Actions speak louder than words:

The Devil's Advocate questioning protocol in opinions about protester actions

Samantha Mann[1]

Aldert Vrij

Haneen Deeb

Sharon Leal

Department of Psychology, University of Portsmouth (UK)

This work was funded by the Centre for Research and Evidence on Security Threats (ESRC Award: ES/N009614/1)

[1] Correspondence concerning this article should be addressed to Samantha Mann, Department of Psychology, University of Portsmouth, King Henry Building, King Henry 1 Street, PO1 2DY, Hants, United Kingdom. Email: samantha.mann@port.ac.uk

None of the listed authors have any conflict of interest to declare

Abstract

The Devil's Advocate protocol has been developed to assist making veracity assessments when someone discusses their opinion (Leal et al., 2010). The present experiment focused on protester actions rather than controversial issues (Leal et al., 2010) and also included an adapted version of the Verifiability Approach.

Participants told the truth or lied about protester actions and the participants' answers to the eliciting opinion and Devil's Advocate questions were compared. The Devil's Advocate approach predicts the difference in answers (residue scores) to be more pronounced in truth tellers than in lie tellers in terms of quantity of the answers (number of words, details and arguments) and quality of the answers (plausibility, immediacy and clarity).

The hypothesis was supported but only in terms of quality: Truth tellers' answers sounded more plausible and immediate and somewhat clearer than lie tellers' answers. Truth tellers also reported more digital verifiable sources than lie tellers.

Actions speak louder than words:

The Devil's Advocate questioning protocol in opinions about protester actions

Research has shown that verbal and nonverbal cues lie tellers spontaneously display are weak and unreliable (DePaulo et al., 2003; DePaulo & Morris, 2004). Thus, researchers have started to develop interview protocols that enhance or elicit such cues. To date, at least four interview protocols have been developed, all assessing verbal cues to deceit: Assessment Criteria Indicative of Deception (Bogaard et al., 2019; Colwell et al., 2015), Cognitive Credibility Assessment (Vrij et al., 2015, 2017; Vrij, Mann et al., 2021), Strategic Use of Evidence (Granhag & Hartwig, 2015; Hartwig et al., 2014) and the Verifiability Approach (Nahari, 2019; Palena et al., 2021). These four protocols have in common that they examine truths and lies when interviewees provide statements about their alleged activities. However, sometimes practitioners, particularly those dealing with immigration, intelligence or security clearance (vetting), are interested in whether interviewees are truthful or deceptive when they express their attitudes (opinions) about issues. Examples include someone expressing their opinion about the actions of the Government or their attitude towards a country, desisting from terrorist activities, becoming an informant, or performing specific protester actions. To not successfully detect deceit when interviewing people about such issues can have far reaching consequences. It can result in violent protests against Government actions, allowing terrorists to enter a country and carry out terrorist activities, introducing double agents in an intelligence agency etc.

The Devil's Advocate Protocol

The Devil's Advocate protocol has been developed to make veracity assessments when someone discusses their opinion (Leal et al., 2010). This, to our knowledge, is the only interview protocol aimed at detecting lying about opinions, and consists of three questions. In

the first 'expressing attitude' question the interviewee is invited to express their attitude about the topic under investigation in terms of good or bad: "What do you think of the Iraq war in terms of good or bad?" The second question is the eliciting opinion question and invites interviewees to report the arguments in favour of the view they just expressed: "Provide all the reasons why you are in favour of the Iraq war?" The third question is the devil's advocate question in which interviewees are asked to give arguments opposed to their expressed view: "Try to play devil's advocate and imagine that you are against the Iraq war. Provide all the reasons why you may be against the Iraq war?"

When people have developed a strong attitude towards a certain issue they typically have thought of the reasons as to why they have that opinion (Azjen, 2001). The confirmation bias predicts that those reasons should be more readily available to respondents than reasons that oppose their view because people tend to seek information that supports rather than opposes their views (Darley & Gross, 1983; Jones & Sugden, 2001). In other words, truth tellers should find it easier to answer the eliciting opinion question than the devil's advocate question. This should be revealed when these two questions are compared with truth tellers giving more information (e.g., more arguments) that is of higher quality (e.g., more plausible) when answering the eliciting opinion question than the devil's advocate question.

The situation is different for lie tellers. They often prepare themselves for interviews (Clemens et al., 2013; Deeb et al., 2018). In this context that involves thinking, prior to the interview, about arguments they can present that support their pretended opinion. They should thus be prepared to answer the eliciting opinion question. Unlike in truth tellers, the devil's advocate question invites lie tellers to give their true opinion and these arguments should be readily available to them. This means that for lie tellers the difference in difficulty in answering the eliciting opinion and devil's advocate questions should be less pronounced

than in truth tellers, making the difference in answering these two questions in terms of quantity and quality less pronounced than in truth tellers.

This different pattern of results was indeed obtained by Leal et al. (2010), the only Devil's Advocate protocol experiment published to date with individual interviewees (Deeb et al. [2020] examined pairs of interviewees). In Leal et al. (2010) participants told the truth or lied when discussing controversial issues (e.g., "The invasion of Iraq was necessary") whereas in the current experiment participants told the truth or lied when discussing their views about protester actions (e.g., whether it is acceptable for a person to graffiti over a bank because it is the largest funder of fossil fuels). We expect the Devil's Advocate Approach also to be an effective veracity assessment tool when participants discuss their opinions about protester actions. If true, it would widen the applicability of the Devil's Advocate Approach, which, to date, has solely been tested to examine deception when discussing societal issues.

We made two changes to the Devil's Advocate interview protocol as used in Leal et al. (2010). First, Leal et al. (2010) did not analyse the answers given to the first question in which the interviewee is asked to express their attitude, but in the present experiment we did. Truth tellers should find it easier to discuss their real attitude than lie tellers should find it to express a false attitude, and this should be reflected in the answer both in terms of quantity and quality.

Second, we added a fourth question to the protocol in which we asked participants to take on the role of a script writer and to write a part in a film for a character who strongly holds the opposite view to the views they supported in the interview. It thus invites truth tellers to report views opposite to their real views but lie tellers to report their real views. It is thus an alternative way of asking the devil's advocate question. We added this question after discussions with practitioners who thought that the phrasing of the devil's advocate question was somewhat complicated to use, and obvious to the interviewee of what was being

assessed, thus open to their implementing counter-strategies. Hence we devised this question as an alternative way of presenting the devil's advocate question, to combat perceived practitioner issues and explore how truth tellers and lie tellers would respond to it.

Also, unlike Leal et al. (2010), we included an adapted version of the Verifiability Approach in our interview protocol in addition to the four Devil's Advocate protocol questions. The Verifiability Approach, discussed in the next section, has a different rationale and examines different verbal cues to the Devil's Advocate Approach. As such, a combination of the Devil's Advocate Approach and the Verifiability Approach may facilitate lie detection, because it examines multiple facets of lying about opinions.

Verifiability Sources

According to the Verifiability Approach, lie tellers are less willing and less able than truth tellers to provide details an investigator can check (Nahari et al., 2014). Verifiable details are activities: (i) carried out, (ii) witnessed by named witnesses, (iii) caught on CCTV or (iv) leaving digital traces (e.g., receipts, photos, phone calls, bank transfers). Lie tellers may be less willing than truth tellers to provide verifiable details because they do not wish to incriminate themselves (Granhag & Hartwig, 2008). Lie tellers may also be less able to do so because it will be difficult to present evidence for their made-up statements. Meta-analyses have supported this hypothesis and found that lie tellers indeed report fewer verifiable details than truth tellers (Palena et al., 2021; Verschuere et al., 2021).

To date the Verifiability Approach has been examined in several settings, including criminal (e.g., Jupe et al., 2020), insurance (e.g., Harvey et al., 2017) and border control settings (e.g., Vrij et al., 2019), but never in an opinion setting. Since people are likely to have discussed their opinions with friends and family or perhaps on social media, we think that the Verifiability Approach could also discriminate between truth tellers and lie tellers in an opinion setting.

In the Verifiability Approach, the number of verifiable *details* are typically counted. Researchers recently started to examine whether the approach also discriminates between truth tellers and lie tellers if verifiable *sources* are counted (Leal et al., 2018; Vrij et al., 2019, 2021). Verifiable details and verifiable sources are linked to each other, because a verifiable source is a requirement for verifiable details to occur. For example, the sentence "I made a phone call to my bank in the afternoon" contains three verifiable details (phone call, bank, afternoon) and these details became verifiable due to the mentioning of a verifiable source (phone call). Research has shown that lie tellers report fewer verifiable sources than truth tellers (Leal et al., 2018; Vrij et al., 2019, 2021). Counting verifiable sources is much easier to do than counting verifiable details, which would make it easier for practitioners to apply the Verifiability Approach in real life. In the present experiment we counted verifiable sources rather than verifiable details.

Hypothesis

In this pre-registered experiment (pre-registration: https://osf.io/8er96/?view_only=509da9433071416490e8461a12e7c995) we introduced seven dependent variables: Three measuring the quantity of information (number of words, number of details and number of arguments) and four measuring the quality of these details (plausibility, immediacy, clarity and verifiable sources). There were a few deviations from Leal et al. (2010) regarding these dependent variables. We did not measure *latency time* and *emotional involvement*. These are nonverbal variables which we could not measure because we employed coding on the transcripts of the interviews. We replaced the subjective *talkative* variable with the more objective *number of details* and *number of words* variables. The variables *clarity* and *verifiable sources* were new.

The experiment tested one hypothesis which was pre-registered: Truth tellers will provide more words, more total details, more arguments defending activist actions, more

verifiable sources, and more plausible, immediate and clear statements than lie tellers. (The number of words variable is mentioned in the OSF but, due to an error, not included in the OSF hypothesis).

Method

Participants

We ran a G*Power analysis to determine the required sample size for a multivariate *F*-test (MANOVA global effects). Assuming an alpha level of .05, a statistical power of 0.85, and a medium to large effect size (f^2 = .13) based on effect sizes found in previous deception studies investigating the Devil's Advocate Approach (Deeb et al., 2018; Leal et al., 2010), we needed at least 118 participants for the experiment. We recruited 123 adult participants but the data for two participants were omitted, because they did not follow instructions correctly. The final sample included 121 participants of whom 34 (28.1%) were male and 87 (71.9%) were female. Their mean age was 27.35 years (*SD* = 10.17). Most participants (65.3%) were from the United Kingdom, 13.2% from Eastern Europe, 9.9% from Western Europe and 11.6% from other countries. Participants were recruited via the departmental database and university staff and student portals; 57% of participants were students at the university, 7.4% were staff, and 35.6% were not from the university (though some were ex-students). Participants received £10 for taking part and were entered into a prize draw for £50 Amazon vouchers. Ethics approval was granted by the University faculty's ethics committee and the sponsor's ethics committee. All the followed procedures conformed with the principles of the Declaration of Helsinki.

Design

We carried out three MANOVAs with Veracity as the only factor and with number of words, number of details, number of pro- and anti-action arguments and 7-point Likert scale ratings for plausibility, immediacy, and clarity as dependent variables. All these variables

were coded from the interview transcripts. In the first MANOVA we analysed the responses to the expressing attitude question (Q1). A direct test of the Devil's Advocate Approach is to compare the answers to the eliciting opinion question (Q2) and the devil's advocate question (Q3). Thus, in the second MANOVA we analysed the difference in answering the eliciting opinion and devil's advocate questions. We made such a comparison by subtracting the devil's advocate mean score from the eliciting opinion mean score (e.g., number of words in the eliciting opinion question minus number of words in the devil's advocate question). We label these mean scores 'residue means'. The Devil's Advocate Approach predicts higher residue means for truth tellers than for lie tellers. Computing the pro-action arguments residue mean perhaps merits a further explanation. It was computed as follows: (eliciting opinion question pro-action arguments + devil's advocate question pro-action arguments) – (eliciting opinion question anti-action arguments + devil's advocate question anti-action arguments). In the third MANOVA we analysed the responses to the script writer question (Q4).

Since the analyses for the original devil's advocate question (Q3) and the alternative devil's advocate question (script writer question, Q4) showed different Veracity effects, we also examined the differences between these two questions in a mixed-subjects MANOVA with Question (Q3 vs Q4) as a within-subjects factor, Veracity (truth vs lie) as a between-subjects factor and the number of words, number of details, number of pro-action arguments, number of anti-action arguments, plausibility, immediacy and clarity as dependent variables.

We finally carried out two ANOVAs with Veracity as the only factor and verifiable-witness and verifiable-digital information sources as dependent variables (Q5).

Procedure

All elements of the experiment (questionnaires and interview) took place online. Prior to taking part, participants signed their consent via Qualtrics and a date for the interview to take place was arranged.

Selection Questionnaire

A Zoom invitation was sent out together with a link to the selection questionnaire (to be completed first) so that we could establish what topics each participant felt strongly about. The selection questionnaire first asked for some background information (gender, age, nationality, and occupation [student, staff, other]) followed by agreement/disagreement ratings with ten different statements. These comprised two items for each of five contentious subjects/topics of which activists have recently protested in the United Kingdom (UK): climate change; meat-eating and its impact on climate change; abortion; COVID-19 lockdowns; and the removal of historical monuments with links to racism.

For each item participants were asked to rate their agreement on a 5-point scale ranging from strongly disagree to strongly agree. For each of the five topics there was one positive and one negative statement, to check that participants had read the items correctly and were consistent in their response. For example, the two COVID-19 related items were:
'I think that COVID-19 lockdown restrictions are authoritarian and irresponsible of the UK Government' and 'I fully support all that is being done to prevent the spread of COVID-19 and think the lockdowns and social restrictions that the UK Government have imposed have been necessary'. Hence if a participant feels strongly against the recent lockdowns they would respond with 'strongly agree' to the first statement above and 'strongly disagree' to the second statement above.

Once the participant had completed the selection questionnaire the experimenter selected a topic that the participant then would focus on for the rest of the experiment. This was achieved by subtracting the negative statement score from the positive statement score

for each of the five topics. Thus, a score of 4 would indicate a participant felt strongly for and a score of -4 that the participant felt strongly against a particular topic. If a participant scored strongly on several topics, the experimenter selected a topic based on the frequencies that each topic has already been discussed by other participants (so as to avoid over-subscribing a particular topic). After selecting a topic, the experimenter checked that the participant was happy to discuss that topic. All participants said that they were. Using absolute scores, the endorsement with a topic was strong for both truth tellers (M = 3.33, SD = 0.90, 95% CI [3.12, 3.55]) and lie tellers (M = 3.48, SD = 0.77, 95% CI [3.26, 3.69]). The difference between truth tellers and lie tellers was not significant, $F(1, 119) = 0.88$, $p = .350$, $d = 0.18$ [-0.18, 0.53].

Pre-Interview Questionnaire

Once the experimenter and participant agreed on a topic to discuss in the interview the participant was sent a link to the pre-interview questionnaire. This investigated what methods/activities the participant condones another person using/doing in order to support or express their opinion about the selected topic. The questionnaire is based on the Activism Orientation Scale (Corning & Myers, 2002). Unbeknown to the participant, the items in the questionnaire distinguish between six levels of activism whereby level 1 indicates the participant condones only low levels of protest (e.g., displaying slogan stickers) and level 6 indicates condoning high levels of protest (e.g., using physical violence). See Table 1 for the six levels and an example of each level. As in the Activism Orientation Scale, the participant was asked to rate their agreement to each item on a 4-point Likert scale (completely disagree/somewhat disagree/somewhat agree/completely agree).

Insert Table 1 about here

Once the participant had completed the pre-interview questionnaire, the experimenter entered the scores into SPSS with values from 0 (completely disagree) to 3 (completely

agree) and calculated the mean score for each of the six levels. The participant was then allocated to the level that obtained the highest score. The average strength level for truth tellers (M = 2.06, SD = 0.47, 95% CI [1.96, 2.17]) and lie tellers (M = 2.09, SD = 0.36, 95% CI [1.99, 2.20]) were very similar, $F(1, 119) = 0.16$, $p = .688$, $d = 0.07$ [-0.29, 0.43]).

Introducing the Vignettes

The participant then met the experimenter at the appointed time via the Zoom invitation. We constructed vignettes for each of the five topics and six levels of action introduced in the experiment. We did so to standardise and direct participants in what to discuss (and steer them towards discussing the actions of others and avoid their discussing their own propensity to protest). See Appendix 1 for a description of the vignettes and levels of action.

The experimenter posted the relevant (topic and level) vignette into the Zoom chat and asked the participant to read it. After the participant had read the vignette, the experimenter would check that the participant agreed with/condoned all the actions of Matt and Jane. If the participant said that they did not agree with any/all of the actions (which occurred for 36 of the 121 participants) the experimenter would ask what they did not agree with and then gave them a lower action level vignette to read (with the exception of one participant who considered the action not high enough and so was given a higher action level to read). Again, the experimenter would check if the participant agreed with/condoned all the actions of Matt and Jane. If the participant did not (which occurred for 12 participants) then the next level would be given until the participant agreed with the actions in the vignette. Table 2 shows to which action level the participants were allocated. Truth tellers and lie tellers showed similar distributions, χ^2 (5, n = 121) = 2.69, p = 0.748, see Table 2.

Insert Table 2 about here

Instructions About the Interview

The experimenter then gave the participant instructions for the interview. They varied depending on whether they were in the truth telling or lie telling condition. Allocation to veracity conditions occurred randomly. Truth tellers were instructed to tell the truth regarding their endorsement of the methods used in the vignette. Lie tellers were instructed to lie about it and convince the interviewer that they did not endorse the methods used. All participants were told that they should, if relevant to their responses, express their truthful opinion about the subject matter (e.g., the environment), although we did not ask questions about the subject matter in the interview. To those in the truth telling condition (n = 60), the experimenter said:

> You will now be interviewed about your opinion about [*chosen topic*] and the information contained within the vignette you just read. Just truthfully answer all questions that the interviewer asks you. If you convince the interviewer you are telling the truth then you will be entered into the draw for £50 in Amazon vouchers. If you do not convince the interviewer you are telling the truth then you will not be entered into the draw, but instead will be asked to write a statement explaining your true view about the actions of the protagonists in the vignette.

To those in the lie telling condition (n = 61) the experimenter said:

> You will now be interviewed about your opinion about [*chosen topic*] and the information contained within the vignette you just read. I want you to answer truthfully regarding your opinion about [*chosen topic*]. However, I want you to lie about your endorsement of the methods used by Matt and Jane who you just read about in the vignette. You need to convince the interviewer that you are strongly opposed to the methods used and consider them to be inappropriate. If you convince the interviewer you are telling the truth then you will be entered into the draw for £50 in Amazon vouchers. If you do not convince the interviewer you are telling the truth then you will

not be entered into the draw, but instead will be asked to write a statement explaining your true view about the actions of the protagonists in the vignette.

The experimenter asked the participant to repeat back what they had to do, to check they understood, and asked if they wanted any time to prepare. Most truth tellers did not need time and lie tellers typically asked for a few minutes. We did not measure the exact time they took because we did not think this to be informative since we could not know whether they spend all their time preparing. When the participant was ready to be interviewed, they were placed into a breakout room where the interviewer was waiting for them.

The Interview

The interviewer introduced themselves and then asked five interview questions. Question 1 (expressing attitude question) invited participants to express their views about the actions taken by Matt and Jane in terms of good or bad. Questions 2 (eliciting opinion question) and 3 (devil's advocate question) represented the Devil's Advocate Approach interview protocol. Question 4 invited participants to put themselves in the role of a writer. They were asked to argue against the view they expressed about the methods taken by Matt and Jane in the interview so far. This means that truth tellers were asked to express a view opposite to their actual view but lie tellers to express their actual view. Question 5 asked about sources that could be consulted to verify the opinion expressed in the interview. The questions varied slightly depending on whether the participant was arguing *for* or *against* the action taken by Matt and Jane, see Appendix 2 for the full interview protocol.

Post-Interview Questionnaire

The interviewer notified the experimenter when the interview had concluded and the experimenter brought the participant back into the main meeting room. Participants were asked to complete a post-interview questionnaire where they rated on a percentage scale the

extent to which they told the truth. They were then debriefed. All participants received £10 (via a bank transfer) and were entered in the £50 Amazon vouchers draw.

Coding

All interviews were recorded. An audio file was retained at the end of the interview for transcribing. All interviews were transcribed and coded for each question. In the Devil's Advocate protocol each question is considered an individual entity; therefore, when repetitions occurred within the same question the information was not coded again, but if they occurred in another question they were coded again. For questions 1 to 4 we coded number of words, number of details, number of pro-action and anti-action arguments for the actions of the protagonists and levels of plausibility, immediacy and clarity. For question 5 we coded the number of verifiable sources. The number of words were counted using the Word count option and then manually counting and subtracting any speech errors (errs, umms and uhhs). This is unlikely to have affected the results because these so-called speech disturbances ($d = 0.04$) or filled pauses ($d = 0.00$) are unrelated to deception (DePaulo et al., 2003). For the remaining variables, two coders were used. The main coder coded all transcripts and a second coder coded 50% of transcripts to ensure reliability of coding. Both coders were blind to the hypothesis and veracity of the participant.

Pro-action arguments were coded according to the participant's response to the opening question. If the participant expressed that they were *for* the actions of Matt and Jane in the opening question, this was coded as a pro-action argument (because being *for* their actions is the stance they are taking for the interview), and accordingly all further arguments *for* the actions of Matt and Jane in the eliciting opinion question were coded as pro-action arguments. If the participant expressed that they were *against* the actions of Matt and Jane in the opening question, this was also coded as a pro-action argument (because being *against* the actions of Matt and Jane is the stance they are taking for the interview) and any further

arguments *against* the actions of Matt and Jane were also coded as pro-action arguments for the opinion eliciting question.

Anti-action arguments were coded when the participant gave an opposing argument to that which they had originally expressed. For a participant who expressed they were *for* the actions of Matt and Jane, any argument against their actions in the opening and eliciting opinion questions was coded as an anti-action argument. For a participant who expressed they were *against* the actions of Matt and Jane an argument in support of their actions was coded as an anti-action argument.

For example, this opening question whereby the participant is arguing for Matt and Jane's actions contained four pro-action arguments (emboldened) and two anti-action arguments (underlined).

> 'Erm I mean I think **their actions *are fine*** erm they're not you know they clearly feel **very *passionately*** about the issue and they're not, from the description I remember they're **not actually doing things that break the *law*** erm or causing trouble as such, they're very active and erm joining *meetings* and joining *protests* and erm they ask *other people* to join in which might *offend* or annoy other people but I mean you know on the *little text* it's hard to sort of say exactly what they're doing but as long as they're asking and *not forcing* erm **I don't have problem with that** and *falling out* with people erm possibly extreme but on the other hand if they erm you know if they feel that passionately about it and feel that they would just be annoyed and angry with a person every time they expressed an *opposite view* erm then it's sort of yeah not you know feeling that they're not people that they want to associate with is okay. So I guess in total I'm positiveish.'

Note that both pro-action and anti-action arguments pertain to the actions of Matt and Jane and not any discussion about the issue about which Matt and Jane were protesting (e.g.,

in the example of protesting about removing monuments with racist associations, only arguments for or against the actions in the vignette were coded for, and not discussion about the issue of the monuments, racism or slavery itself).

Details were coded as the main bits of information within the response to the question, without double counting synonyms or tautologies. As an example, the above excerpt given as an example for pro-action and anti-action arguments contained 11 *details* (italicised).

Questions 1 to 4 were also rated on scales of 1 (not at all) to 7 (extremely) for Plausibility (whether the response sounds reasonable and genuine and if there is enough of an answer to sound convincing); Immediacy (whether the response sounds personal and not distanced); and Clarity (how clear the response is for the reader to understand what the participant is saying by the end of the answer).

For example, this participant argued in favour of the actions Matt and Jane took to protest against meat-eating and its impact on climate change:

"Okay yeah I think their actions are good er obviously they're very passionate about the environment which is completely acceptable erm and er they're going about it in a way that doesn't harm anybody it doesn't cause anyone any bother. Also on their shirts aren't anything like offensive you know I think if they had an image of a dead animal for example you know that would really be pushing it to the limit you know but you know in this case it's pretty harmless and um what they're doing wearing the shirts and having the car stickers this is stuff I've seen practised before and er I've never taken any issue to it um as said it's not causing anyone doing any wrong and it's also not illegal either that's also where a line might be you know whether it's breaking the law and this is far from breaking the law it's permitted under freedom of speech um yeah."

A '7' for plausibility was given because the participant elaborated on their views in a thoughtful manner. Immediacy also received a '7' because it was a personal answer. Clarity received a '6' because the participant put effort into explaining their views and the answer was well understood by the coders.

In the following statement the participant argued against the actions Matt and Jane took to protest against the removal of historical monuments.

"It's probably like recent stuff that's going around the world erm just with all the injustice erm I tend to do everything through religion and I'm trying to get people that are my faith to kind of stop doing these things as well erm because I know it can lead to just sinning so I think it's just more of like a really personal faith thing."

This answer received a '1' for plausibility, a '4' for immediacy and a '1' for clarity. Although there is some personal involvement mentioned, the answer is vague, incoherent and unclear and not well understood by the coders in relation to the question.

Question 5 (verifiability question) was coded for the presence of two criteria: verifiable-witness details (e.g., people who could verify for what the participant is saying) and verifiable-digital information details (e.g., social media related).

For example, the sentence "You could speak to my husband because he will definitely remember this conversation and I've got facebook messages" contained one verifiable witness source (husband) and one verifiable-digital information source (facebook).

Inter-rater reliability between the two coders was measured using the two-way random effects model measuring consistency. It was very good for details (Single Measures ICC = .94), pro-action arguments (Average Measures ICC = .93), anti-action arguments, (Average Measures ICC = .97), plausibility (Average Measures ICC = .90), immediacy (Average Measures ICC = .87), clarity (Average Measures ICC = .90), verifiable-witness

source (Average Measures ICC = 1.00) and verifiable-digital information source (Average Measures ICC = 1.00).

Results

Manipulation Check

An ANOVA with Veracity (truth vs lie) as only factor and the percentage of reporting truth telling as dependent variable revealed a significant effect, $F(1, 119) = 270.82$, $p < .001$, $d = 2.99$, 95% CI [2.43, 3.46]. Truth tellers ($M = 98.17\%$, $SD = 4.69$, 95% CI [92.79, 103.54]) reported to have told the truth more than lie tellers ($M = 35.25\%$, $SD = 29.25$, 95% CI [29.92, 40.58]).

Hypothesis Testing

We tested our data using null hypothesis significance testing (NHST) and equivalence testing to support any null findings demonstrating an absence of differences between truth tellers and lie tellers (see Lakens et al., 2018). We chose our smallest effect size of interest to be 0.5, because our research is applied and we were interested in observing a moderate to large effect size. Thus, the equivalence bounds ranged between -0.5 and 0.5.

A MANOVA analysing the expressing attitude question (Q1) with Veracity as the only factor and the seven variables listed in Table 3 as dependent variables revealed a significant multivariate Veracity effect, $F(7, 113) = 7.58$, $p < .001$, $\eta_p^2 = .32$. Table 3 shows that three univariate effects were significant. Truth tellers' statements were more plausible, immediate and clear than lie tellers' statements. This supports the hypothesis for the quality variables but not for the quantity variables although equivalence testing showed no conclusions can be deduced from the data about the number of words and pro-action arguments.

Insert Table 3 about here

The MANOVA for the residue means (means for the eliciting opinion question minus means for the devil's advocate question) showed a significant multivariate Veracity effect, $F(6, 114) = 3.28$, $p = .005$, $\eta_p^2 = .25$. Table 4 shows that three univariate effects were significant. The residue means were significantly higher for truth tellers than for lie tellers for plausibility and immediacy but significantly lower for truth tellers than for lie tellers for the number of pro-action arguments. This partially supports the hypothesis for the quality variables but does not support the hypothesis for the quantity variables.

We further examined whether the residue means differed from zero for truth tellers and lie tellers. One sample *t*-tests showed that for truth tellers the residue scores were significantly higher than zero in five out of six variables (the exception was 'pro-action arguments'). For lie tellers, the residue scores were significantly higher than zero for number of words, number of details and pro-action arguments.

Insert Table 4 about here

For information, Appendix 3 shows the univariate Veracity results for the eliciting opinion question (Q2) and devil's advocate question (Q3) separately. For the eliciting opinion question, compared to lie tellers, truth tellers reported fewer pro-action arguments and their statements were more plausible, immediate and clear. For the devil's advocate question, compared to lie tellers, truth tellers used more words and gave more plausible answers.

A MANOVA analysing the script writer question (Q4) with Veracity as the only factor and the seven variables listed in Appendix 4 as dependent variables revealed a non-significant multivariate Veracity effect, $F(7, 113) = 1.39$, $p = .215$, $\eta_p^2 = .08$. None of the univariate effects were significant either. Equivalence testing supported the NHST results, see Appendix 4.

Since the analyses for the original devil's advocate question (Q3) and the alternative devil's advocate question (script writer question, Q4) showed different Veracity effects, we

also examined the differences between these two devil's advocate questions in a mixed-subjects MANOVA with Question (Q3 vs Q4) as a within-subjects factor, Veracity (truth vs lie) as a between-subjects factor and the number of words, number of details, number of pro-action arguments, number of anti-action arguments, plausibility, immediacy and clarity as dependent variables. The analysis revealed a multivariate Question effect, $F(7, 113) = 12.98$, $p < .001$, $\eta_p^2 = .45$ and a non-significant Veracity main effect, $F(7, 113) = 1.87$, $p = .081$, $\eta_p^2 = .10$. At a univariate level the Veracity effects for number of words, $F(1,119) = 30.70$, $p < .001$, $d = 0.40$, 95% CI (0.14,0.65), number of details, $F(1,119) = 7.66$, $p = .007$, $d = 0.28$, 95% CI (0.02,0.53), number of anti-action arguments, $F(1,119) = 34.20$, $p < .001$, $d = 0.75$, 95% CI (0.47,0.99), plausibility, $F(1,119) = 9.05$, $p = .003$, $d = 0.29$, 95% CI (0.04,0.54), immediacy $F(1,119) = 8.78$, $p = .004$, $d = 0.29$, 95% CI (0.04,0.54), and clarity, $F(1,119) = 6.65$, $p = .011$, $d = 0.25$, 95% CI (0.00,0.50), were all significant. Only the effect for pro-action arguments was not significant, $F(1,119) = 0.73$, $p = .394$, $d = 0.09$, 95% CI (-0.16,0.35). Answers to the alternative devils' advocate question (Q4) included more words and details than answers to the original devil's advocate question Q3). However, answers to Q4 also included fewer anti-arguments and sounded less plausible, less immediate and less clear than answers to Q3. The multivariate Question X Veracity interaction effect was not significant, $F(7, 113) = 1.95$, $p = .068$, $\eta_p^2 = .11$. Perhaps more relevant are the univariate effects for number of words and plausibility, because these two variables revealed significant Veracity effects in the original but not in the alternative devil's advocate question. The univariate interaction effect was not significant for the number of words, $F(1, 119) = 3.27$, $p = .073$, $\eta_p^2 = .03$, but was significant for plausibility, $F(1, 119) = 5.52$, $p = .020$, $\eta_p^2 = .04$. The significant plausibility effect implies that the Veracity effect was more pronounced in the original devil's advocate question (Appendix 3) than in the alternative devil's advocate question (Appendix 4).

Two ANOVAs were carried out for the verifiable sources (Q5) question with verifiable-witness and verifiable-digital information sources as dependent variables. The analysis for verifiable-digital information sources was significant, $F(1, 119) = 4.78$, $p = .031$, $d = 0.40$, 95% CI [0.03, 0.75). This result was supported by equivalence testing, $t(115) = -0.57$, $p = .287$. Truth tellers reported more verifiable-digital information sources ($M = 0.58$, $SD = 0.74$, 95% CI [0.41, 0.76) than lie tellers ($M = 0.31$, $SD = 0.62$, 95% CI [0.14, 0.49). The analysis for verifiable-witness sources was not significant, $F(1, 119) = 0.59$, $p = .443$, $d = 0.14$, 95% CI [-0.22, 0.49). This partially supports the hypothesis.

Discussion

In this experiment we examined differences in verbal responses between truth tellers and lie tellers when they discussed their opinions about protester actions. We examined the quantity of the answers (number of words, details and arguments) and quality of the answers (plausibility, immediacy and clarity). We examined such differences in three different ways. First, when participants were asked to express their attitude towards the actions of the protagonists shown in a vignette (Q1). Second, when comparing their answers to an eliciting opinion question where they were asked to provide the reasons that support their attitude (Q2) with their answers to a Devil's Advocate question where they were asked to give reasons that were against their attitude (Q3). This second method is a direct test of the Devil's Advocate Approach. Third, when participants were asked to take on the role of a script writer (Q4).

When asked to express their attitude towards the actions of the protagonists (first examination, Q1), truth tellers sounded more plausible, immediate and clear than lie tellers, but no differences emerged in the quantity of information. The Devil's Advocate Approach (second examination comparing answers to Qs 2 and 3) predicts the difference in answers (residue scores) to be more pronounced in truth tellers than in lie tellers in terms of both the quantity and quality of the answers. The hypothesis was supported but only in terms of

quality: Truth tellers' answers sounded more plausible and immediate than lie tellers' answers. For the quantity variables only one difference emerged and it was opposite to the prediction: Lie tellers reported more arguments in favour of the protagonists' actions than truth tellers.

In other words, the two examinations produced similar results and both showed that the expected pattern of results was only supported for the quality of details. A similar pattern of results, truth tellers and lie tellers differ from each other in the predicted way in the quality of their answers but not in the quantity of their answers, was found in another experiment examining differences between truth tellers and lie tellers when they expressed their opinions (Vrij et al., 2022). Leal et al. (2010) found similar quality differences (plausibility and immediacy) as obtained in the current experiment but also found a quantitative difference with truth tellers reporting more details than lie tellers.

The findings regarding opinions show similarities and differences from results obtained when interviewees discuss their alleged activities. The similarity is that when discussing alleged activities, the qualitative variable 'plausibility' seems to result in more pronounced differences between truth tellers and lie tellers than the quantitative variable 'total details' (DePaulo et al., 2003; Sporer et al., 2021; Vrij, Deeb et al., 2021), although some caution is warranted given the lack of studies examining plausibility in deception research about alleged activities.

The difference is that in an alleged activity setting the quantitative variable 'total details' is typically a diagnostic veracity indicator with truth tellers reporting more details than lie tellers, see Amado et al. (2015); Gancedo et al. (2021) and Verschuere et al. (2021) for meta-analyses. Although this was also found in Leal et al. (2010), it was not found in the current experiment or in Vrij et al. (2022). Differences in details when reporting alleged activities emerge because lie tellers are less able and willing to report as many details as truth

tellers. They are less able because they find it difficult to fabricate as many details as truth tellers typically report (Köhnken, 2004) and are less willing out of fear that the information they provide would give leads to investigators that they are lying (Masip & Ces, 2011; Nahari, 2019). Perhaps lie tellers find it easier to fabricate information when they talk about opinions than when they discuss activities, for example because they have heard such opinions from someone else or have read about them in the written media. It could also be that lie tellers talk more freely about fabricated opinions than about fabricated activities because they are less concerned that the information they give incriminates them. Regarding the latter, activities are usually embedded in space and time, because activities take place somewhere at a specific moment. Contextual embeddings can easily incriminate lie tellers and they therefore typically report fewer contextual embeddings than truth tellers (Nisin et al., 2022) which contributes to lie tellers reporting fewer details than truth tellers. Contextual embeddings do not play a role when expressing opinions.

Lie tellers are typically more concerned about their credibility than truth tellers (Kassin, 2005; Kassin et al., 2010) and people often believe that providing a lot of information may sound sincere (Bell & Loftus, 1989). If, indeed, lie tellers find it relatively easy to fabricate information and that they are not concerned about this information incriminating them, lie tellers' desire to provide information could explain why they reported more pro-action arguments than truth tellers.

Comparing the results to the expressing their attitude towards the actions of the protagonists question (Q1) with the Devil's Advocate residue scores (Qs 2-3 comparison) showed that the results for the expressing attitude question were more supportive of the hypothesis than the Devil's Advocate results: It resulted in three significant differences rather than two and the effect sizes for the three quality variables were larger in the expressing attitude results (*d's* ranging from 0.50 to 1.15) than for the Devil's Advocate results (*d's*

ranging from 0.35 to 0.48). Yet, we do not think that the two Devil's Advocate protocol questions are redundant. The benefit is that it entails a comparison within the interviewee (comparing Qs 2 and 3, a within-subjects comparison). In contrast, for the expressing attitude question the comparison is made between interviewees (between-subjects comparison). Practitioners, who may need to judge the veracity of a single interviewee in front of them, favour within-subjects comparisons over between-subjects comparisons (Nahari et al., 2019; Vrij, 2016). Rather than only asking the expressing opinion question, we recommend asking that question as well as the two Devil's Advocate protocol questions, because the combination of these three questions provides more verbal veracity indicators than the individual questions and, hence, the most pronounced differences between truth tellers and lie tellers.

The alternative devil's advocate question, in which participants were asked to take on the role of a script writer (Q4), revealed no significant veracity effects. Since the question in its original format (Q3) resulted in two significant veracity effects (for number of words and plausibility) we ran an additional analysis in which we compared the two types of devil's advocate questions (Q3 and Q4) using a design that included Question and Veracity as factors. This analysis revealed that the plausibility effect was more pronounced in answers to the original devil's advocate question than in answers to the alternative version, showing that the original question is somewhat more effective in eliciting veracity differences than the alternative question. Second, Question effects emerged. Compared to the original devil's advocate question (Q3), the alternative devil's advocate question (Q4) resulted in higher quantity answers (more words and details), but lower quality answers (less plausible, immediate and clear). The lower quality answers give the impression that participants (lie tellers as well as truth tellers) found the alternative devil's advocate question more difficult to answer than the original devil's advocate question. This could be caused by participants

thinking that longer answers were required to the alternative question than to the original question.

The verifiable sources question (Q5) resulted in a significant effect for digital information sources only. This shows support for the Verifiability Approach. Previous research has shown that the Verifiability approach is a diagnostic veracity assessment tool in some settings (e.g. criminal settings and airport settings) but not in other settings (e.g. malingering). The present experiment thus revealed another setting in which the Verifiability Approach can distinguish between truth tellers and lie tellers: Opinion settings. However, the frequency of occurrence of digital sources was very low ($M = 0.58$ for truth tellers and $M = 0.33$ for lie tellers). This raises doubts about the usefulness of this cue as a veracity indicator, because a cue is often not helpful for practitioners if it is hardly ever present. We do not have a certain explanation as to why the predicted effect did not occur for witness sources. Participants may have considered digital sources to be more checkable by the interviewer than witness sources. That is, they may have considered it to be more likely that the next question the interviewer asked might have been to substantiate their claims about their social media content than to substantiate their claims about their family members and friends. In sum, lie tellers may feel happier about giving false source information if they feel it to be unlikely that it will be followed up by the interviewer.

Three limitations merit discussion. First, we asked participants to discuss their views on actions taken by others rather than by themselves. We did this for ethical and practical reasons. If participants were to report being willing to take strong illegal actions themselves, they could put themselves into legal difficulty which would raise ethical concerns. In addition, if participants would discuss their own actions, they may have been less honest and more prone to social desirability than when they refer to other people's actions. Also, we expected most participants not to be activists themselves but would be OK with other people

taking certain actions without feeling strongly enough about it to take such risks themselves. For example, many people would have agreed that Hitler should be killed, but rather than doing that themselves, would condone others to do this. We expect talking about other people's actions to be revealing about talking about their own actions because we think the two would positively correlate with each other. However, this remains an empirical question.

Second, the verifiable sources question was always asked at the end of the interview. In future research its position could be manipulated (at the beginning or at the end of the interview) because discussing opinions may influence the answer to the verifiable sources question and vice versa. Third, most participants only condoned low levels of activism and thus discussed low levels of activism in the experiment (the majority of participants discussed levels 1 or 2, see Table 2). It is worthwhile to replicate the experiment with participants who condone high levels of activism (levels 5 or 6, which were condoned by only a small minority of participants) because these are the people practitioners such as immigration, intelligence and vetting officers worry about. It will not be easy to find such a sample because we expect the activism levels to be generally low amongst students and the general population. A specific target group is thus required. We expect the findings to be stronger in such a group of participants than we obtained in the present experiment. We expect them to have more extreme views than students or the general public and holding extreme views should make it more difficult to consider both sides (both pros and cons) of an argument. The less someone can consider both sides, the more effective the Devil's Advocate protocol should become. This is an empirical question worth pursuing.

In conclusion, practitioners such as immigration officers, intelligence officers and vetting officers are often confronted with the question of whether the opinion an interviewee expresses is true or false. The Devil's Advocate approach predicts the difference in answers (residue scores) to be more pronounced in truth tellers than in lie tellers in terms of quantity

of the answers (number of words, details and arguments) and quality of the answers (plausibility, immediacy and clarity). The hypothesis was supported but only in terms of quality: Truth tellers' answers sounded more plausible and immediate and somewhat clearer than lie tellers' answers. Truth tellers also reported more digital verifiable sources than lie tellers. Although more research in this area is needed, these findings make us feel more confident to recommend practitioners to use the Devil's Advocate Approach when attempting to detect lying about opinions.

References

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

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

Bell, B. E., & Loftus, E. F. (1989). Trivial persuasion in the courtroom: The power of (a few) minor details. *Journal of Personality and Social Psychology*, *56*, 669-679. Doi: 10.1037//0022-3514.56.5.669

Bogaard, G., Colwell, K., Crans, S. (2019). Using the Reality Interview improves the accuracy of the Criteria- Based Content Analysis and Reality Monitoring. *Applied Cognitive Psychology*, *33*(6), 1018–1031. https://doi.org/10.1002/acp.3537

Clemens, F., Granhag, P. A., & Strömwall, L. A. (2013). Counter-interrogation strategies when anticipating questions on intentions. *Journal of Investigative Psychology and Offender Profiling*, *10*(1), 125–138. https://doi.org/10.1002/jip.1387

Colwell, K. James-Kangal, N., Hiscock-Anisman, C., & Phelan, V. (2015). Should police use ACID? Training and credibility assessment using transcripts versus recordings. *Journal of Forensic Psychology Practice*, *15*(3), 226-247. https://doi.org/10.1080/15228932.2015.1035187

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., 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

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 University Press.

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

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

Granhag, P. A., & Hartwig, M. (2015). The Strategic Use of Evidence (SUE) technique: A conceptual overview. In P. A. Granhag, A. Vrij, & B. Verschuere (Eds.), *Deception detection: Current challenges and new approaches* (pp. 231-251). Wiley.

Hartwig, M., Granhag, P. A., & Luke, T. (2014). Strategic use of evidence during investigative interviews: The state of the science. In Raskin, D.C., Honts, C.R., & Kircher, J.C. (Eds.), *Credibility assessment: Scientific research and applications (pp. 1-36)*. Academic Press.

Harvey, A., Vrij, A., Leal, S., Lafferty, M., & Nahari, G. (2017). Insurance based lie detection: Enhancing the Verifiability Approach with a model statement component. *Acta Psychologica*, *174*, 1-8. https://doi.org/10.1016/j.actpsy.2017.01.001

Jones, M., & Sugden, R. (2001). Positive confirmation bias in the acquisition of information. *Theory and Decision*, *50*, 59-99. https://doi.org/10.1023/A:1005296023424

Jupe, L. M., Vrij, A., Leal, S., & Nahari, G. (2020). Fading lies: applying the verifiability approach after a period of delay. *Psychology, Crime & Law*, *26*(4), 327-342. https://doi.org/10.1080/1068316X.2019.1669594

Kassin, S. M. (2005). On the psychology of confessions: Does innocence put innocents at risk? *American Psychologist*, *60*, 215-228. Doi: 10.1037/0003-066X.60.3.215

Kassin, S. M., Appleby, S. C., & Torkildson-Perillo, J. (2010). Interviewing suspects: Practice, science, and future directions. *Legal and Criminological Psychology* (Special issue "What works in investigative psychology"), *15*, 39-56. Doi: 10.1348/135532509X449361

Köhnken, G. (2004). Statement Validity Analysis and the 'detection of the truth'. In P. A. Granhag & L. A. Strömwall (Eds.), *Deception detection in forensic contexts* (pp. 41-63). Cambridge University Press.

Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. *Advances in Methods and Practices in Psychological Science*, *1*(2), 259-269. https://doi.org/10.1177/2515245918770963

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., Vernham, Z., Dalton, G., Jupe, L., Harvey, A., & Nahari, G. (2018). Cross-cultural verbal deception. *Legal and Criminological Psychology*, *23*(2), 192-213. https://doi.org/10.1111/lcrp.12131

Masip, J., & Ces, C. (2011, March 3-5). *Guilty and innocent suspects' self-reported strategies during an imagined police interview* [Paper presentation]. The 4th International Congress on Psychology and Law, Miami, FL, United States.

Nahari, G. (2019). Verifiability approach: Applications in different judgmental settings. In T. Docan-Morgan (Ed.), *The Palgrave handbook of deceptive communication* (pp. 213-225). Palgrave Macmillan.

Nahari, G., Ashkenazi, T., Fisher, R. P., Granhag, P. A., Hershkovitz, I., Masip, J., Meijer, E., Nisin, Z., Sarid, N., Taylor, P. J., Verschuere, B., & Vrij, A. (2019). Language of Lies: Urgent issues and prospects in verbal lie detection research. *Legal and Criminological Psychology*, *24*(1), 1-23. https://doi.org/10.1111/lcrp.12148

Nahari, G., Vrij, A., & Fisher, R. P. (2014). The verifiability approach: Countermeasures facilitate its ability to discriminate between truths and lies. *Applied Cognitive Psychology, 28(1),* 122-128. https://doi.org/10.1002/acp.2974

Nisin, Z., Nahari, G., & Goldsmith, M. (2022). Lies divorced from context: evidence for Context Embedded Perception (CEP) as a feasible measure for deception detection. *Psychology, Crime & Law*, Doi: 10.1080/1068316X.2022.2078825

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

Sporer, S., Manzanero, A. L., & Masip, J. (2021). Optimizing CBCA and RM research: Recommendations for analyzing and reporting data on content cues to deception. *Psychology, Crime, & Law, 27*(1), 1-39. https://doi.org/10.1080/1068316X.2020.1757097

Verschuere, B., Bogaard, G., & Meijer, E. H. (2021). Discriminating deceptive from truthful statements using the verifiability approach: A meta-analysis. *Applied Cognitive Psychology, 35*(2), 374-384. https://doi.org/10.1002/acp.3775

Vrij, A. (2016). Baselining as a lie detection method. *Applied Cognitive Psychology*, *30*(6), 1112-1119. https://doi.org/10.1002/acp.3288

Vrij, A., Deeb, H., & Leal, S., & Fisher, R. P. (2022). The effects of a secondary task on true and false opinion statements. *International Journal of Psychology & Behavior Analysis*, *7*: 185. https://doi.org/10.15344/2455-3867/2022/185

Vrij, A., Deeb, H., Leal, S., Granhag, P. A., & Fisher, R. P. (2021). Plausibility: A verbal cue to veracity worth examining? *The European Journal of Psychology Applied to Legal Context, 13*(2), 47-53. https://doi.org/10.5093/ejpalc2021a4

Vrij, A., Fisher, R., Blank, H. (2017). A cognitive approach to lie detection: A meta-analysis. *Legal and Criminological Psychology, 22*(1)*,* 1-21. https://doi.org/10.1111/lcrp.12088

Vrij, A., Leal, S., Deeb, H., Chan, S., Khader, M., Chai, W., & Chin, J. (2019). Lying about flying: The efficacy of the Information Protocol and Model Statement for detecting deceit. *Applied Cognitive Psychology, 34*(1)*,* 241-255. https://doi.org/10.1002/acp.3614

Vrij, A., Leal, S., Mann, S., Vernham, Z., & Brankaert, F. (2015). Translating theory into practice: Evaluating a cognitive lie detection training workshop. *Journal of Applied Research in Memory and Cognition, 4*(2)*,* 110-120. https://doi.org/10.1016/j.jarmac.2015.02.002

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

Table 1

The Six Levels of Activism and an Example of Each Level

Order Level	Example
1. Public Protest/passive support for cause	Display a poster or car sticker with a political message
2. Civil Disruption/slightly active (lawful) support for cause	Attend an informational meeting of a relevant political group
3. Social Disruption/ Persuade or try to convince others of view and to support cause	Invite a friend to attend a meeting of a relevant political organisation or event
4. Civil Disobedience	Use tactics designed to overwhelm police resources to make their point
5. Criminal Damage	Graffiti an opinion on public property to make their point
6. More extreme criminal actions/ risk personal safety	Use physical violence to make their point

Table 2

Level of Action Allocation as a Function of Veracity

Level	Truth tellers	Lie tellers
1	21 (35.0%)	20 (32.8%)
2	16 (26.7%)	15 (24.6%)
3	17 (28.3%)	22 (36.1%)
4	3 (2.5%)	3 (4.9%)
5	1 (0.8%)	1 (0.8%)
6	2 (1.7%)	0 (0%)

Table 3

Statistical Results for the Expressing Attitude Question (Q1)

	Truth			Lie			NHST			Equivalence Testing	
	M	*(SD)*	95% CI	*M*	*(SD)*	95% CI	*F*	*p*	*d*	*t*	*p*
Expressing attitude question (Q1)											
Number of words	143.35	(128.61)	110.82,175.88	171.67	(125.88)	139.41,203.93	01.50	.223	0.22 (-0.14,0.58)	1.53	.065
Number of details	11.35	(7.97)	8.97,13.73	12.52	(10.49)	10.16,14.89	00.48	.489	0.13 (-0.23,0.48)	2.06	.021
Pro-action arguments	2.88	(1.37)	2.48,3.29	3.21	(1.76)	2.81,3.61	01.32	.253	0.21 (-0.15,0.56)	1.60	.056
Anti-action arguments	0.30	(0.62)	0.14,0.46	0.33	(0.63)	0.17,0,49	00.06	.806	0.05 (-0.31,0.40)	2.50	.007
Plausibility	4.98	(1.00)	4.73,5.24	3.82	(1.01)	3.57,4.07	40.62	< .001	1.15 (0.75,1.52)	3.62	1.00
Immediacy	4.20	(1.07)	3.95,4.46	3.61	(0.92)	3.35,3.86	10.73	.001	0.59 (0.22,0.95)	0.52	.699
Clarity	4.27	(1.04)	4.00,4.54	3.74	(1.06)	3.47,4.00	07.66	.007	0.50 (0.14,0.86)	0.02	.507

Note. NHST = Null-hypothesis significance testing. Variables in bold showed significant Veracity differences.

Table 4

Statistical Results for the Devil's Advocate Approach Comparison (Residue Means)

	Truth			Lie			NHST			Equivalence Testing	
	M	*(SD)*	95% CI	*M*	*(SD)*	95% CI	*F*	*p*	*d*	*t*	*p*
Eliciting-opinion question minus Devil's advocate question (Residue means)											
Number of words	26.78*	(79.98)	07.90,45.67	34.80*	(67.32)	16.07,53.53	00.36	.552	0.11 (-0.25,0.46)	2.15	.017
Number of details	02.00*	(07.71)	-01.19,05.20	04.20*	15.87	01.-3,07.37	00.93	.337	0.18 (-0.18,0.53)	1.79	.039
Pro-action arguments	00.33	(02.03)	-00.13,00.81	01.16*	(01.65)	00.70,01.63	06.12	.015	0.45 (0.08,0.80)	0.28	.391
Plausibility	00.55*	(00.93)	00.29,00.77	00.11	(00.95)	-0.12,00.35	06.00	.016	0.47 (0.10,0.82)	-0.30	.382
Immediacy	00.47*	(01.14)	00.21,00.73	-00.02	(00.88)	-0.28,00.24	06.78	.010	0.48 (0.11,0.84)	-0.15	.441
Clarity	00.38*	(01.04)	00.13,00.64	00.03	(00.97)	-0.22,00.29	03.68	.057	0.35 (-0.02,0.70)	-0.83	.204

Note. Mean scores with an '*' differ significantly from zero. NHST = Null-hypothesis significance testing. Variables in bold showed significant Veracity differences.

APPENDIX 1

Vignettes used

Note – each vignette comprised a first paragraph outlining the topic (listed A to E below) followed by a paragraph pertaining to the action level. Here we have listed all first paragraphs followed by just the action levels pertaining to A – anti-meat-eating vignette, for illustration.

A – Anti-meat-eating vignette

Matt and Jane, a married couple, passionately believe that the meat industry is a very strong, if not the strongest, contributor to climate change. They both avoid using any animal products, but they want to take action to change the practice of breeding and keeping animals for meat and, in particular, to end intensive farming.

B – Environmentalist vignette

Matt and Jane, a married couple, passionately believe that the human race is destroying the environment and that we must take action now to save the planet from climate change and organism extinctions. They both do as much as they can to live in an environmentally friendly way, but they want to take action to change other people's behaviour also.

C – COVID restrictions vignette

Matt and Jane, a married couple, passionately believe that the Government's imposing of lockdowns in 2020-21 to contain the spread of COVID-19 is a basic infringement of their human rights. They own a bar which is on the brink of collapse from the lack of trade, and go to great efforts to use PPE and take preventative actions to avoid spreading the virus. They believe they should be allowed to remain open and for people to take responsibility for themselves to avoid spreading the virus whilst living without restrictions, thus allowing them to continue trading. They rally for other people to support their view to avoid economic catastrophe and social isolation.

D – Pulling down statues vignettes

FOR protecting monuments/statues

Matt and Jane, a married couple, are passionately patriotic and whilst they believe in racial equality, they are proud of their British heritage and do not agree with anti-racism protestors pulling down monuments of historic figures. They both want to take action to make other people understand their point of view.

AGAINST (believe monuments should be pulled down)

Matt and Jane, a married couple, are passionately anti-racist and ashamed of the British slave-trading history and do not agree with the national display of monuments and statues of such

slave-related historic figures. They both want to take action to make other people understand their point of view.

E – Abortion vignettes

FOR (for Abortion)

Matt and Jane, a married couple, are passionately pro-choice and believe a woman has the right to end a pregnancy. They are living in Malta, where abortion is disallowed, and want to take action to make other people understand their point of view.

AGAINST (are anti-abortion)

Matt and Jane, a married couple, are passionately anti-abortion and believe the practice should be outlawed in the UK. They both want to take action to make other people understand their point of view.

Level 1

To demonstrate their support for the cause, both Matt and Jane have stickers in their cars and windows and at weekends wear tee-shirts which depict four cartoon cows carrying placards with messages such as 'act now' and 'tell the truth'.

Level 2

To demonstrate their support for the cause, both Matt and Jane attend meetings organised by a relevant protest group, and join in peaceful protest marches and stand in the high street in butcher/supermarket doorways peacefully holding placards which depict four cartoon cows carrying placards with messages such as 'act now' and 'tell the truth'.

Level 3

To demonstrate their support for the cause, both Matt and Jane attend meetings organised by a relevant protest group, and regularly ask their friends and family to join them at those meetings. Both often talk about the subject with others, and have been known to fall out with family, friends and colleagues who disagree with them about the issue.

Level 4

To demonstrate their support for the cause, both Matt and Jane are members of a relevant protest group and joined other protestors in occupying the headquarters of a large corned-beef manufacturer, known to import beef from South America, not allowing any staff to enter the building that morning.

Level 5

To demonstrate their support for the cause, both Matt and Jane are members of a relevant protest group, and recently broke into a pig farm and released the animals onto neighbouring fields and graffitied messages saying 'tell the truth' and 'act now' over the farm buildings.

Level 6

To demonstrate their support for the cause, both Matt and Jane are members of a relevant protest group and recently joined other protestors to march in protest about the issue. The march ended at a supermarket where they stood with other protesters in a large group, making entry for shoppers awkward. When police intervened, both Matt and Jane used physical violence to attempt to resist arrest.

Appendix 2

The interview protocol

If the interviewee reported being *for* their actions, the questions were as follows:

Q1- expressing attitude question: Some people are of the view that the actions taken by Matt and Jane is a good thing, whilst others are of the view that it is a bad thing. Could you please start by telling me what your opinion is about their actions in terms of good or bad?

Q2 - eliciting opinion question: What do you think has led to you having that opinion about being in favour of the actions taken by Matt and Jane?

Q3 - devil's advocate question: You must have considered your opinion. If you were to look at it from the point of view of an opposer to the actions taken by Matt and Jane, is there anything you can say against their actions?

Q4 - script writer question: Imagine you are writing a part in a film for a character who strongly holds the view that the actions taken by Matt and Jane are NOT appropriate? What arguments would you develop in the script for that character for them to justify their position?

Q5 - verifiable sources question: Could you provide me with any details that I could check regarding your opinion about the actions taken by Matt and Jane in the vignette? An example of information I could check is that you may have mentioned your opinion to friends or talked about it on social media? Basically, anything at all that you think I could check to verify that this is your true opinion?

If the interviewee reported being *against* the actions taken by Matt and Jane, questions 1 and 5 were the same, and questions 2, 3 and 4 differed and were as follows:

Q2 - eliciting opinion question: What do you think has led to you having that opinion about being against the actions taken by Matt and Jane?

Q3 - devil's advocate question: You must have considered your opinion. If you were to look at it from the point of view of a supporter of the actions taken by Matt and Jane is there anything you can say in favour of their actions?

Q4 - script writer question: Imagine you are writing a part in a film for a character who strongly holds the view that the actions taken by Matt and Jane are completely appropriate. What arguments would you develop in the script for that character for them to justify their position?

APPENDIX 3

Statistical Results for the Eliciting Opinion Question (Q2) and the Devil's Advocate Question (Q3)

	Truth			Lie			NHST			Equivalence Testing	
	M	*(SD)*	95% CI	*M*	*(SD)*	95% CI	*F*	*p*	*d*	*t*	*p*
Eliciting opinion question (Q2)											
Number of words	151.40	(118.17)	126.09,176.71	118.24	(75.53)	93.14,143.34	03.39	.068	0.33 (-0.03,0.69)	-0.91	.183
Number of details	11.98	(09.10)	09.45,14.52	11.74	(10.68)	09.22,14.25	00.02	.892	0.02 (-0.33,0.38)	-2.62	.005
Pro-action arguments	01.23	(01.11)	00.97,01.50	01.92	(00.99)	01.65,02.18	12.85	<.001	0.66 (0.28,1.01)	-0.83	.797
Anti-action arguments	00.15	(00.58)	00.03,00.27	00.10	(00.30)	-00.02,00.22	00.38	.537	0.11 (-0.25,0.46)	-2.13	.018
Plausibility	04.92	(01.12)	04.65,05.19	03.93	(01.00)	03.67,04.20	25.85	<.001	0.93 (0.54,1.29)	2.33	.989
Immediacy	04.08	(01.31)	03.78,04.38	03.57	(01.02)	03.28,03.87	05.72	.018	0.43 (0.07,0.79)	-0.36	.359
Clarity	04.10	(01.05)	03.83,04.37	03.67	(01.08)	03.40,03.94	04.89	.029	0.40 (0.04,0.76)	-0.54	.295
Devil's advocate question (Q3)											
Number of words	124.62	(87.04)	106.75,65.72	83.44	(47.35)	65.72,101.16	10.50	.002	0.59 (0.22,0.94)	0.48	.684
Number of details	09.98	(07.48)	07.33,12.63	07.54	(12.56)	04.92,10.17	01.68	.198	0.24 (-0.13,0.59)	-1.46	.074
Pro-action arguments	00.68	(01.05)	00.45,00.92	00.77	(00.74)	00.54,01.00	00.28	.598	0.10 (-0.26,0.45)	-2.72	.004
Anti-action arguments	01.43	(01.18)	01.14,01.73	01.43	(01.14)	01.13,01.72	00.01	.973	0.00 (-0.36,0.36)	2.22	.014
Plausibility	04.38	(01.12)	04.12,04.65	03.82	(00.92)	03.56,04.08	09.13	.003	0.55 (0.18, 0.90)	0.27	.606
Immediacy	03.62	(01.06)	03.38,03.86	03.59	(00.80)	03.35,03.83	00.02	.877	0.03 (-0.32, 0.39)	-2.59	.005
Clarity	03.72	(01.06)	03.46,03.98	03.64	(00.97)	03.38,03.90	00.18	.676	0.08 (-0.28, 0.43)	-2.33	.011

Note. NHST = Null-hypothesis significance testing.

APPENDIX 4

Statistical Results for the Script Writer Question (Q4)

	Truth			Lie			NHST			Equivalence Testing	
	M	*(SD)*	95% CI	*M*	*(SD)*	95% CI	*F*	*p*	*d*	*t*	*p*
Script writer question (Q4)											
Number of words	147.03	(103.44)	123.48,170.59	127.56	(79.52)	140.20,150.93	1.35	.248	0.21 (-0.15,0.57)	-1.59	.058
Number of details	12.02	(7.12)	10.28,13.76	10.36	(6.49)	8.63,12.09	1.79	.184	0.24 (-0.12,0.60)	-1.41	.080
Pro-action arguments	1.38	(1.41)	1.04,1.73	1.70	(1.31)	1.36,2.05	1.69	.197	0.24 (-0.13,0.59)	1.45	.075
Anti-action arguments	0.23	(0.56)	0.11,0.36	0.15	(0.40)	0.02,0.27	0.93	.336	0.16 (-0.19,0.52)	-1.78	.039
Plausibility	3.85	(1.04)	3.60,4.10	3.75	(0.92)	3.51,4.00	0.29	.593	0.10 (-0.26,0.46)	-2.21	.014
Immediacy	3.40	(0.96)	3.17,3.63	3.26	(0.83)	3.03,3.49	0.71	.401	0.16 (-0.20,0.51)	-1.91	.029
Clarity	3.40	(1.14)	3.13,3.67	3.44	(0.96)	3.18,3.71	0.05	.824	0.04 (-0.32,0.39)	2.53	.006

Note. NHST = Null-hypothesis significance testing.