences to the eliciting-opinion and devil's advocate requests (residue scores), only the Veracity effect for eloquence was significant in the same example statement-absent condition. It also had sufficient Bayes Factor support. The difference was larger for truth tellers than for lie tellers. The residue scores differed from '0' for truth tellers for all five variables: eloquence, $t(41) = 9.34$, $p < .001$, predictability, $t(41) = 3.30$, $p = .002$, plausibility, $t(41) = 7.17$, $p < .001$, immediacy, $t(41) = 6.56$, $p < .001$, and clarity, $t(41) = 8.80$, $p < .001$. For lie tellers the residue scores also differed from '0' for all five variables: eloquence, $t(42) = 6.01$, $p < .001$, predictability, $t(42) = 5.59$, $p < .001$, plausibility, $t(42) = 3.66$, $p < .001$, immediacy, $t(42) = 4.70$, $p < .001$, and clarity, $t(42) = 5.98$, $p < .001$.

In the same example statement-present condition all effects apart the effect for predictability were significant. The Bayes Factor analyses showed sufficient support for eloquence, plausibility and clarity. For truth tellers the residue scores differed from '0' for all five variables: eloquence, $t(44) = 9.16$, $p < .001$, predictability, $t(44) = 7.39$, $p < .001$, plausibility, $t(44) = 8.50$, $p < .001$, immediacy, $t(44) = 7.96$, $p < .001$, and clarity, $t(44) = 7.22$, $p < .001$. For lie tellers the residue scores also differed from '0' for all five variables: eloquence, $t(40) = 6.16$, $p < .001$, predictability, $t(40) = 6.01$, $p < .001$, plausibility, $t(40) = 4.85$, $p < .001$, immediacy, $t(40) = 6.54$, $p < .001$, and clarity, $t(40) = 4.13$, $p < .001$.

Table 4B
Interview responses as a function of same example statement.

Interview responses	Same example statement-present			Same example statement-absent			NHST			BF_{10}
	M	*(SD)*	95% CI	*M*	*(SD)*	95% CI	*F*	*p*	*d*	
Main effect results										
Eloquence (video analyses)	4.46	(0.57)	4,34,4.58	4.21	(0.56)	4.09,4.33	8.29	.005	0.44 (0.13,0.74)	7.277
Predictability (video analyses)	3.21	(0.60)	3.07,3.35	3.52	(0.73)	3.38,3.67	9.47	.002	0.46 (0.15,0.76)	12.365
Plausibility (video analyses)	4.37	(0.74)	4.22,4.52	4.11	(0.67)	3.95,4.26	6.03	.015	0.37 (0.06,0.66)	2.631
Immediacy (video analyses)	4.45	(0.58)	4.31,4.58	4.17	(0.69)	4.03,4.31	8.08	.005	0.44 (0.13,0.74)	6.623
Clarity (video analyses)	4.56	(0.72)	4.41,4.71	4.36	(0.68)	4.21,4.51	3.54	.062	0.29 (−0.02,0.58)	1.183

Note. NHST = Null-hypothesis significance testing.

Compared to the same example statement-absent condition, the responses in the same example statement-present condition sounded more eloquent (more plausible and more immediate) and less predictable. The Bayes Factor results showed sufficient support for eloquence (due to.

Table 6A
Summary of the significant veracity results regarding plausibility, immediacy and clarity.

Interview responses	Text analyses			Video analyses		
	All participants	Same example statement-absent	Same example statement-present	All participants	Same example statement-absent	Same example statement-present
	d	*d*	*d*	*d*	*d*	*d*
Eliciting-opinion request						
Plausibility	0.80		1.15	0.44		
Immediacy						
Clarity	0.47		1.00			
Eliciting-opinion minus devil's advocate request						
Plausibility	0.56		1.06	0.58		0.65
Immediacy	0.32		0.81	0.39		
Clarity	0.60		1.01	0.56		0.62

Note. A positive *d*-score indicates a higher score for truth tellers than for lie tellers, whereas a negative *d*-score indicates a higher score for lie tellers than for truth tellers.

Table 6A summarises the Veracity effects which were significant ($p < .05$) with sufficient evidence for the alternative hypothesis ($BF_{10} > 3$). Veracity effects were more pronounced in the text than video analyses for the eliciting-opinion responses. The residue effects are remarkably similar in the text and video analyses for the entire sample, but more pronounced for text analyses than for video analyses in the same example statement-present condition.

Data availability

Data are made available upon reasonable request to the second author.

References

Ajzen, I. (2001). Nature and operation of attitudes. *Annual Review of Psychology, 52*, 27–58. https://doi.org/10.1146/annurev.psych.52.1.27

Amado, B. G., Arce, R., & Fariña, F. (2015). Undeutsch hypothesis and criteria based content analysis: A meta-analytic review. *The European Journal of Psychology Applied to Legal Context, 7*(1), 3–12. https://doi.org/10.1016/j.ejpal.2014.11.002

Amado, B. G., Arce, R., Fariña, F., & Vilarino, M. (2016). Criteria-based content analysis (CBCA) reality criteria in adults: A meta-analytic review. *International Journal of Clinical and Health Psychology, 16*(2), 201–210. https://doi.org/10.1016/j.ijchp.2016.01.002

Blake, K. R., & Gangestad, S. (2020). On attenuated interactions, measurement error, and statistical power: Guidelines for social and personality psychologists. *Personality and Social Psychology Bulletin, 46*(12), 1702–1711. https://doi.org/10.1177/0461672209133613363

Brimbal, L., Dianiska, R. E., Swanner, J. K., & Meissner, C. A. (2019). Enhancing cooperation and disclosure by manipulating affiliation and developing rapport in investigative interviews. *Psychology, Public Policy, and Law, 25*(2), 107–115. https://doi.org/10.1037/law0000193

Clemens, F., Granhag, P. A., & Strömwall, L. A. (2013). Counter-interrogation strategies when anticipating requests on intentions. *Journal of Investigative Psychology and Offender Profiling, 10*, 125–138. https://doi.org/10.1002/jip.1387

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Colwell, K., Hiscock-Anisman, C. K., Memon, A., Taylor, L., & Prewett, J. (2007). Assessment criteria indicative of deception (ACID): An integrated system of investigative interviewing and detecting deception. *Journal of Investigative Psychology and Offender Profiling, 4*(3), 167–180. https://doi.org/10.1002/jip.73

Darley, J. M., & Gross, P. H. (1983). A hypothesis-confirming bias in labelling effects. *Journal of Personality and Social Psychology, 44*(1), 20–33. https://doi.org/10.1037/0022-3514.44.1.20

Deeb, H., Vrij, A., Hope, L., Mann, S., Granhag, P. A., & Lancaster, G. (2017). Suspects' consistency in statements concerning two events when different question formats are used. *Journal of Investigative Psychology and Offender Profiling, 14*(1), 74–87. https://doi.org/10.1002/jip.1464

Deeb, H., Vrij, A., Hope, L., Mann, S., Leal, S., Granhag, P. A., & Strömwall, L. A. (2018). The devil's advocate approach: An interview technique for assessing consistency among deceptive and truth-telling pairs of suspects. *Legal and Criminological Psychology, 23*(1), 37–52. https://doi.org/10.1111/lcrp.12114

Denault, V., Patterson, M. L., Zloteanu, M., & Talwar, V. (2025). The myth of body language and the fallacies of body language analysis and training programs. In D. Chadee, & A. Kostic (Eds.), *Body language communication* (pp. 151–186). London: Palgrave-McMillian. https://doi.org/10.1007/978-3-031-70064-4_15.

DePaulo, B. M., Lindsay, J. L., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129*(1), 74–118. https://doi.org/10.1037/0033-2909.129.1.74

DePaulo, B. M., & Morris, W. L. (2004). Discerning lies from truths: Behavioural cues to deception and the indirect pathway of intuition. In P. A. Granhag, & L. A. Strömwall (Eds.), *Deception detection in forensic contexts* (pp. 15–40). Cambridge, England: Cambridge University Press.

Ewens, S., Vrij, A., Leal, S., Mann, S., Jo, E., Shaboltas, A., Ivanova, M., Granskaya, J., & Houston, K. (2016). Using the model statement to elicit information and cues to deceit from native speakers, non-native speakers and those talking through an interpreter. *Applied Cognitive Psychology, 30*(6), 854–862. https://doi.org/10.1002/acp.3270

Gancedo, Y., Fariña, F., Seijo, D., Vilariño, M., & Arce, R. (2021). Reality monitoring: A meta-analytical review for forensic practice. *The European Journal of Psychology Applied to Legal Context, 13*(2), 99–110. https://doi.org/10.5093/ejpalc2021a10

Gilovich, T. (1990). Differential construal and the false consensus effect. *Journal of Personality and Social Psychology, 59*(4), 623–634. https://doi.org/10.1037/0022-3514.59.4.623

Granhag, P. A., & Hartwig, M. (2008). A new theoretical perspective on deception detection: On the psychology of instrumental mind-reading. *Psychology, Crime & Law, 14*(3), 189–200. https://doi.org/10.1080/10683160701645181

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorial in Quantitative Methods for Psychology, 8*(1), 23–34. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3402032/.

Hart, W., Albarracín, D., Eagly, A. H., Brechan, I., Lindberg, M. J., & Merrill, L. (2009). Feeling validated versus being correct: A meta-analysis of selective exposure to information. *Psychological Bulletin, 135*(4), 555–588. https://doi.org/10.1037/a0015701

Hartwig, M., & Bond, C. F. (2011). Why do lie-catchers fail? A lens model meta-analysis of human lie judgments. *Psychological Bulletin, 137*(4), 643–659. https://doi.org/10.1037/a0023589

Hartwig, M., & Bond, C. F. (2014). Lie detection from multiple cues: A meta-analysis. *Applied Cognitive Psychology, 28*(5), 661–667. https://doi.org/10.1002/acp.3052

Hartwig, M., & Granhag, P. A. (2023). Strategic use of evidence (SUE): A review of the technique and its principles. In G. Oxburgh, T. Myklebust, M. Fallon, & M. Hartwig (Eds.), *Interviewing and interrogation: A review of research and practice since world war II* (pp. 299–318). Brussels, Belgium: Torkel Opsahl Academic (EPublisher).

Hartwig, M., Granhag, P. A., & Strömwall, L. (2007). Guilty and innocent suspects' strategies during police interrogations. *Psychology, Crime & Law, 13*(2), 213–227. https://doi.org/10.1080/10683160600750264

Jarosz, A. F., & Wiley, J. (2014). What are the odds? A practical guide to computing and reporting Bayes factors. *The Journal of Problem Solving, 7*(1), 2. https://doi.org/10.7771/1932-6246.1167

Jawed, S., Amin, H. U., Malik, A. S., & Faye, I. (2019). Classification of visual and non-visual learners using electroencephalographic alpha and gamma activities. *Frontiers in Behavioral Neuroscience, 13*, 86. https://doi.org/10.3389/fnbeh.2019.00086

Knobloch-Westerwick, Mothes, C., & Polavin, N. (2020). Confirmation bias, ingroup bias, and negativity bias in selective exposure to political information. *Communication Research, 47*(1), 104–124. https://doi.org/10.1177/0093650217719596

Knobloch-Westerwick, S. (2015). The selective exposure self- and affect-management (SESAM) model: Applications in the realms of race, politics, and health. *Communication Research, 42*, 959–985. https://doi.org/10.1177/0093650214539173

Leal, S., Vrij, A., Deeb, H., & Dabrowna, O. (2023). Combining the devil's advocate approach and verifiability approach to assess veracity in opinion statements. *The European Journal of Psychology Applied to Legal Context, 15*(2), 53–61. https://doi.org/10.5093/ejpalc2023a6

Leal, S., Vrij, A., Deeb, H., & Fisher, R. P. (2024). In my opinion you're wrong! Adding a model statement to the devil's advocate approach to detect true and false opinions. *Journal of Applied Research in Memory and Cognition, 14*(2), 241–254. https://doi.org/10.1037/mac0000201

Leal, S., Vrij, A., Mann, S., & Fisher, R. (2010). Detecting true and false opinions: The devil's advocate approach as a lie detection aid. *Acta Psychologica, 134*(3), 323–329. https://doi.org/10.1016/j.actpsy.2010.03.005

Leal, S., Vrij, A., Warmelink, L., Vernham, Z., & Fisher, R. (2015). You cannot hide your telephone lies: Providing a model statement as an aid to detect deception in insurance telephone calls. *Legal and Criminological Psychology, 20*(1), 129–146. https://doi.org/10.1111/lcrp.12017

Leins, D., Fisher, R. P., & Ross, S. J. (2013). Exploring liars' strategies for creating deceptive reports. *Legal and Criminological Psychology, 18*(1), 141–151. https://doi.org/10.1111/j.2044-8333.2011.02041.x

Mac Giolla, E., & Luke, T. (2021). Does the cognitive approach to lie detection improve the accuracy of human observers? *Applied Cognitive Psychology, 35*(2), 385–392. https://doi.org/10.1002/acp.3777

Mann, S., Vrij, A., Deeb, H., & Leal, S. (2022). Actions speak louder than words: The devil's advocate questioning protocol in opinions about protester actions. *Applied Cognitive Psychology, 36*, 905–918. https://doi.org/10.1002/acp.3979

Mann, S., Vrij, A., Deeb, H., & Leal, S. (2023). All mouth and trousers?: Use of the devil's advocate questioning protocol to determine authenticity of opinions about protester actions. *Psychiatry, Psychology and Law*. https://doi.org/10.1080/13218719.2023.2242433

Markowitz, D. M. (2024). Deconstructing deception: Frequency, communicator characteristics, and linguistic features of embeddedness. *Applied Cognitive Psychology, 38*(3), Article e4215. https://doi.org/10.1002/acp.4215

McLeod, S. (2023). What does effect size tell you?. https://www.simplypsychology.org/effect-size.html.

Mochon, D., & Schwartz, J. (2024). The confrontation effect: When users engage more with ideology-inconsistent content online. *Organizational Behavior and Human Decision Processes, 185*, Article 104366. https://doi.org/10.1016/j.obhdp.2024.104366

Mullen, B., Atkins, J. L., Champion, D. S., Edwards, C., Hardy, D., Story, J. E., & Vanderklok, M. (1985). The false consensus effect: A meta-analysis of 115 hypothesis tests. *Journal of Experimental Social Psychology, 21*(3), 262–283. https://doi.org/10.1016/0022-1031(85)90020-4

Nahari, G. (2019). Verifiability approach: Applications in different judgmental settings. In T. Docan-Morgan (Ed.), *The Palgrave handbook of deceptive communication* (pp. 213–225). New York, NY: United States: Palgrave Macmillan. https://doi.org/10.1007/978-3-319-96334-1.

Nickerson, R. S. (1998). Confirmation Bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2*(2), 175–220. https://doi.org/10.1037/1089-2680.2.2.175

Nisbett, R. E., & Wilson, T. D. (1977). The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology, 35*(4), 250–256. https://doi.org/10.1037/0022-3514.35.4.250

Oleszkiewicz, S., & Watson, S. J. (2021). A meta-analytic review of the timing for disclosing evidence when interviewing suspects. *Applied Cognitive Psychology, 35*(2), 342–359. https://doi.org/10.1002/acp.3767

Palena, N., Caso, L., Vrij, A., & Nahari, G. (2021). The verifiability approach: A meta-analysis. *Journal of Applied Research in Memory and Cognition, 10*(1), 155–166. https://doi.org/10.1016/j.jarmac.2020.09.001

Patterson, M. L. (1995). Invited article: A parallel process model of nonverbal communication. *Journal of Nonverbal Behavior, 19*(1), 3–29. https://doi.org/10.1007/BF02173410

Patterson, M. L. (2006). The evolution of theories of interactive behavior. In V. Manusov, & M. L. Patterson (Eds.), *The SAGE handbook of nonverbal communication* (pp. 21–40). Thousand Oaks, CA: Sage. https://doi.org/10.4135/9781412976152.n2.

Patterson, M. L., Fridlund, A. J., & Crivelli, C. (2023). Four misconceptions about nonverbal communication. *Perspectives on Psychological Science, 18*(6), 1388–1411. https://doi.org/10.1177/17456916221148142

Rice, M., & Harris, G. T. (2005). Comparing effect sizes in follow-up studies: ROC area, Cohen's d, and r. *Law and Human Behavior, 29*(5), 615–620. https://doi.org/10.1007/s10979-005-6832-7

Ross, L., Greene, D., & House, P. (1977). The 'false consensus effect': An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology, 13*(3), 279–301. https://doi.org/10.1016/0022-1031(77)90049-X

Smith, C. A., & Ellsworth, P. C. (1985). Patterns of cognitive appraisal in emotion. *Journal of Personality and Social Psychology, 48*(4), 813–838. https://doi.org/10.1037/0022-3514.48.4.813

Sommer, J., Musolino, J., & Hemmer, P. (2024). Updating, evidence evaluation, and operator availability: A theoretical framework for understanding belief. *Psychological Review, 131*(2), 371–401. https://doi.org/10.1037/rev0000444

Sporer, S. L., & Masip, J. (2025). A systematic review of the validity of criteria-based content analysis in child sexual abuse cases and other field studies. *Psychology, Crime & Law, 31*(9), 1142–1183. https://doi.org/10.1080/1068316X.2024.2335971

Sporer, S. L., & Schwandt, B. (2006). Paraverbal indicators of deception: A meta-analytic synthesis. *Applied Cognitive Psychology, 20*(4), 421–446. https://doi.org/10.1002/acp.1190

Vallano, J. P., & Schreiber Compo, N. (2011). A comfortable witness is a good witness: Rapport-building and susceptibility to misinformation in an investigative mock-crime interview. *Applied Cognitive Psychology, 25*(6), 960–970. https://doi.org/10.1002/acp.1789

Vallerand, R. J. (2015). *The psychology of passion: A dualistic model, series in positive psychology*. New York: Oxford Academic Press. https://doi.org/10.1093/acprof:oso/9780199777600.001.0001

Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities* (2nd ed.). Chichester: John Wiley and Sons. ISBN: 978-0-470-51624-9.

Vrij, A., & Fisher, R. P. (2019). Nonverbal cues to deception in title IX investigations. *Journal of Applied Research in Memory and Cognition, 8*(4), 417–419. https://doi.org/10.1016/j.jarmac.2019.07.006

Vrij, A., & Fisher, R. P. (2020). Unravelling the misconception about deception and nervous behaviour. *Frontiers in Psychology. Personality and Social Psychology, 11*, 1377. https://doi.org/10.3389/fpsyg.2020.01377

Vrij, A., Fisher, R. P., & Leal, S. (2023). How researchers can make verbal lie detection more attractive for practitioners. *Psychiatry, Psychology and Law, 30*(3), 383–396. https://doi.org/10.1080/13218719.2022.2035842

Vrij, A., & Granhag, P. A. (2012). Eliciting cues to deception and truth: What matters are the questions asked. *Journal of Applied Research in Memory and Cognition, 1*(2), 110–117. https://doi.org/10.1016/j.jarmac.2012.02.004

Vrij, A., Granhag, P. A., Ashkenazi, T., Ganis, G., Leal, S., & Fisher, R. P. (2022). Verbal lie detection: Its past, present and future. *Brain Sciences, 12*(12), 1644. https://doi.org/10.3390/brainsci12121644

Vrij, A., Leal, S., & Fisher, R. P. (2018). Verbal deception and the model statement as a lie detection tool. *Frontiers in Psychiatry, section Forensic Psychiatry, 9*, 492. https://doi.org/10.3389/fpsyt.2018.00492

Vrij, A., Leal, S., & Fisher, R. P. (2023). Interviewing to detect lies about opinions: The devil's advocate approach. *Advances in Social Sciences Research Journal, 10*(12), 245–252. https://doi.org/10.14738/assrj.1012.16027

Vrij, A., Mann, S., Leal, S., & Fisher, R. P. (2021). Combining verbal veracity assessment techniques to distinguish truth tellers from lie tellers. *European Journal of Psychology Applied to Legal Context, 13*(1), 9–19. https://doi.org/10.5093/ejpalc2021a2

Vrij, A., Meissner, C. A., Fisher, R. P., Kassin, S. M., Morgan, A., III, & Kleinman, S. (2017). Psychological perspectives on interrogation. *Perspectives on Psychological Science, 12*(6), 927–955. https://doi.org/10.1177/1745691617706515

Vrij, A., Palena, N., Leal, S., & Caso, L. (2021). The relationship between complications, common knowledge details and self-handicapping strategies and veracity: A meta-analysis. *The European Journal of Psychology Applied to Legal Context, 13*(2), 55–77. https://doi.org/10.5093/ejpalc2021a7