Lying about flying: the first experiment to detect false intent

Aldert Vrij , Pär A. Granhag , Samantha Mann & Sharon Leal

Lying about flying: the first experiment to detect false intent

Aldert Vrij[a]*, Pär A. Granhag[b], Samantha Mann[a] and Sharon Leal[a]

[a]Department of Psychology, University of Portsmouth, Portsmouth, UK; [b]Department of Psychology, Gothenburg, Sweden

This article reports the results of, to our knowledge, the first ever experiment to investigate lying about intentions. Sixty passengers in an airport departure hall told the truth or lied about their forthcoming trip in an interview that comprised nine questions. The interviews were transcribed and raters coded the amount of detail in, and the plausibility of, the answers. Raters also coded whether the transcripts included contradictions and spontaneous corrections. Liars' answers were less plausible than truth tellers' answers but did not differ in terms of detail. Liars also included more contradictions and fewer spontaneous corrections in their answers. A total of 72% of truth tellers and 74% of liars were detected on the basis of these variables. We discuss the implications of the findings together with the limitations of the experiment and ideas for future research.

Keywords: deception; deception detection; interviewing; interviewing to detect deception; interviewing style

Introduction

This article describes an experiment investigating lying about intentions. We define intentions as a person's mental representations of his/her planned future actions. Intentions (a) relate to someone's own activities, (b) are based on some amount of reasoning and planning, and (c) typically come with a strong commitment to perform the intended actions (Malle, Moses, & Baldwin, 2001). When people are asked to discuss their intentions regarding future actions, their answers could be truthful (true intent) or deceptive (false intent). Being able to discriminate between true and false intents is of paramount importance to investigators and intelligence work, as it addresses the issue of *preventing* criminal acts from occurring, including terrorist attacks. Given the importance of preventing crime it is unfortunate that research regarding lying about intentions is virtually non-existent. A possible explanation for this is that the need for such research has only become apparent to the relevant groups of people (e.g. scholars, investigators, and policy makers) after the 11 September 2001 terrorist attacks.

A Guilty Knowledge polygraph Test (GKT) can be used to detect someone's intentions in situations where investigators (a) have gathered intelligence about a forthcoming criminal act and (b) have a suspect who may have been involved in the planning of this forthcoming crime (Ben-Shakhar & Elaad, 2003). In certain situations, however, there is no intelligence available, such as when security personnel

interview passengers about their forthcoming trip at an airport. In addition, an airport setting is not suitable for polygraph testing as it is a time-consuming procedure that requires examinees to be attached to polygraph equipment. In our experiment we examined whether airport passengers who are truthful about their forthcoming trip give a different pattern of answers compared to those who lie about their forthcoming trip in interviews with security personnel.

We designed an interview protocol comprising nine questions about the passengers' forthcoming trip. We focused on five dependent variables that we expected to distinguish truth tellers from liars: (i) length of the response, (ii) amount of detail, (iii) plausibility, (iv) contradictions, and (v) spontaneous corrections. We expected those variables to reveal differences between truth tellers and liars because they are associated with the theoretical principles we expected to play a role in lying about intentions: cognitive load, episodic future thought, avoidance strategies, and impression management. We expected liars to give shorter and less detailed answers to questions about their forthcoming trip, e.g. 'What is the main purpose of this trip?' (Hypothesis 1), because fabricating answers to such questions may be more difficult than truth telling, resulting in shorter and less detailed answers. This hypothesis is further supported by the emerging line of research on 'episodic future thought' (Schacter, Addis, & Buckner, 2008). In brief, episodic future thought (EFT) represents the ability to mentally pre-experience a one-time personal event that may occur in the future (Schacter & Addis, 2007). For the current context it is of particular relevance to note that deceptive passengers who themselves make up a plan for a future event seem to activate a more concrete (detailed) mental image of the upcoming scenario than do those who more casually adopt a plan that they do not tend to execute (Watanabe, 2005). In addition, liars may avoid providing detailed information because they run the risk of lie detectors checking the veracity of the details they provide, or of contradicting themselves. The more a liar says, the more opportunity he gives to lie detectors to check the veracity of the statement, and the more chance there is of contradicting himself. The notion that liars' statements are less detailed than truth tellers' statements is widely supported in deception research (DePaulo et al., 2003; Vrij, 2004, 2008).

We further expected liars' answers to be less plausible than those of truth tellers (Hypothesis 2). Although we expected liars to have thought about a false purpose for their trip, they may not have thought through all the sorts of details that we asked them to provide about this pretend purpose. This results in on-the-spot thinking that leads to less plausible answers. Deception research also typically finds that liars' stories sound less plausible (DePaulo et al., 2003; Vrij, 2004, 2008).

We also examined whether the passengers contradicted themselves. If liars are asked to provide all sorts of detail that they have not prepared about their activities, their mental representation of these activities will not be strong. Weaker mental representations are often unstable and, as a result, liars may start to contradict themselves (Vrij et al., 2009). In contrast, truth tellers will have thought about the trip in more detail and will therefore have a stronger mental representation of their forthcoming trip than liars, which should therefore be more stable. We thus expect liars to contradict themselves more than truth tellers (Hypothesis 3).

We further examined whether the passengers made spontaneous corrections (corrections made without being prompted by the interviewer). Truth tellers may have reason to correct themselves because the odd mistake may slip in to their

elaborations about forthcoming activities. Liars may also have reason to correct themselves, because during the interview they may change their mind about what is best for them to say in order to make a convincing impression. However, despite having reasons to correct themselves, liars may be less inclined to do so than truth tellers, typically thinking that correcting themselves looks suspicious (Strömwall, Granhag, & Hartwig, 2004; Taylor & Hick, 2007; Vrij, Akehurst, & Knight, 2006). Liars tend to refrain from saying things that evoke suspicion (Köhnken, 1996, 2004). Therefore, we expect truth tellers to make more spontaneous corrections than liars (Hypothesis 4).

Method

Participants

A total of 60 participants took part, of whom 41 (68%) were male and 19 (32%) were female. Their age ranged from 21 to 75 with an average age of $M = 45.16$ years (SD $= 12.8$). Most participants ($N = 37$, 62%) were flying for business reasons. The other three reasons for travelling were visiting family ($N = 12$, 23%), going on holiday ($N = 10$, 17%) and organizing a funeral ($N = 1$, 2%). There were 19 different destinations that participants were flying to.

Procedure

The study took place airside in the departure hall of a relatively small international airport. The airport is mainly frequented by businesspeople on short-haul flights. The experimenter approached passengers and said: 'I am from the Psychology Department of the University of Portsmouth and would like to ask you two questions about your travel plans for a study about memory'. Those participants who agreed to be questioned (the majority[1]) were asked the following two questions:

1. Where are you going to fly to today?
2. How would you describe the main purpose of your trip?

The experimenter wrote down the answers.[2] In cases where the participant said that s/he was going home, the interview was terminated. In the remaining cases the experimenter continued:

> 'Are you willing to participate in a study that involves telling a convincing story about your travel plans with the chance of earning £10? The interview will take five minutes and will be audiotaped. My colleague will ask you a few questions about your forthcoming trip. Some people will be asked to tell the truth whereas others will be asked to lie during these interviews. My colleague, who does not know who is lying or telling the truth, will make a veracity judgement at the end of the interview. When my colleague believes that you are telling the truth, you will get £10, if she thinks that you are lying, you will not get any money. [In fact, all participants received £10, see below.] Do you understand the procedure? Are you willing to participate?'

Those who were willing to participate (the majority of people) were randomly allocated to the truth/lie condition. A total of 29 passengers were requested to tell the truth while answering every single question during the interview, whereas the remaining 31 participants were asked to lie about the purpose of the trip but to tell the truth about the destination they were flying to. The experimenter then asked the

participant whether s/he needed preparation time. If the participant expressed a need to prepare then the experimenter gave them as much time as they wanted by asking them to return when they were ready to be interviewed, at which point the experimenter took the participant to one of the two female interviewers. The interviewer, not knowing the status (truth/lie) of the participant, switched on the digital tape-recorder, introduced herself and conducted the interview. A total of nine questions about the trip were asked including the following three questions: 'How long will you be away for?'; 'How many times have you made this trip before?'; and 'What would you say is the main purpose of your trip?' After the interview, she told the interviewee whether or not she believed him or her and gave the interviewee £10 regardless of this verdict (we did this for ethical reasons) together with a debriefing form.

Coding

The audiotaped interviews were transcribed. The length of response to the questions was measured objectively by counting via the Word count option, for each question, the number of words in the response excluding speech fillers such as 'erm' or 'er'.

The other variables were coded subjectively by one coder who was blind to the actual truth/lie status of the interviews. On the basis of the transcripts, the coder rated (i) how detailed the answers to the nine questions combined were on a Likert scale rating from (1) few details to (7) many details; and (ii) how plausible the answers to the nine questions combined were on a Likert scale ranging from (1) implausible to (7) plausible. For reliability purposes, a second coder, blind to the actual truth/lie status of the interviews, carried out the same subjective ratings independently from the first rater. The inter-rater reliability scores (Pearson correlations) were 0.41 for plausibility and 0.59 for details. These inter-rater correlations cannot be considered high, but inter-rater reliability scores in the 0.4–0.6 range are not unusual for verbal lie detection coding schemes. For example, similar rates are frequently obtained in CBCA research (Vrij, 2005).

The coder also rated whether there were contradictions or spontaneous corrections in the transcripts. A contradiction is the delivery of inconsistent information. For example, one participant initially said that he was going to meet one friend, but later on referred to meeting 'friends'. Spontaneous corrections occur when the participant explicitly acknowledges changing information without having been prompted by the interviewer. For example, one participant discussed the last time she had seen the friend she was about to visit, and corrected herself about when this was ('I have not seen her for six months. Sorry, I have seen her once since but that was only for like 10 minutes'). The second coder also rated the contradictions and spontaneous corrections in the text. The correspondence between the two coders was satisfactorily high (contradictions, Kappa = 0.76; corrections, Kappa = 0.93).

Results

All 29 truth tellers informed the experimenter that they did not need preparation time prior to the interview, whereas eight of the 31 liars did request it. This difference in asking for preparation time between truth tellers and liars is significant, $\chi^2(1, N = 60) = 8.64$, $p < 0.01$. However, the number of participants who wanted to

Table 1. The answers of the passengers as a function of veracity.

	Truth ($N=29$)		Lie ($N=31$)		F/χ^2	Eta^2	d
Number of words	185.16	143.7	176.74	123.9	0.06	0.00	0.06
Few–many details	3.79	1.2	3.35	1.2	1.90	0.03	0.37
Implausible–plausible	4.24	1.7	3.03	1.6	8.13**	0.12	0.73
Contradictions (% yes)	7%		39%		8.47**		
Corrections (% yes)	24%		3%		5.67*		

*$p<0.05$, **$p<0.05$.

prepare themselves for the interview was too small to include 'Preparation' as a factor in the analyses.

Table 1 reveals that the answers to the nine questions from truth tellers and liars were of equal length and included similar amount of detail, which means that Hypothesis 1 could not be supported. However, in alignment with our remaining predictions, truth tellers' stories sounded more plausible than liars' stories (Hypothesis 2), and truth tellers made fewer contradictions (Hypothesis 3) and more spontaneous corrections than liars (Hypothesis 4).

We further examined how many truth tellers and liars could be correctly classified on the basis of the differences that emerged between the two groups. In order to examine this we carried out several discriminant analyses. The results are reported in Table 2.

First, we examined how many truth tellers and liars could be correctly classified on the basis of each individual variable that obtained a significant result. Table 2 shows that on the basis of plausibility ratings alone, a substantial number of truth tellers (72%) and liars (74%) were correctly classified.

We conducted one more discriminant analysis, where we examined the accuracy in classifying truth tellers and liars based on a cluster of the three variables (plausibility, contradictions and corrections). We used the stepwise method. This procedure allows us to assess the contribution of each of the variables entered into the analysis. For these analyses we report the cross-validated results. In cross-validation, each case is classified by the functions derived from all cases other than that case. The final model included only plausibility as variable ($F(1,58)=25.89$, $p<0.01$ Wilks' Lambda $=0.69$). As mentioned above, 72% of truth tellers and 74% of liars could be correctly classified on the basis of this variable.[3]

Table 2. The correct classification of passengers based on plausibility, contradictions and corrections.

	Truth accuracy	Lie accuracy	Truth accuracy: stepwise and cross-validated	Lie accuracy: stepwise and cross-validated
Implausible–plausible	72%	74%		
Contradictions	93%	39%		
Corrections	24%	97%		
Accuracy rates based on (1) plausibility, (2) corrections and (3) contradictions				
			72%	74%

Discussion

In the first ever (to our knowledge) experiment on lying about intentions we asked passengers in an airport departure hall to tell the truth or to lie about their forthcoming trip. We were reasonably successful in our attempts to discriminate between true and false intents. Truth tellers' answers were more plausible than liars' answers and based on the plausibility cue 72% of truth tellers and 74% of liars were correctly classified. Traditional deception research, based on people lying about their opinions or past activities, has also showed that liars' stories tend to be less plausible than truth tellers' stories (DePaulo et al., 2003; Vrij, 2008). Plausibility thus discriminates in the same way between true and false intents as it does in discriminating between telling the truth and lying about past behaviour. In addition, liars made more contradictions and fewer spontaneous corrections, two deception cues that have also emerged as cues to deceit in traditional deception research (DePaulo et al., 2003; Vrij, 2008).

No more correct classifications of truth tellers and liars occurred on the basis of plausibility scores, contradictions and spontaneous corrections together than on the basis of plausibility scores alone. The addition of contradictions and spontaneous corrections did therefore not enhance accuracy, but this is perhaps not surprising. Contradictions and spontaneous corrections were probably, to some extent, already accounted for in the raters' plausibility score, as people typically associate contradictions with deception and spontaneous corrections with truthfulness (Strömwall et al., 2004; Taylor & Hick, 2007; Vrij, 2008; Vrij et al., 2006). Indeed, when we correlated the plausibility scores with the contradictions and spontaneous corrections ratings we found that plausibility was negatively correlated with contradictions, $r(60) = 0.38$, $p < 0.01$ and positively, albeit not significantly, correlated with spontaneous corrections, $r(60) = 0.18$, NS.

Although examining plausibility alone would thus suffice to achieve 70–75% accuracy rates, we prefer to examine plausibility together with contradictions and spontaneous corrections when making veracity judgements, as it will make the veracity tool somewhat more objective. A disadvantage of relying on plausibility is that it is a subjective cue and leaves room for interpretation to the individual lie detector. The fact that the inter-rater reliability score between the two coders' plausibility ratings was somewhat low ($r = 0.41$), demonstrates the subjective nature of plausibility ratings. Contradictions and spontaneous corrections can be coded more objectively as the Kappa results in the present experiment demonstrated. Whether plausibility can (i) be more objectively rated and (ii) still discriminate between truth tellers and liars when coders are asked to use guidelines such as 'contradictions are implausible' and 'spontaneous corrections are plausible' when judging plausibility is worth investigating in future research.

One hypothesis was not supported. We expected liars' answers to contain fewer details than truth tellers' answers but no difference in detail emerged. This null effect is somewhat exceptional, as many traditional deception studies have found that truth tellers provide more detail than liars. For example, Vrij (2008) reviewed 29 deception studies in which details were examined and in 22 of these studies liars included less detail than truth tellers. In some of the seven studies where the amount of detail did not differ between liars and truth tellers, the liars were coached. For example, in Vrij, Mann, Kristen, and Fisher's (2007) experiment liars were requested to pretend to

have participated in an event that the truth tellers had actually participated in. Liars had received detailed information about the event in which the truth tellers were involved. Coaching, obviously, has an effect on the amount of detail liars are likely to provide and may negate differences in detail between truth tellers and liars. However, in the present experiment no coaching took place. Although the lack of difference in detail between liars and truth tellers could be the result of limitations in our experiment, such as asking inadequate questions or coding the amount of detail inadequately, we will not rule out that it is peculiar to lying about intentions. That is, maybe descriptions of truthful intentions are generally not more detailed than descriptions of false intentions. One aspect that makes truth tellers' stories about past activities often more detailed than liars' stories is that there is a wealth of perceptual details truth tellers have experienced during these past activities that they can recall (if they still remember them). However, when discussing their intentions about a forthcoming trip truth tellers have not yet experienced anything, and this may mean that differences in detail between truth tellers and liars are less likely to occur. Whether or not the detail cue is more diagnostic cue to deception when people discuss their past activities rather than their intentions is an empirical question that needs to be addressed in future research.

The absence of an effect for details may have a positive side-effect: it provides a rationale for examining lying about intentions. One could argue that the hypotheses tested in this experiment do not just apply to lying about intentions but also to lying about past events. Why to bother with lying about intentions research if lying about intentions lead to the same deceptive cues as lying about past events? The absence of an effect for details suggests that differential findings do emerge between lying about intentions and past behaviour and that research findings regarding lying about past events cannot be automatically generalized to lying about intentions. It thus makes intentions deception research worthwhile.

The way forward

This is as far as we know only the first experiment to investigate lying about intentions and much more research is required to draw meaningful conclusions about this area of lie detection. Our experiment has advantages and disadvantages. An attractive aspect is the setting in which it took place. An airport is an important setting to carry out false intentions research because detecting false intent in airports could prevent potentially large calamity. A further advantage is the ecological validity of the findings for truth tellers. Our findings are clearly representative for air travellers who discuss their true intentions, and one could argue that it is more representative than any laboratory experiment possibly could offer. However, we cannot take the representativeness of liars for granted, at least not for all liars. One limitation of the experiment is that the positive consequences for being believed and the negative consequences for not being believed were not as high as they would be in real-life settings for criminals or terrorists. However, although high stakes may affect the behaviour of liars, that is irrelevant for this experiment as behaviour was not analysed. Whether stakes would affect the verbal criteria that were examined in our experiment is unknown, though we believe that they are less likely to be affected than nonverbal behaviour.

A limitation that could affect the verbal criteria of our experiment is that liars did not have the opportunity to thoroughly prepare themselves. Not being well prepared may be representative of some liars, because not all liars will prepare themselves for an interview. For example, members of the resistance during the Second World War who were caught by the Germans and subsequently interviewed by German police were often poorly prepared for those interviews (J. Vrij, 2001). All their precaution activities were focused on avoiding capture by the German police, and the equally important aspect of what to say if they were caught had been neglected. In terms of an airport setting it would mean that many liars will have thought about what they need to do in order to avoid raising suspicion or being selected for an interview by security personnel. Fewer liars will have prepared themselves for such an interview.

Another reason why not all liars would have prepared themselves for such an interview is that they cannot prepare if they do not know what will be asked. At least, this is the reason given by many passengers in the present experiment and explains why relatively few accepted the opportunity we gave them to prepare themselves.

Liars in real-life settings who prepare themselves for an interview by security personnel face the difficulty of needing to correctly anticipate which questions will be asked. Only if they succeed in doing this can their preparation be of benefit. The challenge for security personnel therefore is to design interview protocols that include questions that are relevant in the context but unlikely to be anticipated by prepared liars. Those unanticipated questions should be the most diagnostic questions for lie detection purposes (see also Vrij et al., 2009).

Our interview protocol encourages passengers to discuss their forthcoming trip and one could argue that this is what intentions are about. However, when passengers arrive at the airport they have done some preparation and planning for the trip. They, or someone else on their behalf, have bought tickets and may have arranged hotels. They may have conducted some research into their destination, etc. Do people, when they prepare their lies about their forthcoming trips, engage in such activities, and if so, do they do this in the same manner and to the same extent as truth tellers? And can questions be asked that can successfully discriminate between truth tellers and liars in preparation and planning activity? Future research could provide the answers to these questions.

Notes

1. We estimated that the uptake rate of those approached to take part was about 55–60%. Passengers who did not take part gave good reasons to do so: They were busy (had work to do) or had to catch their flight. The high response rate may have been a combination of an official looking experimenter combined with the lack of facilities at that particular airport (hence many passengers seemed pleased for something unusual to do whilst waiting).
2. These questions were asked to establish the ground truth in the experiment. All participants (truth tellers and liars) were asked to reveal their true destination during the actual interview, and truth tellers were also asked to reveal the true purpose of their trip. When we compared the answers during the actual interview with the answers given to the experimenter, we noticed that all interviewees honestly reported their destination, and that all truth tellers honestly reported the true purpose of their trip.
3. The two interviewers who conducted the airside interviews also made veracity judgements at the end of each of their interviews. They performed well at the lie

detection task and correctly classified 79% of truth tellers and 65% of the liars (72% total accuracy rate). The truth accuracy rate, $Z = 2.61$, $p < 0.01$, lie accuracy rate, $Z = 2.30$, $p < 0.01$ and total accuracy rate, $Z = 2.90$, $p < 0.01$, were all significantly above the level of chance. The interviewers demonstrated a truth bias and judged more passengers as truthful ($N = 34$) than as deceptive ($N = 26$), $Z = 2.90$, $p < 0.01$. We do not know for certain what the interviewers based their decisions on. They could have been based on the answers the passengers gave, on the nonverbal behaviour these passengers displayed, or a combination of the two. We have reason to believe that the latter option occurred. At the end of the experiment, one interviewer told us that she based her judgements on a combination of plausibility and eye contact (where too much indicated that the participant was lying).

The difference in the distribution of accuracy rates obtained in the discriminant analysis and by the two interviewers approached significance, χ^2 ($N = 60$) $= 3.46$, $p = 0.06$. Interviewers' truth accuracy rate (79%) was somewhat higher than the truth accuracy rate that emerged from the plausibility scores (72%, see Table 1), but the interviewers' lie accuracy rate (65%) was somewhat lower than the outcome that emerged from paying attention to plausibility scores (74%, see Table 1).

References

Ben-Shakhar, G., & Elaad, E. (2003). The validity of psychophysiological detection of information with the guilty knowledge test: A meta-analytic review. *Journal of Applied Psychology*, 88, 131–151.

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

Köhnken, G. (1996). Social psychology and the law. In G.R. Semin, & K. Fiedler (Eds.), *Applied Social Psychology* (pp. 257–282). London: Sage.

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: Cambridge University Press.

Malle, B.F., Moses, L.J., & Baldwin, D.A. (2001). *Intentions and intentionality: Foundations of social cognition*. Cambridge, MA: The MIT Press.

Schacter, D.L., & Addis, R.D. (2007). The cognitive neuroscience of constructive memory: Remembering the past and imagining the future. *Philosophical Transactions of the Royal Society*, 392, 773–786.

Schacter, D.L., Addis, D.R., & Buckner, R.L. (2008). Episodic simulation of future event: Concepts, data and applications. *Annals of New York Academy of Sciences*, 1124, 39–60.

Strömwall., L.A., Granhag, P.A., & Hartwig, M. (2004). Practitioners' beliefs about deception. In P.A. Granhag, & L.A. Strömwall (Eds.), *Deception detection in forensic contexts* (pp. 229–250). Cambridge: Cambridge University Press.

Taylor, R., & Hick, R.F. (2007). Believed cues to deception: Judgements in self-generated serious and trivial situations. *Legal and Criminological Psychology*, 12, 321–332.

Vrij, A. (2004). Invited article: Why professionals fail to catch liars and how they can improve. *Legal and Criminological Psychology*, 9, 159–181.

Vrij, A. (2005). Criteria-Based Content Analysis: A qualitative review of the first 37 studies. *Psychology, Public Policy, and Law*, 11, 3–41.

Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities* (2nd ed.). Chichester: John Wiley and Sons.

Vrij, A., Akehurst, L., & Knight, S. (2006). Police officers', social workers', teachers' and the general public's beliefs about deception in children, adolescents and adults. *Legal and Criminological Psychology*, 11, 297–312.

Vrij, A., Leal, S., Granhag, P.A., Mann, S., Fisher, R.P., Hillman, J., & Sperry, K. (2009). Outsmarting the liars: The benefit of asking unanticipated questions. *Law and Human Behavior*, 33, 159–166.

Vrij, A., Mann, S., Kristen, S., & Fisher, R. (2007). Cues to deception and ability to detect lies as a function of police interview styles. *Law and Human Behavior*, *31*, 499–518.

Vrij, J. (2001). Verzet tegen angst. *Kontakt door Aantreden*, *56*, 4.

Watanabe, H. (2005). Semantic and episodic predictions of memory for plans. *Japanese Psychological Research*, *47*, 40–45.