Journal of Experimental Psychology: Applied
2016, Vol. 22, No. 4, 488–499

1076-898X/16/$12.00 http://dx.doi.org/10.1037/xap0000102

You Must Be Lying Because I Don't Understand You: Language Proficiency and Lie Detection

Elizabeth Elliott and Amy-May Leach
University of Ontario Institute of Technology

We examined the impact of interviewees' language proficiencies on observers' lie detection performance. Observers ($N = 132$) were randomly assigned to make deception judgments about interviewees ($N = 56$) from Four proficiency groups (i.e., native, advanced, intermediate, and beginner English speakers). Discrimination between lie- and truth-tellers was poorest when observers judged beginner English speakers compared to interviewees from any other proficiency group. Observers were also less likely to exhibit a truth-bias toward nonnative than native English speakers. These results suggest that interviewing individuals in their nonnative languages can create inequalities in the justice system.

Keywords: deception detection, language proficiency, cognitive load

A recent census in the United States revealed that 20% of the population—60.6 million people over the age of five—spoke a language other than English at home (Ryan, 2013). If any of these individuals come into contact with law enforcement officials, they might not have an opportunity to speak in their native languages. Instead, they might attempt to communicate in their nonnative, weaker languages. Communication difficulties between officers and interviewees can arise because of a disparity in proficiency levels (Van Compernolle, 2001). The effect of language proficiency on deception and its detection is unclear, however.

A Multifactor Approach

In their multifactor theory, Zuckerman, DePaulo, and Rosenthal (1981) suggested that attempts to control behavior, cognitive processing, emotion, and generalized arousal affect the expression of deception. Emerging evidence suggests that these factors might also underlie nonnative speech (e.g., Caldwell-Harris & Ayçiçeği-Dinn, 2009). We consider the similarities between deceiving and speaking in a nonnative language, below.

Cognitive Processing and Self-Regulation

Behavioral, self-report, and neuroimaging studies indicate that lie-telling involves greater self-regulation and cognitive processing than truth-telling. Lie-tellers monitor the verbal and nonverbal behaviors of others for signs of suspicion, and control their own behaviors to maintain deceit (Sporer & Schwandt, 2006). In turn, they may appear rehearsed and overly rigid (Vrij, 1995): Compared to truth-tellers, lie-tellers use fewer hand movements, nod less often, and show signs of holding back information (e.g., lip pressing; DePaulo et al., 2003; Sporer & Schwandt, 2006, 2007). Lie-tellers must also inhibit the truth, remember details of the lie, and keep track of who heard the deceptive information to ensure consistency within, and between, statements (Buller & Burgoon, 1996). They think harder than truth-tellers (Evans, Michael, Meissner, & Brandon, 2013; Vrij, Fisher, Mann, & Leal, 2008) and exhibit more signs of heightened cognitive load (e.g., greater vocal uncertainty, longer response latency), as a result (DePaulo et al., 2003; Sporer & Schwandt, 2006, 2007). In addition, several regions of the brain—such as the prefrontal cortex—are more active when individuals tell a lie (e.g., Ganis, Kosslyn, Stose, Thompson, & Yurgelun-Todd, 2003; Lee et al., 2005; Spence et al., 2001). A meta-analysis suggests that lying specifically increases the use of areas associated with executive control (e.g., working memory; Christ, Van Essen, Watson, Brubaker, & McDermott, 2009).

Recently, researchers have posited that speaking in a nonnative language places similar demands on cognitive resources (Evans et al., 2013). Individuals with lower language proficiencies are unable to process language automatically and, consequently, rely more heavily on working memory (Volk, Köhler, & Pudelko, 2014). Nonnative speakers must also continuously inhibit their native languages (Abutalebi & Green, 2007). Perhaps as a result, they are slower to respond in tests of word production (Ivanova & Costa, 2008) and understanding (Ransdell & Fischler, 1987). Although native and nonnative languages activate the same regions in the brain, there is increased activity when the age of nonnative language acquisition is higher, there is a low level of proficiency, and there has been less exposure to the nonnative language (Perani & Abutalebi, 2005). As with lying, speaking a nonnative language has been shown to increase activation of the prefrontal cortex (Lazar, Stern, & Cohen, 2014; Perani et al., 2003). Areas associated with executive control functions, such as working memory,

Elizabeth Elliott and Amy-May Leach, Faculty of Social Science and Humanities, University of Ontario Institute of Technology.

We thank Tiffany Barone, Victoria Campbell, Nathmoulica Doung, Maggie Howe, Karen Joseph, Nicole Kingston, Kaylee-Anne Kosoralo, Tyson McLeish, and Alexis Sokarijo for their assistance with data collection. We also want to thank the staff and students at COSTI and Culture-Works for their help and participation.

Correspondence concerning this article should be addressed to Elizabeth Elliott, Faculty of Social Science and Humanities, University of Ontario Institute of Technology, 2000 Simcoe Street North, Oshawa, Ontario L1H 7K4, Canada. E-mail: elizabeth.elliott@uoit.ca

are particularly implicated (Adesope, Lavin, Thompson, & Ungerleider, 2010).

These findings suggest that lie-tellers and nonnative speakers might be cognitively overloaded when subjected to additional demands. Indeed, nonnative speakers provide less detail than native speakers, especially when asked to maintain gaze—a cognitively difficult task (Chiu, Hong, & Krauss, 2001). Lie-tellers are similarly affected (Vrij, Mann, Leal, & Fisher, 2010). Speaking a nonnative language while lying may, thus, doubly tax cognitive resources. Evans et al.'s (2013) work on the Psychologically Based Credibility Assessment Tool supports this hypothesis. Of the nine cognitive processing cues tested, seven differentiated between lie- and truth-telling when interviewees were nonnative English speakers, compared to only three when interviewees were native English speakers. This finding suggests that nonnative speakers exhibit more overt signs of cognitive load during deception than native speakers.

Arousal and Emotion

Lying has been proposed to be accompanied by feelings of guilt, fear, and excitement (Ekman, 1989), as well as underlying physiological arousal (see Vrij et al., 2008 for a discussion). In fact, the polygraph was developed based on the notion that lie-tellers exhibited higher skin conductance, heart rates, and breathing rates (i.e., autonomic arousal) than truth-tellers (Lykken, 1974). Several behavioral studies have demonstrated that deception increases nervousness, vocal pitch, pupil dilation, facial unpleasantness, and negative statements—all of which are signs of emotion and/or arousal (DePaulo et al., 2003; Sporer & Schwandt, 2006, 2007). During deception, brain activity increases in the regions that are associated with emotional processing, as well (Abe, Suzuki, Mori, Itoh, & Fujii, 2007).

Differences, in terms of physiological arousal and the experience of emotions, also exist between native and nonnative speakers of a particular language. Speaking in a nonnative language increases nervousness (Gregersen, 2005), pupil dilation (Duñabeitia, & Costa, 2015; Hyönä, Tommola, & Alaja, 1995), and neural activity in regions implicated in the regulation of emotions (Lazar et al., 2014). However, the emotional blunting hypothesis stipulates that speaking in a less proficient language can reduce emotion and arousal (e.g., Bond & Lai, 1986; Pavlenko, 2006). For example, nonnative speakers report feeling less arousal in response to emotional statements (Caldwell-Harris & Ayçiçeg˘i-Dinn, 2009). Psychotherapists have noted that patients exhibit emotional detachment when speaking in their nonnative languages (Marcos, 1976). Although these patients could readily verbalize emotions, they rarely displayed them. Keysar, Hayakawa, and An (2012) proposed that the emotionality associated with speaking another language even impacts decision-making abilities. In a risk-taking scenario, nonnative speakers were less emotional than native speakers and were able to make decisions without bias. On the whole, the majority of research has indicated that speaking in a nonnative language reduces emotion and arousal; however, a few studies have suggested that arousal can be heightened. Both emotional blunting and second language anxiety are, therefore, likely to be responsible for the emotional involvement of nonnative speakers.

Implications for nonnative speaking deceivers are unclear because lying increases arousal and emotions, whereas nonnative language speakers tend to experience decreases in subjective feelings of emotion. In one test, no differences were found between lie- and truth-tellers who spoke in their native or nonnative languages in terms of physiological arousal (Caldwell-Harris & Ayçiçeğii-Dinn, 2009). However, when observers used the Psychologically Based Credibility Assessment Tool, nervousness and negativity were diagnostic of deception in nonnative speakers but not in native speakers (Evans et al., 2013).

Language Proficiency and Deception Detection

Although there has been considerable research on the effects of deception and speaking in a nonnative language, independently, only a handful of studies to date have been conducted on the impact of language proficiency on implicit deception detection. In a study by Cheng and Broadhurst (2005), interviewees were randomly assigned to lie or tell the truth, in either their native language (i.e., Cantonese) or nonnative language (i.e., English), while providing their opinions about a moral issue. No differences were found between observers' abilities to detect nonnative and native speakers' deception.

Two additional studies yielded different findings (Da Silva & Leach, 2013; Leach & Da Silva, 2013). Both studies utilized the same video stimuli featuring interviewees who were questioned about a transgression (i.e., cheating on a test) in English—their native or nonnative language. Laypersons and police officers were better able to detect deception in native speakers than in nonnative speakers. In keeping with previous findings (Bond & DePaulo, 2006), observers had a tendency to believe that native speakers were telling the truth. Moreover, Da Silva and Leach (2013) found that observers were more likely to assume that nonnative speakers were telling lies; Leach and Da Silva (2013) did not find any bias toward this group. These findings suggest that observers perceived native speakers more positively than nonnative speakers.

These biases were also found in a different context: alibis (Evans & Michael, 2014). Laypersons watched alibi statements made by interviewees who were either lying or telling the truth in English (i.e., their native or nonnative language). Observers were unable to discriminate between lie- and truth-tellers, regardless of whether they spoke in their native or nonnative languages. They were less likely to exhibit a truth-bias toward nonnative speakers, however. Thus, although nonnative speakers were not perceived negatively (i.e., there was no lie-bias), they were also not perceived as positively as native English speakers (i.e., observers did not tend to think that nonnative speakers were telling the truth).

Next, the impact of language and cultural background on observers' deception judgments was examined (Castillo, Tyson, & Mallard, 2014). Columbian and Australian interviewees were instructed to lie or tell the truth about having committed a computer crime (i.e., reading confidential emails). Australians spoke in their native language (i.e., English), whereas Colombians spoke in their native language (i.e., Spanish) or nonnative language (i.e., English). Observers were better able to discriminate between lie- and truth-tellers of varying cultural backgrounds (i.e., Colombian vs. Australian), but they were unable to discriminate between lie- and truth-tellers of varying language proficiencies (i.e., those speaking in a native vs. nonnative language). Observers also exhibited a truth-bias toward native speakers (i.e., Australian and Colombian-English) and a lie-bias toward nonnative speakers (i.e., Colombian-Spanish). Overall, interviewees' language proficiency negatively impacted observers' decision making and bias, whereas cultural background did not.

Interviewees were dichotomously categorized as either native speakers or nonnative speakers in all of these studies. Evans et al. (2013) posited that deception detection effects might not be uniform across all nonnative speakers (i.e., one's level of proficiency in the nonnative language should affect deceptive abilities). In their study, interviewees were categorized as native English speakers, non-English speakers with high levels of proficiency, and nonnative English speakers with low levels of proficiency. All interviewees lied or told the truth about their alibis. Contrary to previous studies, observers were equally accurate at detecting the deception of interviewees who had high and low nonnative English proficiencies, and were significantly less accurate when they viewed native English speakers.

There are discernible similarities and differences across the abovementioned studies. Observers repeatedly believed that native speakers were telling the truth to a greater extent than nonnative speakers (Castillo et al., 2014; Da Silva & Leach, 2013; Evans & Michael, 2014; Leach & Da Silva, 2013). In terms of accuracy, three studies (i.e., Castillo et al., 2014; Cheng & Broadhurst, 2005; Evans & Michael, 2014) failed to reveal effects of language proficiency on deception detection. Those studies in which an effect of language proficiency was found yielded contradictory results. In Da Silva and Leach (2013) and Leach and Da Silva (2013), observers were better able to discriminate between lie- and truth-tellers who were native than nonnative English speakers, whereas Evans et al. (2013) reported that observers were better able to detect deception in nonnative than native English speakers. Thus, despite similarities across studies in terms of bias, there appear to be mixed effects of language proficiency on accuracy; the reasons for these inconsistencies are unclear.

Sources of Variability Between Studies

These mixed findings may be due to a number of factors. First, nonnative speakers' proficiencies differed between studies. Interviewees in Da Silva and Leach's (2013) and Leach and Da Silva's (2013) studies had been assigned a "basic" level of English proficiency (i.e., 4 or less on a scale from 1 to 12) using standardized Canadian Language Benchmarks (Centre for Canadian Language Benchmarks, 2012). Although a different scale was used in Evans and Michael's (2014) study, nonnative speakers appeared to rate their English proficiencies considerably higher (i.e., 6.14 on a 10-point scale). Finally, Evans et al.'s (2013) low- and high-proficiency interviewees attended university in English (i.e., both groups were highly proficient in English). Rather than previous research findings reflecting a lack of replicability, they might suggest that lie detection performance depends on whether nonnative interviewees are beginner (Da Silva & Leach, 2013; Leach & Da Silva, 2013), intermediate (Evans & Michael, 2014), or advanced (Evans et al., 2013) English speakers.[1] Reframing previous findings, it might be easiest to detect deception in advanced nonnative speakers, followed by intermediate nonnative speakers, and finally, beginner nonnative speakers.

Second, interviewees' language use was not uniform in all studies. In the majority of studies, interviewees spoke in either their native or nonnative languages, depending on the assigned condition. Cheng and Broadhurst (2005) allowed interviewees to alternate between speaking their nonnative and native languages, however. This practice—known as code-switching (Gumperz & Hymes, 1986)—might have reduced cognitive load (i.e., nonnative speakers could conserve cognitive resources by switching to a less demanding language). Reductions in cognitive load could account for the absence of proficiency-related differences in that study.

Third, language proficiency has been assessed using differing measures. Interviewees were categorized as native or nonnative speakers using either self-report (Castillo et al., 2014; Cheng & Broadhurst, 2005), responses to a language history questionnaire (Evans et al., 2013; Evans & Michael, 2014), or standardized tests (Da Silva & Leach, 2013; Leach & Da Silva, 2013). The accuracy of subjective measures of language proficiency is unknown. There may not have been effects of language proficiency in some studies—such as Cheng and Broadhurst (2005) and Castillo et al. (2014)—because nonnative and native speakers were not able to accurately self-report their own proficiencies. In turn, there might have been few objective differences between these groups at the outset.

Finally, the type of paradigm that was used to elicit deceptive behavior differed between studies. That is, interviewees discussed alibis (Evans et al., 2013; Evans & Michael, 2014), cheating (Da Silva & Leach, 2013; Leach & Da Silva, 2013), a computer crime (Castillo et al., 2014), or opinions (Cheng & Broadhurst, 2005). Each of the paradigms placed a different combination of demands on emotion and cognitive resources. For example, interviewees who lied about cheating might have felt intense guilt about having committed a transgression and fear of being caught or punished (high emotion, low cognitive load), whereas interviewees who fabricated alibis were required to invent and remember plausible events (low emotion, high cognitive load). The different pattern of results across studies might have been due to the underlying demands of the deception paradigms. Nonnative speakers' deception might have been more difficult to detect than that of native speakers only in those paradigms that imposed emotion (e.g., Da Silva & Leach, 2013).

The Present Research

We addressed the methodological variations in previous research. In this study, the experimenter only spoke English and interviewees were not permitted to code-switch. A single paradigm was used to ensure that the same cognitive and emotional demands were placed on all interviewees. In addition, interviewees' language proficiencies were determined using standardized tests to ensure objectivity, to generate more precise measurements of effects at each level of language proficiency, and to allow for direct comparisons between levels of proficiency. We also examined how accuracy and bias varied across a full range of proficiencies (i.e., beginner, intermediate, advanced, and native English speakers). Finally, we tested how proficiency affected subjective and objective cues to deception—an issue that has received little attention to date.

[1] Proficiency levels were not reported in two studies (i.e., Castillo et al., 2014; Cheng & Broadhurst, 2005).

Hypotheses

Discrimination. Based on previous studies (e.g., Da Silva & Leach, 2013), we hypothesized that discrimination between lie- and truth-tellers would be poorer when interviewees were beginner English speakers than native speakers. We expected observers to be equally able to detect deception in intermediate and native English speakers, as in Evans and Michael's (2014) study. Finally, based on Evans et al.'s (2013) findings, we posited that observers would be more accurate when detecting the deception of advanced speakers than native English speakers.

Bias. We hypothesized that interviewees with native English proficiency would be more likely to be judged as truth-tellers than interviewees in lower proficiency groups (i.e., beginner, intermediate, and advanced English speakers). This hypothesis was consistent with previous findings (e.g., Bond & DePaulo, 2008; Da Silva & Leach, 2013; Evans & Michael, 2014).

Cues to deception. We predicted that beginner English speakers would self-report and exhibit more cues associated with cognitive load than native speakers, particularly when they lied. Given previous mixed effects of proficiency on emotion (e.g., Caldwell-Harris & Ayçiçeği-Dinn, 2009; Evans et al., 2013), analyses related to that factor were exploratory in nature.

Method

We analyzed the impact of language proficiency on deception detection. Native and nonnative English-speaking community members and undergraduate students were randomly assigned to lie or tell the truth about an event. Later, observers watched the recorded interviews and attempted to differentiate between the lie- and truth-tellers. The interviews were also analyzed by coders for the presence of deception cues.

Phase 1

Research design. A 2 (Veracity: lie vs. truth) × 4 (Proficiency: beginner English vs. intermediate English vs. advanced English vs. native English) between-subjects design was used in this study. That is, interviewees from four language proficiency groups were randomly assigned to either lie or tell the truth.

Participants. Prior to data collection, we decided to collect data from as many interviewees as necessary to create a heterogeneous sample of 14 interviewees from each proficiency group, matching for age, gender, and race. Interviewees were recruited from two established centers that provided language training services to prospective university students and new immigrants. The centers used standardized English tests (i.e., Canadian Language Benchmarks) to assign students to language proficiency groups; we recruited interviewees from within the beginner, intermediate, and advanced English groups. Native English speakers were recruited from undergraduate psychology courses at a midsized Canadian university. Participants who were recruited from language centers were paid $10 CAD, whereas university students were compensated with course credit.

Of the large sample of participants collected ($N = 110$), we were able to match 56 interviewees (females = 37, males = 19, $M_{\text{age}} = 27.51$, $SD_{\text{age}} = 9.18$) across proficiencies according to veracity, age, race, and gender (e.g., there was a 19-year-old lie-telling male who self-identified as Latin American in each of the native, advanced, intermediate, and beginner English groups). They self-identified as Arab/West Asian (10.9%), Black (25.5%), Chinese (9.1%), Filipino (5.5%), Hispanic (7.3%), Japanese (3.6%), Latin American (23.6%), South Asian (3.6%), South East Asian (7.3%), White (1.8%), or other (1.8%). Interviewees who confessed (i.e., revealed that they were instructed to lie) during the interview were excluded from the study ($n = 15$).

Materials.

Videos. Interviewees watched one of two videos: the *Innocuous* video or the *Suspicious* video. The focal point of both videos was a computer desk that was covered with office supplies and personal belongings. During each video, the camera zoomed into the background of the scene and revealed a wall decorated with pictures, a map, and newspaper clippings. The items in the *Innocuous* video were meant to replicate a typical office setting, whereas the items in the *Suspicious* video were intended to suggest that a terrorist act was being planned. There were an equal number of items in both videos. Each critical *Suspicious* item was replaced with an equivalent item in *Innocuous* video (e.g., a bomb-making manual vs. an exhaust vent manual).

Demographics questionnaire. Interviewees were asked to provide information related to age, gender, and race. They were asked to rate their English proficiencies on a 5-point Likert scale from 1 (*poor*) to 5 (*excellent*) and nine other questions pertaining to their language proficiencies (e.g., "What language[s] do you consider your native [or first] language[s]?" "What language[s] do you speak at home?").

Experimental questionnaire. We used a manipulation check to confirm that interviewees watched the video, understood its contents, and were aware of the instructions throughout the experiment. Interviewees were required to indicate the items that were present in the video using a checklist. This checklist included 11 items that were not present in the video (e.g., a plant) and nine items that were featured in the video (e.g., a calendar).

The questionnaire also included 19 questions related to cognitive load (e.g., "How hard did you have to think about your answers?") and emotion (e.g., "How nervous were you when answering the interviewer's questions?") which were rated on a 10-point Likert scale (1 = *not at all*, 10 = *extremely*). These questions were created to assess the cognitive state of the interviewee during the interview and were based on the existing literature pertaining to cognitive load (Vrij et al., 2008) and emotion (Caldwell-Harris & Ayçiçeği-Dinn, 2009).

Interview questions. Interviewees were asked closed and open-ended questions regarding the video that they had watched. The questions were increasingly more specific as the interview progressed (e.g., "What did you see on the wall?" vs. "What was marked on the calendar?"). These questions were asked in English and code-switching was not permitted.

Procedure. The experiment was conducted in a small room in the university or language center. A female experimenter greeted the interviewee and explained that he or she must watch a short video clip on a laptop and follow the instructions that were provided on the screen. The experimenter began the MediaLab program (Jarvis, 2014), entered a randomly assigned participant number, and exited the room. The preprogrammed MediaLab file randomly assigned the interviewee to view either the *Innocuous* or *Suspicious* video. The video was delayed for 1 min to ensure that the experimenter had left the room before it begun to play and that she would remain blind to condition. During the delay, text on the computer screen prompted interviewees to remember details from the video because the remainder of the experiment would involve a memory task. At the end of the video, interviewees read on-screen instructions regarding the upcoming interview with the experimenter. The instructions varied based on condition: interviewees who viewed the *Innocuous* video were told to answer all of the experimenter's questions honestly and accurately (i.e., tell the truth), whereas the interviewees who viewed the *Suspicious* video were instructed to lie to the experimenter about the items in the video. Interviewees were given two minutes to prepare for the interview. To motivate the interviewees in both conditions, the instructions stated that a \$50 CAD reward would be offered to an interviewee if the experimenter was convinced that he or she was telling the truth. In fact, all interviewees were entered into a draw for the money.

The experimenter reentered the room after seven minutes, turned on the video camera, and asked the interview questions. Finally, the experimenter provided the interviewee with the demographics questionnaire and the experimental questionnaire. Once the questionnaires were completed, the interviewee was debriefed and entered into a draw to win the reward. The experimenter also obtained consent from the interviewee to utilize video footage of the interview in the next phase of the study. Each session lasted approximately 30 minutes.

Cue analysis. Two independent raters methodically assessed each interview. Half of the interviews ($n = 28$) were transcribed by each rater. Both raters coded all of the footage and corresponding transcripts (Cohen's $\kappa = .67$). They were instructed to record the duration and number of times that each of the following verbal and nonverbal cues were observed: smiling, covering mouth, covering eyes, gaze aversion, facial fidgeting (e.g., touching face), illustrators (i.e., hand movements used for description), ah speech (e.g., "umm," "uhh," "mmm," "hmm," "mhmm"), lack of memory, negative statements and complains, repetition (i.e., repeating a word or phrase without pause), non-'uhh' speech hesitation (e.g., "like," "so," "you know"), admitted uncertainty (e.g., "I'm not sure"), and spontaneous corrections. Using a Likert scale (1 = *not at all*, 5 = *extremely*), coders recorded the extent to which a target appeared to exhibit the following behaviors: nervousness, vocal tension, emotionality, surprise, cooperativeness, plausibility, thinking hard, difficulty answering the experimenter, difficulty understanding the experimenter, and amount of detail provided. The list of verbal and nonverbal cues coded was based on previous research into deception cues (DePaulo et al., 2003) and cognitive load (e.g., Mann, Vrij, & Bull, 2002).

Phase 2

Research design. We used a Veracity × Proficiency mixed-factors design, with veracity as the repeated measure. That is, observers were randomly assigned to view a compilation of videos of interviewees who were lying and telling the truth, and were speaking English either at a beginner, intermediate, advanced, or native level of proficiency.

Participants. An a priori power analysis using G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) and effect sizes from Leach and Da Silva's (2013) study revealed that 132 participants were required to achieve sufficient power (i.e., 0.95). Thus, we recruited 132 undergraduates (females = 80, males = 52, $M_{\text{age}} = 19.21$ years, $SD_{\text{age}} = 1.85$) to participate in the study in exchange for course credit. These observers self-identified as Arab/West Asian (6.8%), Black (9.1%), Chinese (6.1%), Latin American (0.8%), South Asian (32.6%), South East Asian (6.1%), White (34.1%), or other (4.5%).

Materials.

Video footage. We compiled 14 videos—seven truth-tellers and seven lie-tellers—from Phase 1 for each level of English proficiency (i.e., beginner, intermediate, advanced, and native English speakers). The interviewees' facial features and upper bodies were visible throughout the videos. The average length of each video was 187.50 seconds ($SD = 83.10$). We randomized the order in which the videos were shown.

Demographics questionnaire. Observers were asked to provide information related to their ages, genders, and races.

Judgment questionnaire. Using this questionnaire, observers indicated whether each interviewee was lying or telling the truth. Additionally, observers were asked to indicate, on a scale from 0% (*not at all confident*) to 100% (*extremely confident*), the degree to which they were confident in each of their decisions.[2]

Procedure. The observers were tested individually in a quiet room. Upon arrival to the experimental session, an experimenter asked observers to sign a consent form. Each observer was instructed to sit at a computer and watch the randomly assigned video footage of interviewees from one of the four language proficiency conditions. Following each video, observers were prompted to complete the corresponding item on the judgment questionnaire. Once all of the videos were viewed, observers were instructed to complete the demographics questionnaire and then they were debriefed. Each session lasted approximately one hour.

Results

Phase 1: Interviewees

Preliminary analyses on gender and race were conducted to eliminate potential covariates. The following analyses were collapsed across these variables because effects were nonsignificant.

Self-reported proficiency. We conducted a one-way analysis of variance (ANOVA) on interviewees' ratings in response to the question, "What is your English proficiency?" There was a significant main effect of proficiency, $F(3, 53) = 34.39$, $p < .001$, $\eta_p^2 = .66$, 95% confidence interval (CI) [.48, .74]. Post hoc Tukey's tests revealed that native English speakers ($M = 4.93$, $SD = 0.27$) self-reported that they were more proficient than advanced ($M = 3.57$, $SD = 0.51$), intermediate ($M = 3.21$, $SD = 0.58$), and beginner English speakers ($M = 3.29$, $SD = 0.61$), $ps < .001$. The nonnative English proficiencies did not significantly differ from one another (all $ps > .05$).

We also conducted a one-way ANOVA on interviewees' responses to the question, "How many years have you been speaking English?" Again, there was a significant main effect of proficiency, $F(3, 53) = 9.82$, $p < .001$, $\eta_p^2 = .36$, 95% CI [.13, .49]. Post hoc analyses revealed that beginners ($M = 2.74$, $SD = 2.92$)

[2] These data are available from the corresponding author.

had been speaking English for significantly fewer years than intermediate speakers (M = 4.51, SD = 10.34), who had fewer years speaking English than advanced speakers (M = 7.04, SD = 12.67) and, finally, native English speakers (M = 18.36, SD = 2.31; all $ps \leq .004$). These findings are indicative of the fact that interviewees' self-reported proficiency was inconsistent with their proficiency groups; however, their years spent speaking English differed according to their proficiency level.

Manipulation check. We conducted one-sample t-tests to compare the correct number of items that interviewees indicated were present in the video to the actual number of items (i.e., nine). There were no significant differences for beginner (M = 8.93, SD = 0.27), $t(14) = -1.00$, p = .336, $d = -.26$, 95% CI [.00, .75], intermediate (M = 8.85, SD = 0.38), $t(14) = -1.48$, p = .165, $d = -.38$, 95% CI [.00, .89], advanced (M = 8.86, SD = 0.36), $t(14) = -1.47$, p = .165, $d = -.38$, 95% CI [.00, .89], and native English speakers (M = 9.00, SD = 0.00).[3] These results indicate that interviewees accurately recalled the videos' contents.

Motivation. A Proficiency × Veracity ANOVA was conducted on interviewees' motivation ratings. There were no significant differences in interviewees' motivation across proficiency groups, $F(3, 53) = 1.72$, p = .175, η_p^2 = .09, 95% CI [.00, .21], nor veracity conditions, $F(1, 55)$ = .38, p = .543, η_p^2 = .01, 95% CI [.00, .10]. Additionally, a one-sample t test revealed that the average motivation score (M = 6.80, SD = 2.84) was significantly higher than 1 (*not at all motivated*), $t(56) = 17.90$, $p < .001$, d = 2.37, 95% CI [1.86, 2.87]. Thus, interviewees were motivated to convince the experimenter that they were telling the truth.

Emotion. We conducted a Proficiency × Veracity multivariate analysis of variance (MANOVA) on interviewees' responses to the nine emotion items on the experimental questionnaire. There was no statistically significant effect of proficiency (see Table 1), $F(3, 22) = 1.08$, p = .420, Wilks' λ = .12, η_p^2 = .13, 95% CI [.00, .31], or veracity (see Table 2), $F(1, 23)$ = .53, p = .821, Wilks' λ = .65, η_p^2 = .02, 95% CI [.00, .22], on the combined dependent variables. There was also no significant interaction between proficiency and veracity, $F(3, 22)$ = .71, p = .812, Wilks' λ = .21, η_p^2 = .07, 95% CI [.00, .22]. These results indicate that emotions were not effected by proficiency or veracity.

Cognitive load. We conducted a Proficiency × Veracity MANOVA on interviewees' responses to six items on the experimental questionnaire pertaining to cognitive load. There was a significant effect of proficiency (see Table 1), $F(3, 53) = 2.54$, p = .001, Wilks' λ = .40, η_p^2 = .13, 95% CI [.00, .27], on the combined dependent variables. We examined the univariate effects more closely using a Bonferroni adjusted alpha level of .008. Proficiency had a significant effect on interviewees' difficulty understanding the experimenter's questions, $F(3, 53) = 7.37$, $p <$ .001, η_p^2 = .30, 95% CI [.08, .44]. Post hoc Tukey's tests revealed that beginner and intermediate English speakers had a harder time understanding the experimenter's questions than native English speakers, p = .001 and p = .003, respectively; however, there were no significant differences between any of the proficiencies and advanced English speakers. We also found a significant effect of veracity (see Table 2), $F(1, 55) = 3.84$, p = .004, Wilks' λ = .65, η_p^2 = .05, 95% CI [.00, .20] on the combined dependent variables. Veracity had a significant effect on interviewees' difficulty answering the experimenter's questions, $F(1, 55) = 21.172$, $p < .001$, η_p^2 = .29, 95% CI [.10, .45]. An independent samples t test revealed that lie-tellers, had more difficulty answering the experimenter's questions than truth-tellers, $t(55) = 4.41$, $p < .001$, d = .59, 95% CI [.30, .87]. There was no significant interaction between veracity and proficiency, $F(3, 53) = 1.32$, p = .190, η_p^2 = .07, 95% CI [.00, .19]. Thus, language proficiency and veracity impact interviewees abilities to understand and answer the experimenter's questions.

Cues to deception. A Veracity × Proficiency MANOVA was conducted on interviewees' behaviors. There was a significant effect of proficiency, $F(3, 53) = 3.67$, $p < .001$, Wilks' λ = .295, η_p^2 = .17, 95% CI [.00, .32] on the combined dependent variables. Two statistically significant differences were revealed when univariate effects were examined using a Bonferroni adjusted alpha of .007: the use of "uhh" speech hesitation, $F(3, 53) = 5.66$, p = .002, η_p^2 = .24, CI [.04, .39], and repetition, $F(3, 53) = 6.29$, p = .001, η_p^2 = .26, CI [.06, .41]. Post hoc Tukey's tests revealed targets who spoke intermediate (M = 19.86, SD = 11.08), p = .048, and advanced English (M = 22.79, SD = 11.45), p = .005, uttered "uhh" speech hesitations more than those who spoke native English (M = 10.57, SD = 4.94). Beginner English speakers (M = 12.50, SD = 6.47) also produced more "uhh" speech hesitations than advanced English speakers, p = .023. However, there were no differences between beginner and intermediate or beginner and native English speakers, all $ps > .05$.

We also found that intermediate and advanced English speakers (M = 6.57, SD = 4.18) repeated more words or phrases than native English speakers (M = 1.21, SD = 1.72), p = .014 and p = .001, respectively. There were no differences in terms of repetition between advanced or native English speakers and beginner (M = 4.64, SD = 4.07) and intermediate English speakers (M = 5.29, SD = 3.95), all $ps > .05$. There was no main effect of veracity, $F(1, 55) = 1.84$, p = .113, η_p^2 = .03, CI [.00, .16], nor an interaction between veracity and proficiency, $F(3, 53) = 1.23$, p = .252, η_p^2 = .07, CI [.00, .18]. These results indicate that proficiency level does impact certain cues to deception, whereas veracity does not.

Phase 2: Observers

To eliminate covariates, preliminary analyses included observer gender and race. Effects were nonsignificant; therefore, the following analyses collapsed across those variables.

Accuracy. Accuracy was calculated by assigning a "0" to each inaccurate decision and a "1" to each accurate decision, and then averaging each observer's scores. Overall accuracy (M = .57, SD = .14) ranged from a minimum of .21 to a maximum of .93 (see Table 3). Observers' accuracy differed across language proficiency groups. They performed significantly better than chance (i.e., .50) when they viewed native (M = .60, SD = .14), $t(32)$ = 4.04, $p < .001$, d = .70, 95% CI [.32, 1.08], advanced (M = .57, SD = .13), $t(32) = 3.10$, p = .004, d = .54, 95% CI [.17, .90], and intermediate English speakers (M = .62, SD = .12), $t(35) = 5.78$, $p < .001$, d = .96, 95% CI [.56, 1.36]. Observers did not significantly differ from chance when they attempted to detect the deception of beginner English speakers (M = .50, SD = .12), $t(33)$ = .00, p = 1.00.

[3] p Values could not be computed because there was no variability.

Table 1
Table of Means and Standard Deviations in Terms of Proficiency for Measures of Emotion and Cognitive Load

Measure	Native		Advanced		Intermediate		Beginner	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Emotion								
Nervous	4.57	3.00	3.86	3.29	5.29	1.50	3.75	2.75
Excited	3.57	2.64	7.29	1.60	6.43	2.70	2.00	1.41
Guilty	2.86	3.34	5.71	3.30	5.00	2.16	2.75	2.06
Surprised	5.29	1.80	7.29	2.14	5.43	2.94	2.75	1.71
Ashamed	3.71	3.25	4.14	2.55	4.43	2.44	1.75	1.50
Afraid	3.57	3.55	3.14	2.41	3.14	2.55	3.00	2.83
Anxious	5.71	3.04	5.71	3.64	4.29	2.87	2.75	2.06
Negative	2.71	2.63	1.86	1.86	2.29	1.80	1.75	.96
Emotional	3.14	3.34	4.29	1.60	4.43	2.57	3.00	1.83
Cognitive load								
Difficulty answering	4.64	1.34	4.93	2.70	5.46	2.60	5.14	2.48
Difficulty understanding	1.21*	.80	2.50	1.99	4.23*	2.77	4.50*	2.74
Thinking hard	5.57	1.91	5.86	2.77	4.54	2.76	4.86	2.41
Planning	5.57	1.83	7.21	2.58	4.85	2.19	5.14	2.28
Remembering	5.29	2.46	5.93	2.53	5.00	3.06	5.64	2.41
Paying attention	4.86	3.39	6.21	2.99	4.69	2.75	6.14	2.83

* $p < .05$.

A Proficiency × Veracity mixed-factors ANOVA was conducted on observers' accuracy. Observers were significantly more accurate when judging truth-tellers ($M = .64$, $SD = .20$) than lie-tellers ($M = .50$, $SD = .22$), $F(1, 131) = 24.85$, $p < .001$, $\eta_p^2 = .16$, 95% CI [.06, .27]. We also found a main effect of proficiency (see Figure 1), $F(3, 129) = 5.55$, $p = .001$, $\eta_p^2 = .12$, 95% CI [.02, .71]. Post hoc Tukey's tests revealed that observers were less accurate when they judged beginner English speakers than when they judged native English speakers, $t(63) = 3.09$, $p = .003$, $d = .39$, 95% CI [.13, .64], advanced English speakers, $t(63) = 2.23$, $p = .029$, $d = .28$, 95% CI [.03, .53], or intermediate English speakers, $t(66) = 3.99$, $p < .001$, $d = .49$, 95% CI [.23, .74]. However, there were no significant differences between the latter three groups in terms of accuracy. There was no significant interaction between veracity and proficiency, $F(3, 129) = 1.46$, $p = .229$, $\eta_p^2 = .03$, 95% CI [.00, .09]. These results suggest that language proficiency and veracity do impact observers' accuracy judgments.

Signal detection theory. Signal detection theory (Green & Swets, 1966) was used to analyze observers' response biases (i.e., β) and their abilities to discriminate between lie- and truth-tellers (i.e., d'). In signal detection theory, discrimination is described as an individual's ability to correctly detect the presence of a signal (e.g., a lie) and make a correct rejection in the absence of a signal (e.g., the truth). Bias refers to the tendency of an individual to choose one response over another (e.g., tending to indicate that the truth was told).

An a priori decision was made to use Wixted and Lee's (n.d.) formulas to calculate discrimination (i.e., $d' = Z_{FA} - Z_{Hit}$) and bias (i.e., $c = -\frac{Z_{FA}+Z_{Hit}}{2}$). A decision was classified as a hit if the observer rendered a "lie" judgment in response to a lie-teller, whereas a false alarm was registered when the observer indicated that the interviewee was telling a lie but, in fact, the truth was told. A standard correction was performed on all hit rates of "1" or false alarms of "0." That is, false alarms of "0" were changed to the equivalent of one divided by two times the maximum number of false alarms (i.e., 7). Hits were changed to the equivalent of one subtracted by the product of one divided by two times the maximum number of false alarms (i.e., 7). Therefore, false alarms were changed from "0" to ".07" and hits were changed from "1" to ".93," as per Wixted and Lee's (n.d.) suggestion.

Discrimination. We conducted a one-way ANOVA on observers' abilities to discriminate between lie- and truth-tellers.

Table 2
Table of Means and Standard Deviations in Terms of Veracity for Measures of Emotion and Cognitive Load

Measure	Lie-teller		Truth-teller	
	M	*SD*	*M*	*SD*
Emotion				
Nervous	4.82	2.36	4.14	2.85
Excited	4.36	3.08	5.79	2.78
Guilty	5.27	2.94	3.43	2.88
Surprised	5.18	2.44	5.71	2.76
Ashamed	3.91	2.74	3.57	2.62
Afraid	3.45	2.95	3.07	2.59
Anxious	4.91	2.66	4.79	3.47
Negative	2.45	2.12	2.00	1.80
Emotional	4.18	2.44	3.50	2.47
Cognitive load				
Difficulty answering	6.21*	2.03	3.81*	1.90
Difficulty understanding	3.79	2.91	2.37	1.88
Thinking hard	6.00	2.30	4.41	2.42
Planning	6.18	2.43	5.22	2.23
Remembering	6.11	2.35	4.81	2.66
Paying attention	5.93	2.46	5.04	3.47

* $p < .05$.

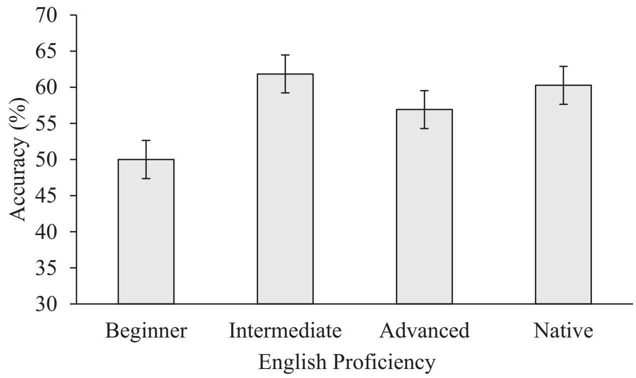

Figure 1. Observers' overall accuracy when judging lie- and truth-tellers across four proficiency groups. Error bars represent 95% confidence intervals.

There was a significant main effect of proficiency, $F(3, 129) = 5.50$, $p = .001$, $\eta_p^2 = .11$, 95% CI [.02, .21]. Post hoc Tukey's tests revealed that observers were less able to discriminate between lie- and truth-tellers who were beginner English speakers than those who were native English speakers, $p = .009$, or intermediate English speakers, $p = .001$ (see Table 3). There were no significant differences between any of the other proficiency groups (i.e., $ps > .05$).

One-sample *t*-tests were used to compare observers' discrimination scores to "0" (i.e., no sensitivity). Observers were able to discriminate between truth- and lie-tellers who were native, $t(31) = 3.99$, $p < .001$, $d = .71$, 95% CI [.31, 1.01], advanced, $t(31) = 3.15$, $p = .004$, $d = .56$, 95% CI [.18, .93], or intermediate English speakers, $t(34) = 5.56$, $p < .001$, $d = .94$, 95% CI [.54, 1.33]. However, observers could not discriminate between lie- and truth-tellers who had just begun to learn English, $t(32) = -.04$, $p = .970$, $d = .01$, 95% CI [–.27, −.28]. These findings indicate observers are able to discriminate between lie- and truth-tellers of certain language proficiencies.

Bias. We also conducted a one-way ANOVA on observers' biases. The results revealed that there was no significant difference between proficiency groups (see Table 3), $F(3, 129) = 1.24$, $p = .298$, $\eta_p^2 = .03$, 95% CI [.00, .08]. One-sample *t*-tests were used to compare observers' response bias to "1" (i.e., no bias). There was a significant truth-bias when observers judged native English speakers, $t(31) = 2.22$, $p = .034$, $d = .39$, 95% CI [.03, .75]. However, there was no indication of bias when observers judged advanced English speakers, $t(31) = .16$, $p = .874$, $d = .03$, 95% CI [–.32, .37], intermediate English speakers, $t(34) = 1.69$, $p = .100$, $d = .29$, 95% CI [–.05, .62], or beginner English speakers, $t(32) = 1.92$, $p = .064$, $d = .33$, 95% CI [–.02, .68]. These results suggest that observers only possess a truth-bias toward native English speakers and no biases toward other language proficiencies.

Discussion

In this study, we analyzed the impact of interviewees' language proficiencies on lie detection. Overall, observers' accuracy was slightly higher than chance (i.e., 50%); in addition, they were more accurate when they judged truth-tellers than lie-tellers. These results were consistent with the previous literature (e.g., Bond & DePaulo, 2006).

Deception Detection

We hypothesized that observers' discrimination between lie- and truth-tellers would be poorer when interviewees were beginner speakers than native English speakers, similar when they were intermediate English speakers or native speakers, and better when they were advanced English speakers than native speakers. These hypotheses were partially supported. Observers were worse at making decisions regarding beginner English speakers than interviewees with any other level of language proficiency (i.e., native, advanced, and intermediate English speakers). As predicted, they were also equally able to detect deception in intermediate and native English speakers. Contrary to our hypotheses, however, we found that discrimination was also similar when detecting advanced English speakers' deception. On the whole, we replicated previous findings (i.e., Da Silva & Leach, 2013; Evans & Michael, 2014; Leach & Da Silva, 2013). We consider several explanations for this pattern of results.

According to theories of cognitive load, telling a lie and engaging in a cognitively challenging task should render deception easier to detect (e.g., Vrij et al., 2008). Because speaking in a nonnative language depletes cognitive resources (Abutalebi & Green, 2007; Volk et al., 2014) those interviewees' deception should be more detectable than that of native English speakers. That was not what we found, however. Observers were less able to detect interviewees' deception in the lowest proficiency group (i.e., highest cognitive load) than advanced, intermediate, and native English speakers. Our cue analysis, just like DePaulo et al.'s (2003), revealed only a few cues to deception. DePaulo et al. (2003) posited that lying may not be as cognitively demanding as originally hypothesized given that lies are commonplace in daily

Table 3
Table of Means and Standard Deviations for Overall Accuracy, Discrimination, and Response Bias From Phase 2

	Native		Advanced		Intermediate		Beginner	
Measure	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Overall accuracy	.60*	.14	.57*	.13	.62*	.12	.50	.12
Discrimination (d')	.41*	.59	.27*	.49	.47*	.50	−.00	.51
Response bias (β)	1.24*	.62	1.01	.32	1.15	.53	1.14	.43

* $p < .05$.

interactions. Depending on the context, it may be easier to make up lies than to remember the truth (Fenn, McGuire, Langben, & Blandón-Gitlin, 2015; Walczyk et al., 2005). In our study, lie-tellers reported more difficulty in answering the experimenter's questions than truth-tellers, whereas beginner and intermediate English speakers had more difficulty understanding the experimenter than native English speakers. Although it was originally anticipated that lying and speaking a nonnative language would be doubly taxing on cognitive resources, this does not appear to be true. There were, in fact, few differences in objective and subjective analyses of cognitive load between native and nonnative English speakers. Therefore, cognitive load might be more complex than originally thought.

Two factors may be responsible for nonnative English speakers' emotional involvement: second language anxiety and emotional blunting. Evidence for the Emotional Blunting Hypothesis suggests that speaking in a nonnative language results in lower emotionality and arousal compared to speaking a native language (Anooshian & Hertel, 1994; Ayçiçeği-Dinn & Harris, 2004). If beginner English speakers were more emotionally distant than members of other groups, then that could account for their deception being more difficult to detect in our study and previous research (e.g., Da Silva & Leach, 2013). Our interviewees' self-reports and our exploratory cue analyses did not appear to support that explanation, however. Researchers have found that nonnative speakers report feeling emotions associated with embarrassment (e.g., anxiety, shame; Dewaele, 2007; Woodrow, 2006). Perhaps, the heightened anxiety associated with speaking in a nonnative language might have offset the decrease in emotions typically associated with deception. Indeed, other researchers have failed to find expected differences between nonnative and native speakers during deception using physiological indices of emotion (Caldwell-Harris & Ayçiçeği-Dinn, 2009).

Although we did not examine the cues that observers used to make their decisions, it is plausible that they were relying on misleading information when viewing beginner English speakers. Observers are known to expect lie-tellers to avoid eye contact and self-manipulate (The Global Deception Research Team, 2006). These behaviors have been reported to be more common with nonnative speakers than native speakers (Gregersen, 2005), but we did not find that effect in our study. Perhaps observers' accuracy was impacted by the contrast between their expectations (e.g., gaze-aversion and self-manipulation) and interviewees' manifest behavior (no cue differences between language groups). Also, as in a previous study (Lev-Ari & Keysar, 2012), observers might have expected nonnative speakers to communicate poorly and paid less attention to the details provided. Detail is an important factor in successful lie detection decisions (Vrij, Evans, Akehurst, & Mann, 2004). If observers disregarded beginner English speakers' message content (e.g., details) due to their heavy accents, then their accuracies would suffer; in fact, that is what we found. Intermediate and advanced speakers were less likely to have strong accents and might have been attended to similarly to native English speakers, accounting for the lack of differences in deception detection performance between those groups.

As speakers' accents increase, their intelligibility also decreases (Wasserman, 2008). Bent and Bradlow (2003) revealed that native English listeners perceived native English speakers to be more intelligible than nonnative English speakers. Lev-Ari and Keysar (2010) found that statements made with an accent were perceived as deceptive compared to statements made without an accent because they were more difficult to understand. Although intelligibility was not examined in our study, all of the coders reported having difficulty understanding the beginner English speakers. Observers might have performed poorly when they judged beginner nonnative English speakers, but not others, because they were unable to understand what was being said and resorted to guessing.

Bias

We hypothesized that observers would be more likely to judge native English speakers as truth-tellers than speakers with any nonnative English proficiencies. This hypothesis was supported, replicating Evans and Michael's (2014); Da Silva and Leach's (2013), and Leach and Da Silva's (2013) results. The source of this effect remains unclear. People with accents have been found to be less credible than those without an accent (Brennan & Brennan, 1981) and accents are related to prejudice (Dixon, Mahoney, & Cocks, 2002). Stimuli that are easier to process, such as native speech, are also perceived as more truthful (Lev-Ari & Keysar, 2010; Reber & Schwarz, 1999). As discussed above, additional work is needed to determine whether it is the mere presence of an accent or the associated unintelligibility that eliminates observers' natural truth-bias.

Deception judgments are no different than other types of interactions in which cultural norms and expectations may give rise to biases. In a study examining the impact of culture on deception detection, Vrij and Winkel (1991) found that individuals of diverse cultural backgrounds presented differing cues (i.e., changes in voice and movement). Observers also noted these differences: they were more suspicious of individuals from a culture that was not their own (Vrij & Winkel, 1994). In our study, we controlled for speakers' cultural backgrounds by matching across conditions. Our analysis also revealed that the cultural background of interviewees did not affect observers' responses. However, the interaction between culture and language, and its effects on deception detection, would be an interesting avenue for future research.

Limitations and Future Research

It is important to examine why observers continuously underperform when judging beginner English speakers. We utilized self-report to obtain data about interviewees' emotions and cognitive load. It is possible that the lack of significant findings—particularly in terms of emotion—was due to limitations associated with this methodology (Nisbett & Wilson, 1977). Using more precise tools (e.g., neuroimaging, the polygraph) might have revealed underlying differences between the language proficiency conditions and explained deception detection effects. Future research should, thus, employ more objective assessments of emotion and cognitive load.

Language proficiency is difficult to measure and analyze. Assessment methods differ across studies: self-report (e.g., Castillo et al., 2014), language history questionnaire (e.g., Evans et al., 2013), and standardized tests (e.g., Da Silva & Leach, 2013) have all been employed in previous work. In our study, nonnative speakers' self-reported language proficiencies did not significantly differ between groups. It is, unlikely, however, that this could account for our nonlinear pattern of results. Self-report measurements are known to be unreliable (Nisbett & Wilson, 1977) and we found that they failed to correspond with government approved objective assessments. For this reason, we ultimately decided to categorize nonnative English speakers' proficiency groups based solely on their placement in groups following standardized testing of their language proficiency. Instead of using self-report measures of language proficiency, our findings suggest that future researchers should continue to recruit participants whose language proficiency was assessed using standardized measures.

Subsequent studies should also inquire into observers' processes by examining the cues that they used to make their decisions (i.e., auditory vs. visual stimuli). For example, using a Lens Model (Hartwig & Bond, 2011), researchers may examine whether it is that observers are relying upon incorrect cues to beginner nonnative speakers' deceit (e.g., they are misled by naturalistic behavior associated with speaking in a weak nonnative language) or that there are few behavioral differences exhibited between lie- and truth-tellers.

Our cue analysis did not yield many significant effects. The nonsignificant findings could be due to the small sample size and a lack of power. Indeed, when we calculated post hoc power using our effect sizes, we found that it was extremely low (i.e., 15%). Yet, proficiency effects have been found with relatively few targets in other studies (e.g., Evans et al., 2013). In those studies, researchers employed subjective measures: minimally trained observers identified behaviors as they rendered lie detection decisions. The mere act of judging interviewees as lie- or truth-tellers might have altered the cues that individuals thought that they had observed. Perhaps behavioral differences between lie- and truth-tellers are minute, accounting for an absence of effects when objective measures were employed in our study (see Hartwig & Bond, 2011 for a full discussion). Cue analysis was not the primary goal of this study, however. Further research is, therefore, needed to determine why proficiency effects on cues to deception are uneven.

In addition, researchers could examine observers' expectations and comprehension of nonnative speakers. Specifically, they could probe for the presence of a stereotype that nonnative speakers communicate poorly (e.g., Lev-Ari & Keysar, 2012). Observers should be questioned about their previous experiences interacting with nonnative speakers and their related expectations. Tsurutani and Selvanathan (2013) found that nonnative speakers' accents made a negative impact on observers' perceptions but that this effect was mitigated by observers' previous contact (i.e., experience) with nonnative speakers. Thus, the impact of observers' preconceived notions and experiences prior to the study may serve to explain their deception judgments.

Observers could also be asked about speakers' intelligibility. That is, researchers could assess the ease with which observers were able to comprehend the interviewees. Presumably, if observers were unable to understand nonnative speakers, then their decision-making processes would be negatively impacted. The relative contributions of intelligibility and accent on lie detection decisions may also be analyzed using alternate modalities (e.g., having observers read transcripts vs. listen to audio).

Implications

Our results revealed that observers were unable to discern between the lies and truths of beginner English speakers. Given that the sample of interviewees was recruited from immigrant centers for adults and students, they all entered the country through a border and would have encountered a border services officer. Our results suggest that these interactions may have been difficult for the officers—unless interpreters were used or the agents spoke the interviewees' native languages—because of the demonstrated difficulties discerning truthfulness. Following debriefing, many beginner English speakers (and other nonnative speakers) spoke about their experiences at airports. Interviewees reported being apprehensive about their arrivals to a new country, which was only made worse when they were met with suspicion and thoroughly questioned regarding their intentions. They recalled repeating their rehearsed itineraries and not being able to understand the officers' questions. Thus, these findings have real implications for borders around the world, as well as other law enforcement contexts in which officers communicate with nonnative speakers.

There is an urgent need to conduct further research to be able to advise law enforcement agencies of best practices with regards to interviewing nonnative speakers. Although the use of an interpreter was once advisable, emerging research suggests that it might hamper deception detection: nonnative English speakers provide less details and show fewer cues to deception when interpreters are used (Ewens et al., 2016). Thus, officers should make every attempt to communicate in the interviewee's native language so that neither the interviewee nor the investigation are placed at risk.

Conclusions

We examined the impact of varying language proficiencies on deception detection accuracy, bias, discrimination, and cues to deception. Overall, observers' accuracy was poorest when judging beginner English speakers. Moreover, nonnative speakers were perceived more negatively than native English speakers. The results of this study have implications for the safety and security of borders and individuals and, thus, more research is needed to establish best practices when interviewing nonnative speakers.

References

Abe, N., Suzuki, M., Mori, E., Itoh, M., & Fujii, T. (2007). Deceiving others: Distinct neural responses of the prefrontal cortex and amygdala in simple fabrication and deception with social interactions. *Journal of Cognitive Neuroscience, 19,* 287–295. http://dx.doi.org/10.1162/jocn.2007.19.2.287

Abutalebi, J., & Green, D. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics, 20,* 242–275. http://dx.doi.org/10.1016/j.jneuroling.2006.10.003

Adesope, O. O., Lavin, T., Thompson, T., & Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Review of Educational Research, 80,* 207–245. http://dx.doi.org/10.3102/0034654310368803

Anooshian, L. J., & Hertel, P. T. (1994). Emotionality in free recall: Language specificity in bilingual memory. *Cognition and Emotion, 8,* 503–514. http://dx.doi.org/10.1080/02699939408408956

Ayçiçeg˘i-Dinn, A., & Harris, C. L. (2004). Bilinguals' recall and recognition of emotion words. *Cognition and emotion, 18,* 977–987. http://dx.doi.org/10.1080/02699930341000301

Bent, T., & Bradlow, A. R. (2003). The interlanguage speech intelligibility benefit. *The Journal of the Acoustical Society of America, 114,* 1600–1610. http://dx.doi.org/10.1121/1.1603234

Bond, C. F., & Depaulo, B. M. (2008). Individual differences in judging deception: Accuracy and bias. *Psychological Bulletin, 134,* 477–492. http://dx.doi.org/10.1037/0033-2909.134.4.477

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

Bond, M. H., & Lai, T. M. (1986). Embarrassment and code-switching into a second language. *The Journal of Social Psychology, 126,* 179–186.

Brennan, E., & Brennan, J. (1981). Accent scaling and language attitudes: Reactions to Mexican American English speech. *Language and Speech, 24,* 207–221.

Buller, D. B., & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory, 6,* 203–242. http://dx.doi.org/10.1111/j.1468-2885.1996.tb00127.x

Caldwell-Harris, C. L., & Ayçiçeg˘i-Dinn, A. (2009). Emotion and lying in a non-native language. *International Journal of Psychophysiology, 71,* 193–204. http://dx.doi.org/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,* 66–82. Retrieved from http://dev.cjcenter.org/_files/apcj/apcjspring2014castillo.pdf_1399047724.pdf

Centre for Canadian Language Benchmarks. (2012). *Canadian language benchmarks: English as a second language for adults*. Retrieved from http://www.cic.gc.ca/english/pdf/pub/language-benchmarks.pdf

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,* 107–118. http://dx.doi.org/10.1375/pplt.2005.12.1.107

Chiu, C., Hong, Y., & Krauss, R. M. (2001). *Gaze direction and fluency in conversational speech*. Unpublished manuscript. Retrieved from http://www.columbia.edu/~rmk7/PDF/Gaze.pdf

Christ, S. E., Van Essen, D. C., Watson, J. M., Brubaker, L. E., & McDermott, K. B. (2009). The contributions of prefrontal cortex and executive control to deception: Evidence from activation likelihood estimate meta-analyses. *Cerebral Cortex, 19,* 1557–1566. http://dx.doi.org/10.1093/cercor/bhn189

Da Silva, C. S., & Leach, A. M. (2013). Detecting deception in second-language speakers. *Legal and Criminological Psychology, 18,* 115–128. http://dx.doi.org/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. *American Psychological Association, 129,* 74–118. http://dx.doi.org/10.1037/0033-2909.129.1.74

Dewaele, J.-C. (2007). Predicting language learners' grades in the L1, L2, L3, and L4: The effect of some psychological and sociocognitive variables. *International Journal of Multilingualism, 4,* 169–197. http://dx.doi.org/10.2167/ijm080.0

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,* 162–168. http://dx.doi.org/10.1177/02627X02021002004

Duñabeitia, J. A., & Costa, A. (2015). Lying in a native and foreign language. *Psychonomic Bulletin & Review, 22,* 1124–1129. http://dx.doi.org/10.3758/s13423-014-0781-4

Ekman, P. (1989). Why lies fail and what behaviors betray a lie. In J. C. Yuille (Ed.), *Credibility assessment* (pp. 71–82). Dordrecht, Holland: Kluwer. http://dx.doi.org/10.1007/978-94-015-7856-1_4

Evans, J. R., & Michael, S. W. (2014). Detecting deception in non-native English speakers. *Applied Cognitive Psychology, 28,* 226–237. http://dx.doi.org/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 & Cognition, 2,* 33–41. http://dx.doi.org/10.1016/j.jarmac.2013.02.002

Ewens, S., Vrij, A., Leal, S., Mann, S., Jo, E., & Fisher, R. (2016). The effect of interpreters on eliciting information, cues to deceit and rapport. *Legal and Criminological Psychology, 21,* 286–304.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39,* 175–191. http://dx.doi.org/10.3758/BF03193146

Fenn, E., McGuire, M., Langben, S., & Blandón-Gitlin, I. (2015). A reverse order interview does not aid deception detection regarding intentions. *Frontiers in Psychology, 6,* 1–12. http://dx.doi.org/10.3389/fpsyg.2015.01298

Ganis, G., Kosslyn, S. M., Stose, S., Thompson, W. L., & Yurgelun-Todd, D. A. (2003). Neural correlates of different types of deception: An fMRI investigation. *Cerebral Cortex, 13,* 830–836. http://dx.doi.org/10.1093/cercor/13.8.830

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York, NY: Wiley.

Gregersen, T. S. (2005). Nonverbal cues: Clues to the detection of foreign language anxiety. *Foreign Language Annals, 38,* 388–400. http://dx.doi.org/10.1111/j.1944-9720.2005.tb02225.x

Gumperz, J. J., & Hymes, D. H. (1986). *Directions in Sociolinguistics: The Ethnography of Communication*. Oxford, United Kingdom: Basil Blackwell.

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

Hyönä, J., Tommola, J., & Alaja, A. M. (1995). Pupil dilation as a measure of processing load in simultaneous interpretation and other language tasks. *Quarterly Journal of Experimental Psychology, 48,* 598–612. http://dx.doi.org/10.1080/14640749508401407

Ivanova, I., & Costa, A. (2008). Does bilingualism hamper lexical access in speech production? *Acta Psychologica, 127,* 277–288. http://dx.doi.org/10.1016/j.actpsy.2007.06.003

Jarvis, B. G. (2014). *MediaLab (Version 2014.1.103)* [Computer Software]. New York, NY: Empirisoft Corporation.

Keysar, B., Hayakawa, S. L., & An, S. G. (2012). The foreign-language effect: Thinking in a foreign tongue reduces decision biases. *Psychological Science, 23,* 661–668. http://dx.doi.org/10.1177/0956797611432178

Lazar, J. N., Stern, A., & Cohen, R. (2014). Decision-making in foreign language reduces emotional arousal. *Psychology, 5,* 2180–2188.

Leach, A. M., & Da Silva, C. S. (2013). Language proficiency and police officers' lie detection performance. *Journal of Police and Criminal Psychology, 28,* 48–53. http://dx.doi.org/10.1007/s11896-012-9109-3

Lee, T. M., Liu, H. L., Chan, C. C., Ng, Y. B., Fox, P. T., & Gao, J. H. (2005). Neural correlates of feigned memory impairment. *NeuroImage, 28,* 305–313. http://dx.doi.org/10.1016/j.neuroimage.2005.06.051

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,* 1093–1096. http://dx.doi.org/10.1016/j.jesp.2010.05.025

Lev-Ari, S., & Keysar, B. (2012). Less-detailed representation of non-native language: Why non-native speakers' stories seem more vague. *Discourse Processes, 49,* 523–538. http://dx.doi.org/10.1080/0163853X.2012.698493

Lykken, D. T. (1974). Psychology and the lie detector industry. *American Psychologist, 29,* 725–739. http://dx.doi.org/10.1037/h0037441

Mann, S., Vrij, A., & Bull, R. (2002). Suspects, lies, and videotape: An analysis of authentic high-stake liars. *Law and Human Behavior, 26,* 365–376. http://dx.doi.org/10.1023/A:1015332606792

Marcos, L. R. (1976). Bilinguals in psychotherapy: Language as an emotional barrier. *American Journal of Psychotherapy, 30,* 552–560.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84,* 231–259. http://dx.doi.org/10.1037/0033-295X.84.3.231

Pavlenko, A. (2006). *Emotions and multilingualism*. New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511584305

Perani, D., & Abutalebi, J. (2005). The neural basis of first and second language processing. *Current Opinion in Neurobiology, 15,* 202–206. http://dx.doi.org/10.1016/j.conb.2005.03.007

Perani, D., Abutalebi, J., Paulesu, E., Brambati, S., Scifo, P., Cappa, S. F., & Fazio, F. (2003). The role of age of acquisition and language usage in early, high-proficient bilinguals: An fMRI study during verbal fluency. *Human Brain Mapping, 19,* 170–182. http://dx.doi.org/10.1002/hbm.10110

Ransdell, S. E., & Fischler, I. (1987). Memory in a monolingual mode: When are bilinguals at a disadvantage? *Journal of Memory and Language, 26,* 392–405. http://dx.doi.org/10.1016/0749-596X(87)90098-2

Reber, R., & Schwarz, N. (1999). Effects of perceptual fluency on judgments of truth. *Consciousness and Cognition, 8,* 338–342. http://dx.doi.org/10.1006/ccog.1999.0386

Ryan, C. (2013). Language use in the United States: 2011 (Report No. ACS-22). Retrieved from http://www.census.gov/content/dam/Census/library/publications/2013/acs/acs-22.pdf

Spence, S. A., Farrow, T. F. D., Herford, A. E., Wilkinson, I. D., Zheng, Y., & Woodruff, P. W. R. (2001). Behavioural and functional anatomical correlates of deception in humans. *Neuoreport: For Rapid Communication of Neuroscience Research, 12,* 2849–2853. http://dx.doi.org/10.1097/00001756-200109170-00019

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

Sporer, S. L., & Schwandt, B. (2007). Moderators of nonverbal indicators of deception: A meta-analytic synthesis. *Psychology, Public Policy, and Law, 13,* 1–34. http://dx.doi.org/10.1037/1076-8971.13.1.1

The Global Deception Research Team. (2006). A world of lies. *Journal of Cross-Cultural Psychology, 37,* 60–74. http://dx.doi.org/10.1177/0022022105282295

Tsurutani, C., & Selvanathan, E. (2013). Influence of generational cohort and experience with non-native speakers on evaluation of speakers with foreign-accented speech. *International Journal of the Sociology of Language, 224,* 43–62.

Van Compernolle, D. (2001). Recognizing speech of goats, wolves, sheep, and . . . non-natives. *Speech Communication, 35,* 71–79. http://dx.doi.org/10.1016/S0167-6393(00)00096-0

Volk, S., Köhler, T., & Pudelko, M. (2014). Brain drain: The cognitive neuroscience of foreign language processing in multinational corporations. *Journal of International Business Studies, 45,* 862–885. http://dx.doi.org/10.1057/jibs.2014.26

Vrij, A. (1995). Behavioural correlates of deception in a simulated police interview. *The Journal of Psychology, 129,* 15–28. http://dx.doi.org/10.1080/00223980.1995.9914944

Vrij, A., Evans, H., Akehurst, L., & Mann, S. (2004). Rapid judgements in assessing verbal and nonverbal cues: Their potential for deception researchers and lie detection. *Applied Cognitive Psychology, 18,* 283–296. http://dx.doi.org/10.1002/acp.964

Vrij, A., Fisher, R., Mann, S., & Leal, S. (2008). A cognitive load approach to lie detection. *Journal of Investigative Psychology and Offender Profiling, 5,* 39–43. http://dx.doi.org/10.1002/jip.82

Vrij, A., Mann, S., Leal, S., & Fisher, R. (2010). "Look into my eyes": Can an instruction to maintain eye contact facilitate lie detection? *Psychology, Crime & Law, 16,* 327–348. http://dx.doi.org/10.1080/10683160902740633

Vrij, A., & Winkel, F. W. (1991). Cultural patterns in Dutch and Surinam nonverbal behaviour: An analysis of simulated police/citizen encounters. *Journal of Nonverbal Behavior, 15,* 169–184. http://dx.doi.org/10.1007/BF01672219

Vrij, A., & Winkel, F. W. (1994). Perceptual distortions in cross-cultural interrogations: The impact of skin color, accent, speech style, and spoken fluency on impression formation. *Journal of Cross-Cultural Psychology, 25,* 284–295. http://dx.doi.org/10.1177/0022022194252008

Walczyk, J. J., Schwartz, J. P., Clifton, R., Adams, B., Wei, M., & Zha, P. (2005). Lying person-to-person about life events: A cognitive framework for lie detection. *Personnel Psychology, 58,* 141–170. http://dx.doi.org/10.1111/j.1744-6570.2005.00484.x

Wasserman, S. J. (2008). *Accent and intelligibility: An investigation of self-perceived communication abilities of non-native English speakers* (Master's thesis). Available from ProQuest Dissertations and Theses database. (UMI No. 1450518)

Wixted, J. T., & Lee, K. (n.d.). *Signal detection theory* [Excel worksheet for SDT calculations]. Retrieved from http://www.kangleelab.com/signal%20detection%20theory.html

Woodrow, L. (2006). Anxiety and speaking English as a second language. *RELC Journal, 37,* 308–328. http://dx.doi.org/10.1177/0033688206071315

Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). In L. Berkowitz (Ed.), *Verbal and nonverbal communication of deception* (Vol. 15, pp. 1–57). *Advances in experimental social psychology*. New York, NY: Academic Press.

Received February 18, 2016
Revision received August 2, 2016
Accepted August 16, 2016 ■

Correction to Elliott and Leach (2016)

In the article "You Must Be Lying Because I Don't Understand You: Language Proficiency and Lie Detection," by Elizabeth Elliott and Amy-May Leach (*Journal of Experimental Psychology: Applied*, Vol. 22, No. 4, pp. 488–499. http://dx.doi.org/10.1037/xap0000102), in the Results section, under "Signal detection theory," the first sentence of the second paragraph should read as follows: "An a priori decision was made to use Wixted and Lee's (n.d.) formulas to calculate discrimination (i.e., $d' = Z_{FA} - Z_{Hit}$) and bias (i.e., $\beta = e^{\left\{\frac{[z_{FA}]^2 - [z_{Hit}]^2}{2}\right\}}$).

http://dx.doi.org/10.1037/xap0000121