

Investigating deception in second language speakers: Interviewee and assessor perspectives

Lucy Akehurst¹, Alina Arnhold¹, Isabel Figueiredo¹, Sarah Turtle¹ and
Amy-May Leach²

¹University of Portsmouth, UK

²University of Ontario, Institute of Technology, Canada

ACCEPTED FOR PUBLICATION IN LEGAL AND
CRIMINOLOGICAL PSYCHOLOGY, FEBRUARY 2018

*Requests for reprints should be addressed to Dr Lucy Akehurst, Department of Psychology, University of Portsmouth, King Henry I Street, Portsmouth, UK (e-mail: lucy.akehurst@port.ac.uk).

Abstract

Purpose. The first of two experiments investigated the effect that speaking in a non-native language has on interviewees' perceptions of their interview experience. A second experiment investigated evaluators' perceptions of the credibility of interviewees who spoke in their native or non-native language.

Method: For the first experiment, 52 participants told the truth or lied about their identity during a mock border control interview. All of the participants were interviewed in English, for half of the sample this was their native language and for the other half of the sample English was not their native tongue. Post interview, all participants completed a self-report questionnaire relating to their perceptions of their interview experience. For the second experiment, 128 participants evaluated the credibility of interviewees from the first experiment. The modality of presentation of interview clips was varied and included 'Visual and Audio', 'Visual Only', 'Audio Only' and 'Transcript Only'.

Results: Non-native speakers were more likely than native speakers to report being nervous and cognitively challenged during their interviews and were more likely to monitor their own behaviour.

Overall, evaluators were better able to distinguish between truth tellers and liars who were speaking in their native language than between truth tellers and liars who were non-native speakers. Relative to native speakers, there was a smaller truth bias for evaluations of non-native speakers. When evaluators were considering the non-native speakers, they achieved higher discrimination accuracy when they were exposed to 'Visual Only' or 'Transcript Only' presentations than when they were shown the 'Visual and Audio' or 'Audio Only' interview clips.

Conclusions: Self-reported experiences of a mock border control interview differed dependent on whether interviewees were speaking in their native or non-native language. Discrimination accuracy was better for native speakers than it was for non-native speakers and was at its worst when evaluators heard the accents of the non-native speakers.

Investigating Deception in Second Language Speakers: Interviewee and Assessor Perspectives

With increasing global threats to security and high levels of migration, border control interviews are commonplace and it is very often the case that interviewees are required to communicate in a language other than their native tongue. The first experiment presented here focusses on the experiences of truthful and lying interviewees who spoke in their native or non-native language at interview. The second experiment, which utilised the data gathered for the first experiment, was designed to investigate the effect that the language status of an interviewee (native or non-native) and the presentation mode of interview clips had on evaluators' judgments of credibility.

Speaking in a second language can result in increased communication apprehension, nervousness, anxiety and cognitive strain (Gregersen, 2005). The verbal and non-verbal behaviours associated with these processes are often linked with deception, and may be the reason that non-native speakers are perceived as less credible than native speakers (Castillo, Tyson & Mallard, 2014; DaSilva & Leach, 2013; Elliott & Leach, 2016; Evans, Pimentel, Michael & Pena, 2017; Leach, Snellings & Gazaille, 2017; Lev-Ari & Keysar, 2010). But how do interviewees, speaking in their non-native language, perceive their own experience? Our first experiment explored the perceptions of interviewees in terms of three processes that have been linked with deceptive behaviour: Emotion (Ekman & Friesen, 1969; Frank & Svetieva, 2013; Frank & Ekman, 1997; Matsumoto, Keltner, Shiota, O'Sullivan & Frank, 2008; Porter & tenBrinke, 2008; Porter & tenBrinke, 2009), cognitive load (Vrij, Fisher, Mann & Leal, 2008; Vrij, Granhag, Mann & Leal, 2011) and behaviour monitoring (DePaulo, et al, 2003; Mann, Vrij & Bull, 2002; Vrij, 2008).

Emotion

Bond and Lai (1986) were amongst the first to explore verbal responses to emotion-inducing and neutral questions provided in their participants' first language (Cantonese) or second language (English). They found that if the question topic was embarrassing participants were more likely to

speak for longer in English than in Cantonese. They suggested that speaking in a second language could serve as a distancing function and reduce emotion.

However, Caldwell-Harris and Ayçiçeği-Dinn (2009) showed that the underlying reasons for Bond and Lai's findings might be more complex. Caldwell-Harris and Ayçiçeği-Dinn (2009) included a measure of skin conductance as an index of anxiety and found that, in line with Bond and Lai's findings, participants in their study exhibited less physiological arousal when listening to emotional phrases in a non-native language compared to their native language. Yet, more arousal was elicited when participants were asked to read statements out loud in their non-native language than their native language. Caldwell-Harris and Ayçiçeği-Dinn (2009) concluded that two separate factors influence the arousal experienced by bilingual speakers when they lie in their two languages: arousal due to emotions associated with lying, and arousal due to anxiety about managing speech production in the non-native language. In sum, the management of second language speech production has been shown to increase anxiety levels and displays of emotion. More recently, Evans, Michael, Meissner and Brandon (2013) investigated the perceptions of evaluators with regard to the nervousness exhibited by lie-telling and truth-telling interviewees who were speaking in their native or non-native language. Although they did not measure the self-reported nervousness of their interviewees, they found no difference between their evaluators' perceptions of nervousness for liars and truth tellers when the interviewees were speaking in their native language. However, when speaking in their non-native language (with high and low proficiency), liars were rated as significantly more nervous than truth tellers. In light of the previous research, it was hypothesised that those interviewees in the current experiment speaking in their non-native language would report being more nervous than those speaking in their native language and that this finding would be accentuated for those interviewees who were lying compared to those who were truth-telling (Hypothesis 1).

Cognitive load

Speaking in a second language may increase cognitive load (Gregersen, 2005). In addition, it places demands on neural processing, causing difficulty in engaging in word and event recall (Perani & Abutalebi, 2005; Ullman, 2001). Research has proposed that the act of deception similarly taxes cognitive resources (Kozel, Padgett & George, 2004; Vrij, Fisher, Mann & Leal, 2008; Vrij, Granhag, Mann & Leal, 2011). Furthermore, Broadbent (1957) suggested that cognitive load increases when attention is divided between two tasks. When one task (e.g., communicating in a non-native language) requires a great deal of effort, the second task suffers as a consequence (e.g., lying).

Cheng and Broadhurst (2005) found that native and non-native liars self-reported higher cognitive load compared to native and non-native truth tellers, but they found no evidence that speaking in a non-native language was perceived by interviewees as more demanding than speaking in their native language. Cheng and Broadhurst (2005) did not, however, report the results of statistical tests on their data. It is therefore not clear what the effect sizes were in terms of the differences in self-reported cognitive load between their experimental conditions. Moreover, some non-native speaking participants in Cheng and Broadhurst's study 'code switched' during their interviews (i.e., they used words from both their first and second language). It has been suggested that code switching helps to lessen cognitive load when lying and telling the truth (Silva-Corvalán, 1994). As participants in Cheng and Broadhurst's study were not prevented from code switching, it is still unclear as to whether non-native speakers (instructed to stick with their non-native language and avoid code-switching) would self-report higher cognitive load compared to native speakers.

Subsequent studies, in which code switching was not permitted, have yielded mixed findings. Duñabeitia and Costa (2015) found that, independently, deceptive statements and second language use resulted in pupil dilations and longer speech durations (both signs of cognitive load; Goldinger & Papesh, 2012; Strijkers, Baus, Runnqvist, Fitzpatrick & Costa, 2013). However, the two effects

did not interact. This could suggest that there is no additional cognitive load associated with speaking a non-native language while deceiving. Yet, Duñabeitia and Costa (2015) asked participants to either tell the truth or lie about which animal they saw pictured on a screen. While this does establish ground truth, it fails to replicate the forensic context of an investigative interview. Additionally, participants were asked to limit their response to a pre-determined sentence structure (I see a [colour] [animal] with [number] legs), a very manufactured lab scenario that may have compromised participants' motivation to lie convincingly (Vrij, 2015).

In their study of the effectiveness of a Psychologically Based Credibility Assessment Tool (PBCAT), Evans et al. (2013) investigated cognitive load by measuring evaluators' perceptions of how hard interviewees had to think during their interviews. They found no difference between their evaluators' perceptions of how hard truth tellers and liars were having to think when the interviewees were speaking in their native language. However, when speaking in their non-native language (with low proficiency), liars were rated as having to think harder than truth tellers.

Based on existing theories of cognitive load and non-native language speaking, and using a forensically relevant scenario, it was hypothesised that non-native language speakers in the current experiment would find their interviews more cognitively demanding than native speakers and that this difference would be accentuated when interviewees were lying as opposed to truth-telling (Hypothesis 2).

Behaviour monitoring

Liars are typically less likely to take their credibility for granted than truth tellers (Kassin & Gudjonsson, 2004; Kassin & Norwick, 2004). As such, liars, more so than truth tellers, are motivated to be perceived as truthful and thus monitor their own behaviour for signs of suspicion (DePaulo et al., 2003; DePaulo & Kirkendol, 1989; Kassin, 2005; Kassin et al., 2010; Kassin & Gudjonsson, 2004; Kassin & Norwick, 2004). In addition, they monitor interviewers' reactions more carefully in order to assess whether they appear to be lying successfully (Buller & Burgoon,

1996; Schweitzer et al., 2002). However, with increased cognitive load, second language speakers may be sufficiently distracted to the point that they are incapable of employing an impression management strategy that involves simultaneously monitoring their own behaviour as well as that of the interviewer. It was hypothesised that native speakers, more so than non-native speakers, would report monitoring their own behaviour and that of the interviewer and that these differences would be accentuated when interviewees were lying rather than telling the truth (Hypothesis 3).

Detecting deceit

So far we have discussed the possible emotions, cognitive load and impression management dilemma faced by truth-telling and fabricating interviewees speaking in their own, or a foreign, language. But what of the credibility assessment task? More often than not, the only evidence available to border control officers is a person's account. To decide whether this person is telling the truth or trying to deceive is one of the most important yet difficult tasks faced by the authorities. Accuracy rates are typically around 50% (i.e., chance level, see Bond & DePaulo, 2006; Vrij, 2008).

As Bond and DePaulo (2008) reported in their meta-analysis, individual differences in deception detection accuracy are minimal. However, the inclination to regard statements as truthful varies (Bond & DePaulo, 2008). Most individuals tend to trust others which results in truth-biased credibility judgments (Bond & DePaulo, 2008; DePaulo et al., 2003; Vrij, 2008). This process may be explained by the availability heuristic (O'Sullivan, Ekman & Friesen, 1988; Tversky & Kahneman, 1973). According to this theory, people are simply more exposed to truthful behaviour in their daily lives and therefore conclude that deceptive behaviour is rare, even in experimental settings. When a greater base rate of truths is expected, people are more likely to judge others as truthful (Street & Richardson, 2015). Another explanation for the high frequency of truth ratings relates to social norms. As it would be deemed impolite to constantly question the truthfulness of others, conversation rules prevent individuals being suspicious of one another and questioning

everything that is said (Vrij, 2008; Vrij & Baxter, 1999), and mean that they are mindful of the social costs of signalling distrust (ten Brinke, Vohs, & Carney, 2016).

On the other hand, the majority of researchers have found that evaluators exhibit, at best, less of a truth bias and, at worst, a lie bias toward non-native speakers (e.g., Castillo, Tyson & Mallard, 2014; DaSilva & Leach, 2013; Elliott & Leach, 2016; Evans et al., 2017; Evans & Michael, 2014; Leach et al., 2017; Levi-Ari & Keysar, 2010). A large body of research spanning the last fifty years shows that foreign-accented speakers tend to be evaluated more negatively on various traits, including those related to credibility (see Dragojevic, 2016; Garrett, 2010; Giles & Rakić, 2014 and Giles & Watson, 2014). For example, a meta-analysis by Fuertes, Gottdiener, Martin, Gilbert, and Giles (2012) showed that accents negatively influenced perceptions of intelligence, education, attractiveness and trustworthiness. Bond and Atoum (2000) stated that listeners attributed blame when they could not understand foreign accents. Recent research has reported a clear bias of perceiving non-native speakers as less truthful than native speakers (Dragojevic & Giles, 2016; Hansen & Dovidio, 2016; Lev-Ari and Keysar, 2010). It could be that if a speaker has an accent, statements are judged to be deceptive simply because they are harder to understand. In fact, research has shown that perceptual fluency can increase belief in a message (Unkelbach, 2007).

Lay people and professionals intent on detecting deceit tend to hold incorrect beliefs about deceptive behaviour. For example, liars are frequently expected to avoid eye contact and fidget (Akehurst, Köhnken, Vrij & Bull, 1996; DePaulo et al., 2003; Global Deception Research Team, 2006; Strömwall & Granhag, 2003; Vrij, Akehurst & Knight, 2006). Whilst these behaviours have not been found to be strongly diagnostic across studies (see DePaulo et al., 2003; Bond & DePaulo, 2008; Vrij, 2008 for meta-analyses) it is of note that they are amongst the behaviours exhibited by non-native speakers, especially when they are feeling anxious (Gregerson, 2005). This may therefore lead to the erroneous labelling of a non-native speaker as a liar.

Non-native speakers also exhibit verbal behaviours indicative of lying. They tend to use

simple and concrete words, which are easier to access, and avoid abstract terms (Newman, Pennebaker, Berry & Richards, 2003). Non-native speech is also less diverse than native speech with significantly higher use of redundant and repetitive words (Kormos & Dénes, 2004). Exactly the same can be found in deceptive speech (Arciuli, Mallard, & Villar, 2010; Vrij, 2008; Zhou, Burgoon, Nunamaker, & Twitchell, 2004).

Taken together it is clear that the similarity between cues exhibited by liars and cues exhibited by non-native speakers may well lead to a lie bias in judgments of credibility of non-native speakers. It was predicted that in our second experiment non-native speakers would be labelled liars more often than they would be labelled truth tellers thus resulting in a lie bias (Hypothesis 4). It was further hypothesised that evaluators would be better able to discriminate between truths and lies for native speakers compared to non-native speakers (Hypothesis 5).

In order to tease apart the influence of accent, nonverbal cues and verbal cues on discrimination and bias, truthful and deceptive interviews with native and non-native persons were presented to participants in one of four ways. If participants were allocated to the Visual and Audio condition they watched videos of interviews which contained both audio and visual information (accent, nonverbal cues and verbal cues present). If they were allocated to the Visual Only condition they watched interviews without sound (nonverbal cues present). If participants were allocated to the Audio Only condition they heard interviews but with no picture (accent and verbal cues present) and if they were allocated to the Transcript Only condition they read typed verbatim transcripts of interviews (verbal cues present).

Generally speaking, a reliance on visual cues (e.g. nonverbal behaviours) is reported to decrease lie detection ability (Vrij, Granhag, & Porter, 2010). DePaulo et al. (2003) and Bond and DePaulo (2006) showed that audio recordings facilitated the detection of deception, whereas participants in mute video conditions used (inaccurate) visual cues and exhibited decreased accuracy rates. However, considering the possibility that accent and linguistic cues play a significant

role in triggering a lie bias in judgments of the credibility of non-native speakers, for this group of interviewees, audio information may 'muddy the waters' when it comes to accurate assessments. As such, no firm predictions were made in terms of the effect of presentation modality on discrimination accuracy and bias.

Some previous research investigating lying by native and non-native speakers has failed to establish ground truth (Bond & Atoum, 2000; Evans et al., 2013; Evans et al., 2017; Evans & Michael, 2014) and/or has not used a forensically relevant setting (e.g., Cheng & Broadhurst, 2005; Evans et al., 2017). In the current experiment, interviewees lied or told the truth about their identity during a mock border control interview. Ground truth was established by checking the content of interviewees' responses with their passports (for truth tellers) or with details of a fake identity that was given to liars. Furthermore, this investigation did not limit itself to only one type of non-native speaker (e.g., only Chinese or only Hispanic interviewees speaking English) but instead recruited from a diverse population of people who did not have English as their first language. This resulted in a sample that better reflected the many different cultures and ethnicities that pass through border control each day.

Experiment 1

Method

Participants. A total of 52 people (28 females, 24 males) participated in the study, ranging from 18 to 47 years of age ($M = 24.87$ years, $SD = 5.07$ years). The sample was predominantly white (62% White European, 12% Black African, 8% Asian Pakistani, 7% Asian Chinese, 3% Black Caribbean and 8% of participants did not state their ethnicity). The native speakers were undergraduate students at an English-speaking university and received a course credit in return for their participation. All of the native English speakers were British and did not have a 'foreign' accent (e.g. there was no-one in this group from USA, Canada, Australia etc.). The non-native speakers were students at an International school who were in England to improve their English

proficiency; they received a certificate of attendance. In an attempt to control language proficiency, all native English speakers self-reported that English was their first language and they had all been educated in English-speaking schools. Non-native English language speakers self-reported speaking English as a second language and attended non-English speaking schools until the age of 18 years. All participants were asked to rate their English verbal proficiency on a 7-point Likert scale (1 = not at all proficient to 7 = extremely proficient). There was no overlap in the ratings made by the native English language speakers ($M = 6.15$, $SD = .73$) and those made by the non-native English language speakers ($M = 4.08$, $SD = .69$)¹. That is, all of the native speakers rated themselves 6 or above on the scale and all non-native speakers rated themselves 5 or below on the 7-point scale. Furthermore, we asked all non-native speakers when they started to learn English and all responded that they had not received formal education in English until they were 11 years or older².

Participants were recruited in groups of four. Within each group, two were native English speakers (one was assigned to be a truth teller and one to be a liar) and two spoke English as a second language (one truth teller and one liar). All of the participants in a group were matched for age (within 3 years), skin colour (White, Black, Asian Pakistani/Indian or Asian Oriental) and gender. The distribution of gender, age and skin colour was therefore comparable across the native and non-native speaking groups and across veracity conditions.

Design.

The experiment comprised a 2 (Veracity of Interviewee: truth teller or liar) x 2 (Language Status of Interviewee: native or non-native speaker) between-subjects design. The dependent

¹ There was a significant difference between the self-reported proficiency ratings of the two groups, $t(50) = 10.54$, $p < .001$.

² The non-native speakers in this study can be considered of medium to high English language proficiency. Evans et al. (2017) classed those who had started to learn English at the ages of 6-12 years as the second highest English proficiency group in their study. Furthermore, the non-native speakers in this study were, at the time of the experiment, learning English to help them with future university study in the UK.

variables were responses to a post-interview questionnaire regarding participants' perceptions of their interview experience.

Procedure. All participants were asked to bring their passports to their interviews to ensure that the ground truth (i.e., their identity) could be established. At the beginning of each experimental session, whilst participants read a consent form, a researcher recorded details from all participants' passports. The details that were recorded related to the questions that were asked at interview. Therefore, for the truth tellers, the researchers were able to check that the details given at interview matched those in their passports. Each liar was provided with details of a false identity. By providing a false identity for the participants in the lie condition, the researchers could ensure that those participants were indeed lying at interview and not simply embedding some lies in otherwise truthful answers regarding their identity.

We drew up paperwork depicting false identities of people of different ages, both genders and different countries of birth, and provided accompanying photographs. Depending on the skin colour, age and gender of the lying participant, an appropriate false identity was provided (i.e., we assigned an identity that matched the characteristics of the participant). As all of the truth telling native speakers were British, all of the lying native speakers were given a false identity that was for a British person with the same age, gender and skin colour as their truth telling counterpart. The lying non-native speakers were given a false identity that was for a person of their own nationality with the same age, gender and skin colour.

Truth tellers. In the truth condition the researcher read the following instructions to participants;

"You have had to leave your country of origin in a hurry and have now arrived in another country. You are at border control. The officer is suspicious regarding your identity and will now interview you. Please respond to the officer's questions truthfully".

Liars. Participants in the lie condition were presented with a new identity and the following instructions were then read aloud by the researcher;

"You have had to leave your country of origin in a hurry and were given a false identity to enable you to do this. The details of your false identity are outlined on this sheet. You have now arrived in another country and are at border control. The officer is suspicious regarding your identity and will now interview you. Please respond to the officer's questions *not* as yourself but as the person whose identity you have been given".

Participants in both conditions were free to ask questions and were given as much time as they wished³ to memorise their passport information (truth tellers) or their false identity information (liars) as they were not allowed to take their documentation in to their interviews.

All participants were then interviewed by the same interviewer and their interviews were video recorded. The interviewer was blind to the hypotheses and to the allocation of participants to the veracity conditions. Although this was a laboratory based study, every effort was made by the interviewer to assume the character of a border agency officer and to conduct a typical border control interview. The interviewer received guidance from the fifth author who had worked at Canadian Customs. The questions asked can be found in Appendix 1. The interviewer spoke with a Standard English accent.

Once the interview was over, participants were asked to complete a questionnaire⁴. Nine questions relating to the interview experience were rated on 7-point Likert scales ranging from 1 (not at all) to 7 (very). Emotion was measured with one item; participants rated how nervous they felt when answering the questions. Cognitive load was measured with five items (Cronbach's alpha = .77) including measurement of how difficult it was for participants to remember their passport

³ Though it was clear to participants that they could take as long as they wished to prepare for their interviews, none took longer than 5 minutes.

⁴ The questionnaires were in English. For the non-native speakers, it was made clear by the Experimenter that should they have any problems understanding any of the questions they should seek clarification and explanation. None did so.

information, how difficult it was for them to attend to the interviewer's questions, how difficult it was for them to understand the interviewer's questions, how hard they had to think about their answers and how difficult it was to explain their answers clearly. Finally behaviour monitoring was measured with three items (Cronbach's alpha = .62) including measurement of how motivated participants were to convince the interviewer that they were being truthful, how much they monitored their own behaviour and how much they monitored the interviewer's behaviour.

Results

A MANOVA was conducted with Veracity of Interviewee (truth teller or liar) and Language Status of Interviewee (native or non-native speaker) as the independent variables and responses to the nine scale items as the dependent variables. At a multivariate level, there was a significant main effect for Veracity of Interviewee, $F(10, 39) = 2.14, p = .045, \eta p^2 = .35$, a significant main effect for Language Status of Interviewee, $F(10, 39) = 4.26, p < .001, \eta p^2 = .52$ and a significant Veracity X Language Status of Interviewee interaction, $F(10, 39) = 2.74, p = .012, \eta p^2 = .41$.

We examined each of these effects more closely at a univariate level. Table 1 gives the means and standard deviations for each condition. Table 2 gives the F values and effect sizes for the univariate level main effects and interaction effects. The univariate main effect for Veracity showed that liars ($M = 4.69, SD = 1.32$) self-reported that they were more motivated to be perceived as truthful than were truth tellers ($M = 3.85, SD = 1.29$). Liars ($M = 4.50, SD = 1.14$) also reported monitoring the interviewer's behaviour more than truth tellers ($M = 3.85, SD = .97$).

There were four significant main effects of Language Status of Interviewee. Non-native speakers ($M = 3.88, SD = 1.21$) self-reported that they were more nervous than native speakers ($M = 2.77, SD = 1.14$). Non-native speakers ($M = 2.73, SD = .92$) reported that they found it more difficult to understand the interviewer's questions than native speakers ($M = 1.96, SD = 1.04$) and non-native speakers ($M = 3.96, SD = 1.25$) found it more difficult to explain their answers than

native speakers ($M = 3.08$, $SD = 1.29$). Finally, native speakers ($M = 4.54$, $SD = 1.36$) reported monitoring their own behaviour to a greater extent than non-native speakers ($M = 3.42$, $SD = 1.36$).

At a univariate level, the Veracity X Language Status of Interviewee interaction was only significant for the item about perceived difficulty in explaining answers. A simple main effects analysis demonstrated that lying non-native speakers ($M = 4.62$, $SD = .96$) had significantly more difficulty explaining their answers than truth-telling non-native speakers ($M = 3.31$, $SD = 1.18$). However, there was no significant difference for difficulty in explaining answers between lying native speakers ($M = 2.77$, $SD = 1.54$) and truth-telling native speakers ($M = 3.38$, $SD = .96$).

Discussion

Experiment 1 aimed to examine the self-reported experiences of native and non-native speakers who lied or told the truth during mock border control interviews. First, in partial support of Hypothesis 1, we found that non-native speakers reported being more nervous than native speakers. This replicates the previous work of Caldwell-Harris and Ayçiçeği-Dinn (2009). However, we failed to find support for the second part of Hypothesis 1 that the effect would be exacerbated for the interviewees who were lying. There are several reasons for this finding. Non-native speakers might have been at ceiling in terms of nervousness; thus, any additional factors, such as lying, would have little noticeable effect. However, this seems unlikely as scores were around the midpoint of the scale. Alternatively, as posited by Caldwell-Harris and Ayçiçeği-Dinn (2009), perhaps there are several pathways to nervousness. That is, non-native speakers may have been nervous about communicating in their second language and being understood but indifferent about lying, especially as the stakes were low for this experiment. Because we did not ask participants to expand upon their responses (i.e., indicate why they were nervous) or provide a clear operational definition of nervousness for the participants, we cannot be certain as to why the predicted interaction was not observed. These limitations could be addressed in future research as could increasing the motivation of participants to better mimic that of real world border control interviews.

Second, we found partial support for Hypothesis 2 that non-native speakers would find their interviews more cognitively demanding than native speakers, particularly if they lied. Non-native speakers reported more difficulty in understanding the interview questions than native speakers. That this effect was not accentuated during deception was not surprising as the experience of listening to, and understanding a question, should not differ for truth tellers and liars.

Non-native speakers found it harder to explain their answers when they were lying compared to when they were telling the truth, this was not the case for native speakers. As predicted by Broadbent (1957), engaging in two cognitively taxing activities (i.e., lying and speaking in a non-native language) impaired perceived performance on one of the tasks. There is a significant body of research that indicates that cognitive overload affects interviewees' abilities to deceive (e.g., Vrij et al., 2008). Our findings indicate that interviewees' subjective experiences, in terms of effortful cognitive processing, are affected by other sources of load (i.e., speaking in a non-native language).

That said, no significant differences were found between the self-reports of the native and non-native interviewees nor between the self-reports of the liars and truth tellers, for questions relating to how difficult it was for them to remember their passport information and to pay attention to the interviewer's questions. Similarly, these groups did not differ in their ratings of how hard they had to think about their answers. This suggests that cognitive load is multifaceted and not all components will be equally affected by language proficiency (or any factor, for that matter). It could, for example, be argued that the current task was difficult for truth tellers as well as for liars as the former group had to memorise the details of their passports in order to recall accurately at interview. The five 'cognitive load' questions asked in the current study allowed for a more nuanced view of cognitive load – rather than the all-or-none approach the field has been adopting. In this case, memory for detail, attention to the interviewer's questions and processing answers before verbalising seemed impervious to the 'cognitive load' effect. It was only verbalising answers

that triggered the predicted difference in ratings for truth tellers and liars. It should also be noted that the non-native speakers in this experiment had medium to high English language proficiency and anecdotally reported that their listening skills were superior to their verbalising ability. Furthermore the stakes were low for this experiment, nothing like they might be during a real border control interview, and therefore participants were arguably not motivated to lie convincingly.

Third, liars reported monitoring the interviewer's behaviour more so than truth tellers (thus supporting Buller & Burgoon, 1996 and Schweitzer, Brodt & Croson, 2002). We did not, however, find that liars monitored their own behaviour more than truth tellers. Rather, in line with Hypothesis 3, native speakers reported monitoring their own behaviour more so than non-native speakers. With increased cognitive load, and more anxiety, second language speakers may be sufficiently distracted to the point that they are incapable of employing an impression management strategy that involves simultaneously monitoring their own behaviour and paying attention to the other demands of the interview.

In terms of monitoring behaviour, there was no interaction effect between veracity and language proficiency of the interviewees. It is likely, as discussed above in terms of cognitive demands, that self-monitoring of behaviour and monitoring the behaviour of others could be multifaceted. Future research should endeavour to pick apart the motivations of interviewees who state that they monitor their own behaviour, and that of others, during investigative interviews.

Experiment 2

The first experiment focused on the self-reported experience of non-native and native speakers. However, simply because there are underlying differences between groups, does not mean that these are correctly interpreted by observers. The data from Experiment 1 was used in Experiment 2 to focus on whether there are differences between how native and non-native speakers are perceived; specifically in terms of deception detection.

Method

Participants. One hundred and twenty-eight participants were recruited for this experiment (81 females, 47 males), ranging from 18 years to 65 years of age ($M = 32.88$ years, $SD = 13.13$ years). The sample was predominantly white (86% White European, 3% Black African, 5% Asian Chinese and 6% did not state their ethnicity). All recruits were students who received a course credit in return for their participation and self-reported being native English language speakers whose entire education had been in English and who had lived in the UK all of their lives.

Design. This experiment comprised a 2 (Veracity of Interviewees: truth tellers or liars) x 2 (Language Status of Interviewees: native or non-native speakers) x 4 (Modality of Presentation of Interview: visual and audio, visual only, audio only and transcript only) mixed design. Veracity of Interviewee was a within subjects factor and Language Status of Interviewee and Modality of Presentation of Interview were between subjects factors. The dependent variable was a dichotomous rating made by each participant concerning the truthfulness of each of 12 interviewees ('truth' or 'lie').

Materials. Of the 52 interviews conducted during Experiment 1, four were not used in this experiment as the quality of the footage was poor (this constituted one group of participants). Therefore, 48 interviews were used: 24 native speakers (12 liars and 12 truth tellers) and 24 non-native speakers (12 liars and 12 truth tellers). Four sets of each interview were created to provide material for each of the Presentation Modalities. That is, there were Visual and Audio, Visual Only, Audio Only and Transcript Only versions of each of the interviews. Interviews ranged from 70 seconds to 205 seconds ($M = 151.73$ seconds, $SD = 32.17$ seconds). An unrelated t-test revealed no significant difference in the length of the interviews for native ($M = 141.88$ seconds) and non-native ($M = 161.58$ seconds) speakers, $t(46) = -2.30, p = .28$.

Procedure. Participants were misled into thinking that they would be shown 20 interview clips and that they may see any combination of truths and lies. These instructions were included to reduce the expectation of participants that they would be shown an equal number of truths and lies

which may have influenced the decisions they made. All participants were actually presented with 12 interviews (6 lies and 6 truths). They either watched native speakers or they watched non-native speakers. For each participant, six interview clips were chosen at random from the pool of 12 possible clips (e.g. six from the 12 clips of native truth tellers and six from the 12 clips of native liars *or* six from the 12 clips of non-native truth tellers and six from the 12 clips of non-native liars). When an interview clip had been evaluated it was removed from the random selection process until a participant had seen all 12 of his/her clips. Within a condition, the next participant to be recruited was shown the remaining 12 clips and then the process started again. Participants were assigned to one of the presentation modality conditions; Visual and Audio, Visual Only, Audio Only or Transcript Only.

Participants took part in the study individually and were given time between clips to judge the credibility of each interviewee. They were simply asked to rate, by ticking a box, whether or not they believed each interviewee was being truthful about his/her identity during the mock border control interviews.

Results

Judgment accuracy. Table 3 shows the proportion of correct judgments made by evaluators as a function of the Language Status of the Interviewee and the Modality of Presentation of the interview clips. For native speakers, evaluators were able to judge credibility at around chance level for all modalities of presentation. One-sample *t*-tests demonstrated that the proportion of native speakers correctly evaluated were not significantly above chance level of .50 (*p*-values ranged from .074 to .554). For non-native speakers, chance level accuracy was achieved for the proportion of clips correctly evaluated in the Transcript Only presentation mode (*p* = .198) and Visual Only presentation mode (*p* = .654). The proportion of clips correctly evaluated for the Visual and Audio and Audio Only presentation modes was significantly below chance level (*p*-values = <.001 and .006, respectively).

Signal detection analyses. Meissner and Kassin (2002) have suggested that Signal Detection Theory (SDT) can be used to analyse the accuracy of credibility assessments. As such, the performance of our evaluators was calculated in terms of response bias (β) and discrimination accuracy (d'). Beta (β) is a measure of response bias (i.e., the tendency to respond 'truth' or 'lie'). A value of 1 indicates that participants did not favour a particular response, whereas >1 signifies a bias towards a truth judgment and <1 signifies a bias towards a lie judgment. D prime (d') is a pure measure of sensitivity, and it measures the signal and the noise means in standard deviation units. It corrects for response bias and guessing. A value of 0 indicates evaluators showed an inability to distinguish liars from truth tellers. Values above 0 indicate an ability to distinguish liars from truth tellers, and negative values indicate response confusion.

Response bias. A 2 (Language Status of Interview) x 4 (Modality of Presentation) ANOVA was performed with participants' response bias (β) entered as the dependent variable.

First, there was a significant main effect of Language Status. Evaluators showed more of a truth bias toward native speakers ($M = 1.24$, $SD = .54$) compared to non-native speakers ($M = 1.06$, $SD = .32$). Second, there was a significant main effect of Modality of Presentation. Pairwise comparisons using Bonferroni adjustment found that when evaluators were considering Visual Only interviews ($M = 1.37$, $SD = .58$) they exhibited a significantly larger truth bias than when they were considering Audio Only interviews ($M = .97$, $SD = .19$), $p < .001$, $d = .96$. There were no other significant differences in response bias when comparing the other combinations of presentation modality (all p values $> .05$). Third, there was no significant Language Status X Modality of Presentation interaction effect. See Tables 4 and 5 for M , SD , F and p values and effect sizes.

Using one-sample t -tests with Bonferroni adjustments, each β was compared to 1 (no bias). With regard to Language Status of Interviewee, evaluators who judged native speakers were significantly biased to respond 'truth', whereas no significant response bias was found for evaluations of non-native speakers. In terms of Modality of Presentation, evaluators in the 'Visual

Only' condition displayed a significant truth bias whereas evaluators in the 'Visual and Audio', 'Audio Only' and 'Transcript Only' conditions showed no bias. See Table 6 for M , SD , t , and p values and effect sizes.

Discrimination accuracy. A 2 (Language Status of Interview) x 4 (Modality of Presentation) ANOVA was performed with participants' sensitivity scores (d') entered as the dependent variable to examine their ability to discriminate between liars and truth tellers.

First, there was a significant main effect of Language Status. Evaluators were significantly better at discriminating between the truthful and fabricated accounts of native speakers ($M = .21$, $SD = .58$) compared to non-native speakers ($M = -.20$, $SD = .73$). Second, there was a significant main effect of Modality of Presentation. Pairwise comparisons using Bonferroni adjustment found that when evaluators read transcripts of interviews they were significantly better able to discriminate liars from truth tellers ($M = .19$, $SD = .64$) than when they listened to audio only interviews ($M = -.18$, $SD = .73$), $p = .019$, $d = .69$. There were no other significant differences in ability to discriminate when comparing the other combinations of presentation modality (all p values $> .05$). Third, there was a significant Language Status X Modality of Presentation interaction effect. We performed t-tests to test the effect of Language Status of the Interviewee within each Modality of Presentation condition. When evaluators were judging the Visual and Audio presentations, they were significantly better able to discriminate between truth tellers and liars who were native speakers ($M = .26$, $SD = .53$) than non-native speakers ($M = -.49$, $SD = .59$), $t(30) = 3.75$, $p = .001$, $d = 1.32$. Similarly, when evaluators were judging the Audio Only presentations they were significantly better able to discriminate between truth-telling and lie-telling native speakers ($M = .21$, $SD = .52$) than non-native speakers ($M = -.56$, $SD = .73$), $t(30) = 3.42$, $p = .002$, $d = 1.21$. There were no effects of Language Status when evaluators were rating clips based on Visual Only or Transcript Only presentations (all p values $> .05$). See Tables 4 and 5 for M , SD , F and p values and effect sizes.

Using one-sample t -tests with Bonferroni adjustments, d' values were compared to 0 (no ability to differentiate between truths and lies). With regard to Language Status of Interviewees, evaluators could discriminate truth tellers from liars who were native speakers at a level significantly above 0 however this was not the case for discrimination between truthful and fabricating non-native speakers. In terms of Modality of Presentation of the clips, taking native and non-native speakers together, evaluators could not reliably discriminate between liars and truth tellers for any of the presentation modalities. See Table 6 for M , SD , t , and p values and effect sizes.

Discussion

Evaluators did not exhibit a lie bias when judging the credibility of non-native speakers therefore no support was found for Hypothesis 4. That said, they showed a truth bias for judgments of native speakers that was not present for judgments of non-native speakers who were viewed more negatively. Furthermore, evaluators were significantly better at discriminating truths from lies told by native speakers compared to non-native speakers thus supporting Hypothesis 5.

More of a truth bias for native speakers, compared to non-native speakers, seems an omnipresent and prevalent finding in deception detection studies. Dixon, Mahoney and Cocks (2002), Frumkin (2007) and Lev-Ari and Keysar (2010) all found that participants who spoke with a foreign accent were rated as significantly less credible and more deceptive than participants without a foreign accent. The findings of Experiment 2 also replicated those of Elliott and Leach (2016), Evans et al. (2017), Leach et al. (2017) and Leach and Da Silva (2013) who found more of a truth bias for native speakers than for non-native speakers. The only, seemingly, contradictory finding comes from Bond and Atoum (2000) who found a truth bias towards non-natives. However, in that study the non-natives spoke in their native language. Thus, they were not non-native speakers struggling to communicate in a second language.

Although previous research has elicited higher lie detection accuracy rates for audio (vs. video) presentation of native speakers (Bond & DePaulo, 2006; Kassin, Meissner & Norwick, 2005), this finding was not replicated in the present study. According to Bond and DePaulo (2006, p. 225) “the usual stereotype of a liar is largely visual, hence is most strongly evoked by video images of people speaking”. However, the present study provides evidence that stereotypes regarding accent may also mislead evaluators. When evaluators were judging the Visual and Audio presentations and the Audio Only presentations (i.e., when accent was available) they were significantly better able to discriminate between the truths and lies of native speakers compared to non-native speakers. There were no effects of Language Status on discrimination accuracy when evaluators were rating clips based on Visual Only or Transcript Only presentations (i.e., when the accent of the interviewees was not available).

Future Research

Some researchers have found, as have we, that when it comes to judging differences between truth tellers and liars, accuracy is better for native speakers than for non-native speakers (e.g. DaSilva & Leach, 2013; Leach & DaSilva, 2013), others have found no differences in discrimination accuracy across the groups (e.g. Evans & Michael, 2014, Castillo et al., 2014) and recently Evans et al. (2017) found the highest discrimination accuracy for their speakers with lowest English proficiency. A detailed review of the differing methodologies including varied proficiencies for non-native speakers, varied interview topics/lengths/styles and varied non-native speaker group composition might reveal possible reasons for the mixed findings.

The use of interpreters during interviews with non-native speakers would eliminate the biasing impact of accent which we found, in Experiment 2, to be detrimental to lie detection. Recently four studies have examined the effects of using interpreters on cues to deceit. Ewens, Vrij, Leal, Mann, Jo and Fisher (2016) found that interviewees who spoke through an interpreter provided less detail than interviewees who spoke without an interpreter and that cues to deceit

occurred more frequently when interviewees spoke without an interpreter than with an interpreter. However, Ewens et al. (2017) found that the amount of detail given by interviewees differentiated truth tellers from liars across their interpreter-absent and interpreter-present conditions. In an attempt to encourage interviewees with an interpreter to say more, Ewens et al. (2016) looked at the effectiveness of a model statement. The model statement resulted in native speakers and those interviewed with an interpreter providing more detail than the non-native speakers interviewed without an interpreter. However, no difference was found in the amount of extra detail provided by liars and truth tellers across the interpreter conditions. Ewens, Vrij, Mann, and Leal (2016) found that the reverse order recall technique revealed two cues to deceit when an interpreter was present, whereas no cues to deceit emerged when interviewees spoke in a non-native language. More research needs to be conducted in this area, for example, research is yet to explore how interpreters affect lie detection and bias specifically.

Finally, studies have found that different cultures have different behavioural norms (Matsumoto, Yoo & Fontaine, 2008) and that culture defines coding, encoding and interpretation of behaviours (Ekman & Friesen, 1971). The majority of past research in this area has used homogeneous groups of non-native speakers (e.g., only Hispanics; Evans & Michael, 2014). Similar to DaSilva and Leach (2013), we chose to recruit non-native speaking and native-speaking interviewees from a diverse population. During the analysis of data for our experiments, no distinction was made in terms of sub-groups within the non-native and native speaker groups based on ethnicity, culture or beliefs. Different cultures have different norms in terms of verbal and nonverbal communication, these may impact upon observers' impressions. Further, it is likely that our interviewer's behaviour was perceived by our interviewees in different ways depending upon their culture and beliefs. This has repercussions for detecting deception as individuals from different cultures, experiencing interviews in different ways, may behave differently (see Taylor, Lerner, Conchie, & van der Zee, 2015 for a review of cultural issues in this field). Future research might

place emphasis on investigating systematic differences between specific groups of non-native speakers in terms of their experience of interviews, their verbal and nonverbal behaviours and the accuracy of credibility judgments as a function of, for example, skin colour and ethnicity.

References

- Akehurst, L., Köhnken, G., Vrij, A., & Bull, R. (1996). Lay persons' and police officers' beliefs regarding deceptive behaviour. *Applied Cognitive Psychology, 10*(6), 461-471. doi: 10.1002/(SICI)1099-0720(199612)10
- Arciuli, J., Mallard, D., & Villar, G. (2010). “Um, I can tell you're lying”: Linguistic markers of deception versus truth-telling in speech. *Applied Psycholinguistics, 31*(03), 397-411. doi: 10.1017/S0142716410000044
- Bond, C. F., & Atoum, A. O. (2000). International deception. *Personality and Social Psychology Bulletin, 26*(3), 385-395. doi: 10.1177/0146167200265010
- Bond, C. F., & DePaulo, B. M. (2008). Individual differences in judging deception: Accuracy and bias. *Psychological Bulletin, 134*(4), 477. doi: 10.1037/0033-2909.134.4.477
- Bond, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10*(3), 214-234. doi: 10.1207/s15327957pspr1003_2
- Bond, M. H., & Lai, T. M. (1986). Embarrassment and code-switching into a second language. *Journal of Social Psychology, 126*(2), 179-186.
- Broadbent, D.E. (1957). A mechanical model for human attention and immediate memory. *Psychological Review, 64*(3), 205–215. doi: 10.1037/h0047313
- Buller, D. B., & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory, 6*(3), 203-242. doi: 10.1111/j.1468-2885.1996.tb00127.x
- Caldwell-Harris, C. L., & Ayçiçeği-Dinn, A. (2009). Emotion and lying in a non-native language. *International Journal of Psychophysiology, 71*(3), 193-204. doi: 10.1016/j.ijpsycho.2008.09.006

- Castillo, P.A., Tyson, G., & Mallard., D. (2014). An investigation of accuracy and bias in cross-cultural lie detection. *Applied Psychology in Criminal Justice*, 10(1), 66-82.
- Cheng, K. H., & Broadhurst, R. G. (2005). Detection of deception: the effects of language on detection ability among Hong Kong Chinese. *Psychiatry, Psychology and Law*, 12(1), 107-118. doi: 10.1375/pplt.2005.12.1.107
- Da Silva, C. S., & Leach, A.M. (2013). Detecting deception in second-language speakers. *Legal and Criminological Psychology*, 18(1), 115–127. doi: 10.1111/j.2044-8333.2011.02030.x
- DePaulo, B.M., Lindsay, J.J., Malone, B.E., Muhlenbruck, L., Charlton, K. & Cooper, H. (2003). Cues to deception. *Psychological Bulletin*, 129(1), 74 – 118. doi: 10.1037/0033-2909.129.1.74
- DePaulo, B. M., & Kirkendol, S. E. (1989). The motivational impairment effect in the communication of deception. In J. C. Yuille (Ed.), *Credibility assessment* (pp. 51–70). Dordrecht, The Netherlands: Kluwer.
- Dixon, J. A., Mahoney, B., & Cocks, R. (2002). Accents of guilt: Effects of regional accent, race, and crime type on attributions of guilt. *Journal of Language and Social Psychology*, 21(2), 162-168. doi: 10.1177/02627X02021002004
- Dragojevic, M. (2016). Language attitudes as intergroup terrain. In H.Giles & A. Maass (Eds.). *New horizons in the study of intergroup communication* (pp. 51-66). New York: Peter Lang.
- Dragojevic, M. & Giles, H. (2016). I don't like you because you're hard to understand: The role of process fluency in the language attitudes process. *Human Communication Research*, 42, 396-420. doi: 10.1111/hcre.12079
- Duñabeitia, J. A., & Costa, A. (2015). Lying in a native and foreign language. *Psychonomic Bulletin and Review*, 22(4), 1124-1129. doi: 10.3758/s13423-014-0781-4

- Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, 17(2), 124. doi:10.1037/h0030377
- Ekman, P., & Friesen, W. V. (1969). Nonverbal leakage and clues to deception. *Psychiatry: Journal for the Study of Interpersonal Processes*, 32(1), 88–106.
- Elliott, E., & Leach, A. M. (2016). You must be lying because I don't understand you: Language proficiency and lie detection. *Journal of Experimental Psychology: Applied*, 22(4), 488 – 499. doi: 10.1037/xap0000102
- Evans, J. R., Pimentel, P. S., Pena, M. M., & Michael, S. W. (2017). The ability to detect false statements as a function of the type of statement and the language proficiency of the statement provider. *Psychology, Public Policy, and Law*, 23(3), 290 - 300. doi: 10.1037/law0000127
- Evans, J. R., & Michael, S. W. (2014). Detecting deception in non-native English speakers. *Applied Cognitive Psychology*, 28(2), 226-237. doi: 10.1002/acp.2990
- Evans, J. R., Michael, S. W., Meissner, C. A., & Brandon, S. E. (2013). Validating a new assessment method for deception detection: Introducing a Psychologically Based Credibility Assessment Tool. *Journal of Applied Research in Memory and Cognition*, 2(1), 33-41. doi: 10.1016/j.jarmac.2013.02.002
- Ewens, S., Vrij, A., Mann, S., Leal, S., Jo, E., & Houston, K. (2017). The effect of the presence and seating position of an interpreter on eliciting information and cues to deceit. *Psychology, Crime & Law*, 23(2), 180-200. doi: 10.1080/1068316x.2016.1239100
- Ewens, S., Vrij, A., Leal, S., Mann, S., Jo, E., & Fisher, R. P. (2016). The effect of interpreters on eliciting information, cues to deceit and rapport. *Legal and Criminological Psychology*, 21(2), 286-304. doi: 10.1111/lcrp.12067

- Ewens, S., Vrij, A., Mann, S., & Leal, S. (2016). Using the reverse order technique with non-native speakers or through an interpreter. *Applied Cognitive Psychology, 30*(2), 242-249. doi: 10.1002/acp.3196
- Ewens, S., Vrij, A., Leal, S., Mann, S., Jo, E., Shaboltas, A., ... & Houston, K. (2016). Using the model statement to elicit information and cues to deceit from native speakers, non-native speakers and those talking through an interpreter. *Applied Cognitive Psychology, 30*(6), 854-862. doi: 10.1002/acp.3270
- Frank, M., & Ekman, P. (1997). The ability to detect deceit generalizes across different types of high-stakes lies. *Journal of Personality and Social Psychology, 72*, 1429-1439.
- Frank, M. G., & Svetieva, E. (2013). Deception. In D. Matsumoto, M. Frank, & H. S. Hwang (Eds.), *Nonverbal communication: Science and applications* (pp. 121–144). Newbury Park: Sage.
- Frumkin, L. (2007). Influences of accent and ethnic background on perceptions of eyewitness testimony. *Psychology, Crime and Law, 13*(3), 317-331. doi: 10.1177/02627X02021002004
- Fuertes, J. N., Gottdiener, W. H., Martin, H., Gilbert, T. C., & Giles, H. (2012). A meta-analysis of the effects of speakers' accents on interpersonal evaluations. *European Journal of Social Psychology, 42*(1), 120-133. doi: 10.1002/ejsp.862
- Garrett, P. (2010). *Attitudes to language*. Cambridge, UK: University Press.
- Giles, H. & Rakić, T. (2014). Language attitudes: Social determinants and consequences of language variation. In T. M Holtgraves (Ed.). *The Oxford handbook of language and social psychology* (pp.11-26). Oxford, UK: University Press.
- Giles, H. & Watson, B. (2013). *The social meanings of language, dialect and accent: International perspectives on speech styles*. New York, USA: Peter Lang.

- Global Deception Research Team. (2006). A world of lies. *Journal of Cross-Cultural Psychology*, 37(1), 60-74. doi: 10.1177/0022022105282295
- Goldinger, S. D., & Papesh, M. H. (2012). Pupil dilation reflects the creation and retrieval of memories. *Current Directions in Psychological Science*, 21(2), 90-95. doi: 10.1177/0963721412436811
- Gregersen, T.S. (2005). Nonverbal cues: Clues to the detection of foreign language anxiety. *Foreign Language Annals*, 38(3), 38-400. doi: 10.1111/j.1944-9720.2005.tb02225.x
- Hansen, K. & Dovidio, J.F. (2016). Social dominance orientation non-native accents and hiring recommendations. *Cultural Diversity and Ethnic Minority Psychology*, 22(4), 544 – 551. doi: 10.1037/cdp0000101
- Kassin, S. M. (2005). On the psychology of confessions: Does innocence put innocents at risk? *American Psychologist*, 60(3), 215–228. doi: 10.1037/0003-066X.60.3.215
- Kassin, S. M., Drizin, S. A., Grisso, T., Gudjonsson, G. H., Leo, R. A., & Redlich, A. D. (2010). Police-induced confessions: Risk factors and recommendations. *Law and Human Behavior*, 34(1), 3-38. doi: 10.1007/s10979-009-9188-6
- Kassin, S. M., & Gudjonsson, G. H. (2004). The psychology of confessions: A review of the literature and issues. *Psychological Science in the Public Interest*, 5(2), 33-67. doi: 10.1111/j.1529-1006.2004.00016.x
- Kassin, S. M., Meissner, C. A., & Norwick, R. J. (2005). " I'd know a false confession if I saw one": A comparative study of college students and police investigators. *Law and Human Behavior*, 29(2), 211. doi: 10.1007/s10979-005-2416-9
- Kassin, S. M., & Norwick, R. J. (2004). Why people waive their Miranda rights: The

power of innocence. *Law and Human Behavior*, 28(2), 211-221. doi:

10.1023/B:LAHU.0000022323.74584.f5

Kormos, J., & Dénes, M. (2004). Exploring measures and perceptions of fluency in

the speech of second language learners. *System*, 32(2), 145-164. doi:

10.1016/j.system.2004.01.001

Kozel, F. A., Padgett, T. M., & George, M. S. (2004). A replication study of the

neural correlates of deception. *Behavioral Neuroscience*, 118(4), 852.

doi: 10.1037/0735-7044.118.4.852

Leach, A., & Da Silva, C. S. (2013). Language proficiency and police officers' lie detection

performance. *Journal of Police and Criminal Psychology*, 28, 48–53. doi:10.1007/s11896-

012-9109-3

Leach, A. M., Snellings, R. L., & Gazaille, M. (2017). Observers' language proficiencies and the

detection of non-native speakers' deception. *Applied Cognitive Psychology*, 31(2), 247-257.

doi: 10.1002/acp.3322

Lev-Ari, S., & Keysar, B. (2010). Why don't we believe non-native speakers? The influence of

accent on credibility. *Journal of Experimental Social Psychology*, 46(6), 1093 – 1096. doi:

10.1016/j.jesp.2010.05.025.x

Mann, S., Vrij, A., & Bull, R. (2002). Suspects, lies, and videotape: An analysis of authentic high-

stake liars. *Law and Human Behavior*, 26(3), 365-376. doi: 10.1023/A:1015332606792

Matsumoto, D., Keltner, D., Shiota, M., O'Sullivan, M. & Frank, M. (2008). Facial expressions of

emotion. In M. Lewis, J.M. Haviland-Jones, & L. Feldman Barrett (Eds). *Handbook of*

emotions (3rd Ed). London: The Guilford Press.

Matsumoto, D., Yoo, S. H., & Fontaine, J. (2008). Mapping expressive differences around the

world the relationship between emotional display rules and individualism versus

collectivism. *Journal of Cross-Cultural Psychology*, 39(1), 55-74. doi:

10.1177/0022022107311854

Meissner, C. A., & Kassin, S. M. (2002). " He's guilty!": investigator bias in judgments of truth and deception. *Law and Human Behavior*, 26(5), 469. doi: 10.1023/A:1020278620751

Newman, M. L., Pennebaker, J. W., Berry, D. S., & Richards, J. M. (2003). Lying words:

Predicting deception from linguistic styles. *Personality and Social Psychology*

Bulletin, 29(5), 665-675. doi: 10.1177/0146167203029005010

O'Sullivan, M., Ekman, P., & Friesen, W. V. (1988). The effect of comparisons on detecting deceit.

Journal of Nonverbal Behavior, 12(3), 203-215. doi: 10.1007/BF00987488

Perani, D., & Abutalebi, J. (2005). The neural basis of first and second language

processing. *Current Opinion in Neurobiology*, 15(2), 202-206.

doi: 10.1016/j.conb.2005.03.007

Porter, S., & ten Brinke, L. (2008). Reading between the lies: Identifying concealed and falsified emotions in universal facial expressions. *Psychological Science*, 19(5), 508-514. doi:

10.1111/j.1467-9280.2008.02116.x

Porter, S., & ten Brinke, L. (2009). Dangerous decisions: A theoretical framework for

understanding how judges assess credibility in the courtroom. *Legal and Criminological*

Psychology, 14(1), 119 – 134. doi: 10.1348/135532508281520

Schweitzer, M. E., Brodt, S. E., & Croson, R. T. (2002). Seeing and believing: Visual

access and the strategic use of deception. *International Journal of Conflict Management*,

13(3), 258-375. doi.org: 10.1108/eb022876

Silva-Corvalán, C. (1994). *Language contact and change: Spanish in Los Angeles*.

New York: Oxford University Press.

Strijkers, K., Baus, C., Runnqvist, E., FitzPatrick, I., & Costa, A. (2013). The

temporal dynamics of first versus second language production. *Brain and Language*, 127(1), 6-11. doi: 10.1016/j.bandl.2013.07.008

Street, C. N., & Richardson, D. C. (2015). Lies, damn lies, and expectations: How base rates inform lie-truth judgments. *Applied Cognitive Psychology*, 29(1), 149-155. doi: 10.1002/acp.3085

Strömwall, L., & Granhag, P. A. (2003). How to detect deception? Arresting the beliefs of police officers, prosecutors and judges. *Psychology, Crime and Law*, 9(1), 19-36. doi: 10.1080/10683160308138

Taylor, P. J., Larner, S., Conchie, S. M., & van der Zee, S. (2015). Cross-cultural deception detection. In P. Granhag, A. Vrij, & B. Verschuere (Eds.), *Detecting deception: Current challenges and cognitive approaches* (pp. 175–201). Chichester, UK: Wiley.

ten Brinke, L., Vohs, K. D., & Carney, D. R. (2016). Can ordinary people detect deception after all? *Trends in Cognitive Sciences*, 20(8), 579-588. doi: 10.1016/j.tics.2016.05.012

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5(2), 207-232. doi: 10.1016/0010-0285(73)90033-9

Ullman, M.T. (2001). The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism: Language and Cognitive*, 4(2), 105 – 122. doi: 10.1017/S1366728901000220

Unkelbach, C. (2007). Reversing the truth effect: Learning the interpretation of processing fluency in judgments of truth. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(1), 219-230. doi: 10.1037/0278-7393.33.1.219

Vrij, A. (2015). Deception detection. In B.L. Cutler, & P.A. Zapf (Eds), *APA handbook of forensic psychology, Vol. 2: Criminal investigation, adjudication, and sentencing outcomes* (pp. 225-244). Washington, DC: American Psychological Association. doi: 10.1037/14462-008

Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities*. Chichester: Wiley & 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*(2), 297-312. doi: 10.1348/135532505X60816
- Vrij, A., & Baxter, M. (1999). Accuracy and confidence in detecting truths and lies in elaborations and denials: Truth bias, lie bias and individual differences. *Expert Evidence, 7*(1), 25-36. Doi: 10.1023/A:1008932402565
- Vrij, A., Fisher, R., Mann, S., & Leal, S. (2008). A cognitive load approach to lie detection. *Journal of Investigative Psychology and Offender Profiling, 5*(1), 39 – 43. doi: 10.1002/jip.82
- Vrij, A., Granhag, P.A., Mann, S., & Leal, S. (2011). Outsmarting the liars: Toward a cognitive lie detection approach. *Current Directions in Psychological Science, 20*(1), 28 – 32. doi: 10.1177/0963721410391245
- Vrij, A., Granhag, P. A., & Porter, S. (2010). Pitfalls and opportunities in nonverbal and verbal lie detection. *Psychological Science in the Public Interest, 11*(3), 89-121. doi: 10.1177/1529100610390861
- Zhou, L., Burgoon, J. K., Nunamaker, J. F., & Twitchell, D. (2004). Automating linguistics-based cues for detecting deception in text-based asynchronous computer-mediated communications. *Group Decision and Negotiation, 13*(1), 81-106. doi: 10.1023/B:GRUP.0000011944.62889.6f

Appendix 1.

Questions asked at each interview were standardised

I need to check some details with you;

Where is your place of birth?

In which country was your passport issued? ... and the city?

Hmmm.... I think you're lying to me.

Did someone give you a false identity to enter this country?

Okay then, when was your passport issued?

When does your passport expire?

Please describe the photograph in your passport, give as much detail as you can.

The information you've given me doesn't match your passport information, can you explain this to me?

Table 1.

Descriptive statistics for the nine rating scales (1 = not at all, 7 = very) as a function of Veracity of Interviewee and Language Status of Interviewee

	Truth tellers		Lie tellers	
	Native speakers (<i>N</i> = 13) <i>M</i> (<i>SD</i>)	Non-native speakers (<i>N</i> = 13) <i>M</i> (<i>SD</i>)	Native speakers (<i>N</i> = 13) <i>M</i> (<i>SD</i>)	Non-native speakers (<i>N</i> = 13) <i>M</i> (<i>SD</i>)
How nervous?	3.00 (1.08)	4.00 (1.15)	2.54 (1.20)	3.77 (1.30)
How difficult to remember passport information?	3.38 (1.33)	3.31 (1.84)	4.00 (1.78)	4.00 (1.41)
How difficult to attend to interviewer's questions?	2.08 (1.12)	2.77 (1.83)	2.62 (1.39)	2.00 (1.15)
How difficult to understand interviewer's questions?	2.07 (.95)	2.77 (.93)	1.84 (1.14)	2.69 (.95)
How hard needed to think about answers?	4.23 (1.30)	3.46 (1.71)	3.62 (1.19)	3.84 (1.21)
How hard it was to explain answers?	2.27 (.96)	3.31 (1.18)	3.38 (1.54)	4.62 (.96)
How motivated to appear truthful?	3.77 (1.17)	3.92 (1.44)	4.23 (1.30)	5.15 (1.21)
How much monitored own behaviour?	4.62 (1.26)	3.31 (1.44)	4.46 (1.51)	3.54 (1.33)
How much monitored interviewer's behaviour?	4.00 (1.00)	3.69 (.95)	4.77 (1.24)	4.23 (1.01)

Table 2.

Univariate main and interaction effects (p values) [effect sizes] for the nine rating scales as a function of Veracity of Interviewee and Language Status of Interviewee

	Veracity of Interviewee main effect	Language Status of Interviewee main effect	Veracity x Language Status interaction effect
	<i>F (p) [d]</i>	<i>F(p)[d]</i>	<i>F(p)[η_p^2]</i>
How nervous?	1.10 (.298) [.27]	11.49 (.001) [.95]	.12 (.727) [.01]
How difficult to remember passport information?	2.15 (.149) [.41]	.01 (.932) [.03]	.01 (.932) [.01]
How difficult to attend interviewer's questions?	.09 (.768) [.08]	.01 (.922) [.02]	2.83 (.099) [.06]
How difficult to understand interviewer's questions?	.31 (.580)[.14]	7.74 (.001) [.78]	.08 (.782) [.01]
How hard needed to think about answers?	.09 (.763) [.09]	.50 (.483) [.20]	1.73 (.195)[.04]
How hard it was to explain answers?	1.11 (.297)[.25]	7.26 (.010) [.70]	8.58 (.005) [.15]
How motivated to appear truthful?	5.64 (.022) [.65]	2.28 (.137) [.40]	1.17 (.286) [.02]
How much monitored	.01 (.921) [.03]	8.41 (.006) [.82]	.25 (.619) [.01]

own behaviour?

How much monitored **5.00 (.030) [.61]** 2.09 (.155) [.39] .16 (.695) [.01]

interviewer's

behaviour?

Bold text highlights significance at the $p < .05$ level

Table 3.

Proportion of clips correctly evaluated

	Native speakers		Non-native speakers	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Modality of Presentation				
Visual and Audio	.55	.10	.38	.11
Visual Only	.55	.12	.48	.14
Audio Only	.54	.11	.39	.14
Transcript Only	.52	.14	.55	.14

Table 4.

Descriptive statistics for response bias (β) and discrimination accuracy (d') as a function of Language Status of Interviewee and Modality of Presentation

	β		d'	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Native Language Speakers ($N = 64$)				
Visual and Audio	1.13	.25	.26	.53
Visual Only	1.50	.71	.27	.65
Audio Only	1.03	.12	.21	.52
Transcript Only	1.29	.72	.11	.64
Non-native Language Speakers ($N = 64$)				
Visual and Audio	1.10	.34	-.49	.59
Visual Only	1.24	.39	-.02	.64
Audio Only	.92	.22	-.56	.73
Transcript Only	.99	.21	.27	.65

Table 5.

Univariate main and interaction effects (p values) [effect sizes] for response bias (β) and discrimination accuracy (d') as a function of Language Status of Interviewee and Modality of Presentation

	Language Status of Interviewee main effect	Modality of Presentation main effect	Language x Presentation interaction effect
	$F(p)[d]$	$F(p)[\eta_p^2]$	$F(p)[\eta_p^2]$
Response bias (β)	5.48 (.021) [.40]	4.83 (.003) [.11]	.70 (.556) [.02]
Discrimination accuracy (d')	13.93 (.001) [.63]	2.68 (.050) [.06]	4.02 (.009) [.09]

Bold text highlights significance at the $p < .05$ level

Table 6.

Response bias (β) and discrimination accuracy (d') tested against chance (1 and 0 respectively)

using one sample t-tests across Language Status of Interviewee and Modality of Presentation

	Response bias (β)		Discrimination accuracy (d')	
	<i>M (SD)</i>	<i>t (p) [d]</i>	<i>M (SD)</i>	<i>t (p) [d]</i>
Language Status of Interviewee				
Native Language Speakers	1.24 (.54)	3.49 (.006) [.44]	.21 (.58)	2.91 (.033) [.51]
Non-native Language Speakers	1.06 (.32)	1.50 (.84) [.19]	-.20 (.73)	-2.21 (.186)[.39]
Modality of Presentation				
Visual and Audio	1.12 (.29)	2.17 (.228) [.38]	-.12 (.67)	-.97 (.338)[.25]
Visual Only	1.37 (.58)	3.64 (.006) [.64]	.13 (.65)	1.08 (.289) [.28]
Audio Only	.98 (.19)	-.87 (.391) [.10]	-.18 (.73)	-1.39 (.176)[.35]
Transcript Only	1.14 (.55)	1.42 (.165) [.17]	.19 (.64)	1.68 (.104) [.42]

Bold text highlights significance at the $p < .05$ level, Bonferroni adjusted