AN INVESTIGATION OF ACCURACY AND BIAS IN CROSS-CULTURAL LIE DETECTION

Paola A. Castillo, Graham Tyson and David Mallard
Charles Sturt University

The detection of deception has been relatively unexplored from a cross-cultural standpoint. In such context, the observers and communicators may have different cultural norms and expectations of behavior. These differences may create the potential for bias and errors. To investigate this, 71 Australian students (20 males and 51 females), with a mean age of 21.3 years (*SD* = 5.98), were asked to make credibility judgments of 24 video clips in which culture and language use (first vs second) were manipulated. It was found that participants were generally poor not only at classifying lies and truths within cultures, but also across cultures. Although the results indicated that the language spoken had no impact on observers' ability to discriminate between lies and truths, observers were generally more suspicious of Colombian clips and, in particular, of those that depicted Colombians speaking in a second language. Practical and theoretical implications are discussed.

Keywords: deception, lying, cross-cultural deception, cultural

Deception researchers have long sought explanations of the process by which people make credibility judgments. One of the most prominent explanations suggests that people make inferences of deception based on norms and expectancy violations of behavior. According to the expectancy violation model (Bond et al., 1992), when communicating with others people have pre-existing expectancies regarding the other person's behavior. These expectancies stem from social norms and denote "a range of acceptable behavior" (Levine et al., 2000, p. 124). A violation of social norms is likely to occur when an individual's behavior falls outside the range of acceptable behavior and is deemed deviant or inappropriate. Thus, when norm violations occur, expectancy violations also occur.

These expectancy violations would then demand an explanation and raise suspicion of the person, which may result in arousal and "finer-grain information processing" (Burgoon & Ebesu Hubbard, 2005, p. 153). An evaluation process is then initiated in order to "make sense" of this violation. This evaluation process may result in a non-specific activation of judgments of deception as a possible cause for the violation. Causal attributions of these violations are then made and, in the absence of a plausible alternative explanation, perceivers may suspect that unexpected behaviors may indicate that the communicator is lying. Thus, deception might be inferred from any behavior that violates a social norm. For example, if the norm for a social interaction includes relatively high levels of eye contact, a person who avoids eye contact may be suspected of deception as a result of violating that

Author Notes: Correspondence concerning this article should be addressed to Paola Castillo, School of Psychology, Charles Sturt University, 68 Clarence Street Port Macquarie Campus, NSW 2444, Australia. E-mail: pcastillo@csu.edu.au.

norm. Furthermore, the more frequently the person displays behavior that violates a given social norm, the more dishonest or deceptive the person is likely to appear. Support for this model was found by Bond and colleagues (1992) which showed that individuals who performed unexpected and strange behaviors (e.g., head tilting and staring) were perceived as more deceptive compared to those that did not violate the observer's expectancies.

However, Levine et al. (2000) argued that Bond et al.'s (1992) expectancy violation model confounded norm and expectancies violations. They asserted that even though norms may create expectations, they are conceptually and empirically distinct from expectations. For instance, polite behavior, depicted by someone who is expected to be rude, would violate expectations but would not violate social norms. Therefore, Levine et al. (2000) proposed the norm violation model, which postulates that norm violations are evaluated independently of expectations, and that norm violations are inappropriate whether they are anticipated or unexpected. According to the model, someone who violates a norm will be viewed as less credible regardless of whether the norm violation was expected or not. Consistent with their model, Levine et al. (2000) found that those who displayed unusual behaviors were seen as less credible regardless of whether the observers were primed or not to expect those norm violations. Current literature (Vrij, Granhag, & Porter, 2010) suggest that these violations may lead to errors, but to date, no research has examined the potential and extent of these errors in cross-cultural contexts.

The notion that individuals might attribute dishonesty to people whose behavior violates a social norm has significant implications for cross-cultural lie detection. Cultures have different behavioral norms and patterns that stem from the nature of culture itself (e.g., Ekman, 1972; Matsumoto, 1989, 1991; Matsumoto et al., 2002; Matsumoto & Kudoh, 1987). Culture not only defines the coding and encoding of verbal and nonverbal messages but also the interpretation and meaning that is given to these behaviors (Ekman & Friesen, 1971). For instance, Matsumoto, Seung Hee and Fontaine (2008) found that individualistic and collectivistic cultures hold different display rules and norms concerning the expressivity of positive and negative emotions.

Behavioral differences that stem from the nature of culture itself may then have the potential to introduce biases when cross-cultural judgments of deception are made. For instance, in a situation in which a communicator and an observer are from different cultures, the observer will apply social norms concerning behavior that may differ from the communicator's own norms. If observers attribute deception based on norm or expectancy violations, a person who is behaving consistently with their cultural norms but inadvertently violates the norms of the observer may be more likely to be judged as deceptive.

Limited deception research has examined the impact of norm and expectancy violations in cross-cultural judgments of credibility. Vrij and Winkel (1991, 1992, 1994) conducted a series of studies to examine whether normative behavior had an effect on perceptions of credibility. In their initial study Vrij and Winkel (1991) found that Black Surinamese participants made more speech errors, spoke more slowly, smiled more, averted their gaze more, spoke with a higher pitched voice, and performed more self-manip-

ulations compared to native Dutch participants regardless of whether they were lying or not. In their subsequent studies (Vrij & Winkel, 1992, 1994) Surinamese and Dutch actors were videotaped while giving a statement and displaying gestures and smiling behavior of typical white (Dutch) or typical black (Surinam). White Dutch police officers were then asked to judge the credibility of these videotapes. They found that both Surinamese and Dutch actors made a more suspicious impression when they displayed behavior that was inconsistent with the observer's norms (i.e., Surinamese norms) than when they displayed consistent normative behavior (i.e., Dutch behavior). These findings highlight the possibility that during cross-cultural credibility judgments, the behavioral patterns of one culture may be vulnerable to being misinterpreted as attempts to hide the truth by the other culture. However, Vrij and Winkel's studies did not examine whether these normative differences between cultures actually result in deception bias. Accordingly, the first aim of the present study was to investigate whether behavioral differences in norms between cultures would result in misinterpretations of a culture's baseline behavior and, consequently, in more dishonest judgments being made (i.e., deception bias).

Language bias

While behavioral differences may increase the potential for bias, language differences between interactants may also play a significant role. Frumkin (2007) argued that characteristics of language and speech provide ways for receivers to judge, form opinions and determine believability of the communicator because these features often trigger stereotypes. Therefore, a communicator's behavior and receiver's evaluations may be affected by characteristics of speech and language. An important feature of cross-cultural interactions is that one participant may be required to communicate in a second or non-native language. According to Spencer-Rodgers and McGovern (2002), this often results in differences in communication styles and variations in norms and customs between interactants, which may result in more errors and miscommunications.

However, limited deception research has been carried out examining the impact of second language use on deception and lie detection. One such study was conducted by Cheng and Broadhurst (2005). They videotaped 31 students from a university in Hong Kong who were randomly assigned to lie or tell the truth about their opinion of capital punishment, and who were interviewed in their mother tongue (i.e., Cantonese) or their second language (i.e., English). Twenty-seven other students then viewed a subset of these interviews and judged whether the person depicted in each video was lying or telling the truth. Overall, they found that observers were more successful in identifying liars speaking in their second language (English) than liars speaking in their native language (Cantonese), but were more successful in identifying truth-tellers speaking in Cantonese than in English. These findings suggest a language bias, such that people speaking in their first language are more likely to be judged as credible compared to people speaking in their second language, irrespective of veracity. More importantly, they found that participants speaking in English tended to display more nonverbal movements indicating nervousness and anxiety (increased gaze aversion, hand and arm movements, and leg and foot movements) compared to those speaking Cantonese.

The perceived indicators of deception literature suggests that lie detectors often look for behavioral signs which indicate that the person is thinking hard (i.e., cognitive load), is anxious or nervous (e.g., DePaulo et al., 2003; Global Deception Research Team, 2006). One reason for this is because these behaviors are often associated with the cognitive demands of lying. Therefore, if observers attribute deception based on the presence of behavioral cues that suggest cognitive load, anxiety or nervousness, an individual speaking in a second language and who is displaying signs of cognitive load that stem from language-based cognitive demands (and not message veracity) may be more likely to be judged as deceptive than someone who is speaking in their first language. It is plausible, then, that the display of these language-dependent behaviors may increase the likelihood of the observer attributing deception as an explanation for 'unusual' behavior. Accordingly, the second aim of the present study was to examine whether second language use results in deception bias.

The present study

In order to achieve the aims of this study, Australian participants were asked to watch a number of video clips that depicted Australians either lying or telling the truth and Colombians either lying or telling the truth in their first and second language (i.e., Spanish) about a mock computer crime. Observers' lie/truth classification judgments were analyzed using a signal detection framework. Such a framework allows for a more comprehensive understanding of credibility judgments because it provides a measure of discrimination (*d'*) between two groups of items (i.e., signal and noise trials), and also provides information about response bias (*c*)[1] (Stanislaw & Todorov, 1999). Discrimination accuracy refers to participants' sensitivity to correctly detecting a "signal" (i.e., in the present study's context, deceit) when it is presented and correctly rejecting "noise" (i.e., truth) when the signal is not presented (Green & Swets, 1966). Response bias (*c*) estimates the extent to which observers are biased to ascribe more "deceitful" than "truthful" judgments.

It was expected that Australian participants would show a cultural bias in that they would ascribe more deception judgments to Colombian video clips than Australian video clips. Similarly, it was expected that Australian participants would show a language bias and would ascribe more dishonest judgments to those video clips that depict Colombians speaking in their second language (English) than when the clips depict Colombians speaking in their first language (Spanish). Furthermore, discriminability would be around chance levels for all video clip types, as suggested by the literature on accuracy of deception judgments (Bond & DePaulo, 2006).

METHOD

Participants

A total of 71 Australian students (20 males and 51 females) with a mean age of 21.3 years (*SD* = 5.98) participated in this study as part of a course requirement.

Materials

Credibility Judgment task. This task comprised of 24 videos clips, a lie/truth judgment task, confidence ratings, and a trustworthiness scale.

The video clips used in this task were recorded as part of the principal investigator's doctoral research program. Of these videos clips, 24 were selected[1] for the present study which depicted Australians and Colombians either lying or telling the truth about a mock computer crime. Specifically, the video clips depicted four Australians lying and four telling the truth (four male and four female) in English; four Colombians lying and four telling the truth (four male and four female) in Spanish; and four Colombians lying and four telling the truth (four male and four female) in English. The duration of the video clips ranged from 45 seconds up to 2 minutes. The video clips were also edited so that each video clip would have a sound and a no sound version. The audio of the clips was manipulated between subjects to ensure that response bias and discriminability were the direct result of non-verbal cues (i.e., no audio), or a combination of verbal and non-verbal cues (i.e., audio).

After watching each video clip, participants were asked to complete the lie/truth judgments task where they indicated whether they thought the person depicted in the video was either lying or telling the truth (dichotomous answer). They were then asked to rate their confidence of their answer on a 5 point rating scale (1 = *not confident* to 5 = *very confident*).

The trustworthiness scale from the ethos/credibility measure developed by McCroskey and Teven (1999) was used in this study. The scale is a 7 point rating scale consisting of six bipolar adjective items: honest/dishonest, trustworthy/untrustworthy, moral/immoral, honourable/dishonourable, ethical/unethical, genuine/phoney. The trustworthiness scale has been shown to be internally reliable (Cronbach's alpha of .92) and to correlate highly $r = .92$ with an overall credibility score.

Inquisit Computer Software (Version 3.0). The credibility judgment task and trustworthiness scale were presented using the Inquisit program ("Inquisit 3.0," 2007). The software recorded participant's responses for each video clip and their responses on the trustworthiness scale. The program also automatically randomized the audio manipulation, the gender and veracity of the video clips.

Procedure

Participants were recruited by posting advertisements on the Psychology Department notice board. Those participants who were enrolled in a first year psychology subject were given credit for their participation.

The experimental sessions were conducted in groups of approximately six participants. When they arrived, participants were asked to sit in front of a computer and complete the informed consent procedures. The credibility judgment task was then explained

1 The video clips selected did not differ significantly in terms of attractiveness, friendliness, interestingness, and likeability, characteristics thought to affect attributions of deception.

to them and questions were answered. Once ready to begin, participants were asked to read further instructions in the computer screen and commence the tasks in their own time.

RESULTS

Participants' lie and truth classifications for each clip were analyzed using a signal detection framework, in which the 'signal' was the occurrence of a lie. Table 1 shows the proportion of hits (i.e., deceptive messages correctly classified as deceptive) and false alarms (i.e., truthful messages incorrectly classified as deceptive) as well as the overall accuracy rate in classifying lies and truths for each of the three types of clips.

As can be seen, participants were poor at correctly classifying lies and truths across all clips. Participants' overall accuracy rates were only slightly above the level of chance, and participants had a higher proportion of incorrect classifications than correct ones. A signal detection analysis was conducted to examine whether participants' truth and lie accuracy rates were significantly different across clips and conditions.

Table 1

Proportion of classification rates across video clips

	Clips		
	Australian Clips	Colombian English Clips	Colombian Spanish Clips
Hits	.25(0.26)	.58 (0.29)	.40 (0.24)
False alarms	.49 (0.24)	.58 (0.22)	.33 (0.24)
Overall accuracy rate	.38 (0.17)	.50 (0.16)	.54 (0.17)

Note: standard deviation shown in parentheses.

Signal detection analysis

In order to analyze participants' lie detection performance, signal detection statistics were calculated. The statistic *d'* provides a measure of discriminability – which refers to the extent to which the participant could discriminate between 'signal' trials (lies) and 'no signal' trials (truths). If the rate at which a participant classified clips as being lies was the same for clips that were actually lies as for clips that were truths – in other words, if the hit rate and the false alarm rate were similar to one another – then discriminability would be poor and *d'* would tend towards zero. On the other hand, if a participant tended to classify as lies a greater proportion of the clips that actually were lies than they did clips that actually were truths – in other words, if the hit rate was higher than the false alarm rate – then discriminability would be good and *d'* would tend to have a positive value. The greater the disparity between hit and false alarm rates, the larger *d'* would become.

The statistic *c* provides a measure of response bias, that is, the general tendency of the participant to classify trials as 'signal' trials (lies) or as 'no signal' trials (truths), inde-

pendent of the actual presence or absence of the signal. If a participant tended to classify more clips as lies than as truths (a 'lie bias'), then the value of *c* would be negative, whereas if a participant tended to classify more clips as truths than lies (a 'truth bias'), the value of *c* would be positive. The magnitude of *c* indicates the degree of response bias, with a value of zero indicating a neutral approach to judgment.

The signal detection statistics for each combination of clip and audio condition are presented in Table 2.

Table 2

Mean (SD) rates of discriminability and response bias across clips and audio conditions

	Clips		
Audio Condition	Australian Clips	Colombian English Clips	Colombian Spanish Clips
Audio			
Discriminability (d')	-0.43 (0.86)	0.17 (0.66)	0.16 (0.80)
Response bias (c)	0.36 (0.44)	-0.17 (0.53)	0.38 (0.43)
No Audio			
Discriminability (d')	-0.73 (0.78)	-0.16 (0.86)	0.16 (0.77)
Response bias (c)	0.27 (0.41)	-0.21 (0.45)	0.24 (0.41)

To examine whether participants' discrimination between lies and truths or response bias differed across clip and audio conditions, a series of planned contrast analyses[2] were performed based on a 2 (audio, no audio) x (3) (Australian English, Colombian English, Colombian Spanish) mixed model design. The same contrast analyses were applied to each of the dependent variables in turn: discriminability (*d'*) and response bias (*c*).

For each dependent variable, contrast coefficients of clip condition were defined to test (i) whether participants' results for the dependent variable differed between Australian video clips and Colombian video clips (contrast coefficients: 2, -1, and -1 for the Australian-English, Colombian-Spanish and Colombian-English clip conditions respectively) and (ii) whether participants' results for the dependent variable differed between Colombian video clips in English and Colombian video clips in Spanish (contrast coefficients: 0, 1, -1). Thus, the two contrasts tested for differences in lie detection judgments based on the culture of the person depicted in the video clip and, where the person was of a different culture, whether that person was speaking in his or her first or second language. As a dichotomous independent variable, the main effect of audio condition was examined as a

2 Given the theoretical direction of our predictions a series of planned contrast analyses were considered more appropriate than a post hoc analysis of variance. Wilcox (1987) argued that the "procedure chosen depends on what the researcher wants to know and the type I error that is allowed" (p.36).

between-subjects pairwise comparison. Therefore, for each set of planned contrasts on clip type (i.e., across nationalities and language use), the corresponding interaction between that contrast and audio condition was examined.

Bonferroni adjustments were used in order to keep the family wise Type I error rate at .05 (i.e., an alpha of .025 was used) for each family of two contrasts. All contrast analyses were conducted using the computer program PSY developed by Bird, Hadzi-Pavlovic, and Isaac (2000). Effect sizes (*d)* for each contrast analysis were also calculated.

Discriminability (d')

The analysis of the first contrast comparing discriminability across clip condition showed that participants' overall ability to correctly discriminate between liars and truth tellers varied significantly as a function of culture, F (1, 69) = 30.94, $p < .025$, and that this effect was large, d= 0.84. On average, participants' ability to discriminate between lies and truths was considerably poorer for clips depicting Australians than for clips depicting Colombians. However, it should be noted that this difference appears to result from participants performing below chance levels (mean d' = -0.58) with Australian clips, whereas the overall discriminability for Colombian clips remained close to chance levels (mean d' = 0.08).

The analysis of the second contrast comparing discriminability across clip conditions showed that participants' discrimination abilities did not significantly differ across language conditions within the Colombian clips, F (1, 69) = 1.56, $p > .025$. That is, participants' capacity to discriminate between the lies and truths told by Colombians did not appear to depend on whether the person was speaking in his or her first or second language. The main effect of audio condition was also not statistically significant, F (1, 69) = 3.79, $p > .025$ (d= 0.27), nor were the audio condition's interactions with the clip condition contrasts statistically significant (Fs ≤ 1.71, ps > .025, $d \leq 0.4$), indicating that the presence or absence of verbal cues did not appear to influence participants' ability to discriminate between lies and truths in the clips.

Response bias (c)

The first contrast analysis comparing response bias across clip conditions indicated that participants had a greater truth bias (positive c) when judging Australian than Colombian clips, F (1, 69) = 24.70, $p < .025$, and that the effect was medium d = 0.58. Overall, as shown in Table 2, participants were likely to ascribe more truthful than deceptive judgments to the Australian clips than to the Colombian clips (mean c = 0.32, and mean c = 0.06 respectively).

The analysis of the second contrast comparing response bias across language conditions suggested that participants showed a lie bias (negative c) for the Colombian – English clips, whereas for the Colombian – Spanish clips they showed a truth bias, F (1, 69) = 48.72, $p < .025$, and the effect was large d = -1.12. The results of both contrast analyses (language and culture) suggest that even though there was a significant difference in response bias across clip conditions this may have been largely due to the lie bias shown

in the Colombian – English video clips. The Colombian – Spanish video clips actually showed a similar truth bias to that seen in the Australian – English video clips.

The main effect of audio condition was not statistically significant, F (1, 69) = 1.73, $p > .025$ (d= 0.22*)*, nor were audio condition's interactions with the clip condition contrasts statistically significant (Fs ≤ 0.49, ps > .025), indicating the presence or absence of verbal and nonverbal cues did not seem to influence their tendency to classify lies and truths ($d \leq 0.22$*)*.

Trustworthiness

A reliability analysis indicated that the trustworthiness scale was internally consistent, (Cronbach's alpha of .97). Therefore, scores for the six items were combined to calculate a trustworthiness score for each participant.

In order to assess whether lying and truth telling videos were rated differently in terms of the perceived trustworthiness of the person depicted in the video clip, a series of planned contrast analyses were performed using 2 (audio, no audio) x (2)(lies, truths) x (3) (Australian – English, Colombian – English, Colombian – Spanish) mixed model design. Table 3 shows the means and standard deviations for the trustworthiness scale across a combination of video clips, veracity, and audio conditions. Higher scores indicated that participants perceived the person in the video clip as more trustworthy.

Table 3

Means and standard deviations for trustworthiness ratings across clips, veracity, and audio conditions

	Clips		
Condition	Australian clips	Colombian Spanish clips	Colombian English clips
Audio			
Lying	29.35 (5.26)	27.15 (4.30)	21.65 (4.75)
Truth-telling	25.00 (4.28)	28.51 (4.16)	23.10 (4.17)
No Audio			
Lying	29.00 (4.96)	25.27 (3.67)	22.44 (5.20)
Truth -telling	23.29 (3.89)	26.97 (4.77)	22.84 (4.53)

Note: Higher scores indicate that participants perceived the actor as trustworthy.

Contrast coefficients of clip and veracity conditions were defined to test (i) whether trustworthiness ratings differed for truths versus lies across all clips (contrast coefficients: -1, 1, -1, 1, -1, 1, lying Australian clips, truthful Australian clips, lying Colombian – Spanish clips, truthful Colombian – Spanish clips, lying Colombian – English clips, truthful Colombian – English clips respectively), (ii) whether ratings differed for Australian versus

Colombian clips (contrast coefficients: 2, 2, -1, -1, -1, -1) and Spanish versus English clips (contrast coefficients: 0, 0, 1, 1, -1, -1), and (iii) whether the differences between ratings for truths versus lies vary depending on culture/language in the clip (contrast coefficients for veracity x culture interaction: -2, 2, 1, -1, 1, -1, and for the veracity x language interaction: 0, 0, -1, 1, 1, -1). Thus, the three families of contrasts tested for differences in trustworthiness ratings based on whether the person in the clip was either lying or telling the truth, the culture of the person depicted in the video clip, and, when the person was of a different culture, whether that person was speaking in his or her first or second language. As a dichotomous independent variable, the main effect of audio condition was examined as a between-subjects pairwise comparison. Bonferroni adjustments were made to keep the family wise Type I error rate at .05 such that each pair of contrasts for clip type and for the veracity x clip type interaction was made using alpha levels of .025.

The analysis of the first contrast comparing trustworthiness ratings between truthful and deceptive clips across all clip type conditions showed that participants' trustworthiness rating significantly differed as a function of veracity, F (1, 69) = 5.60, $p < .05$, and that the effect size was small, $d = 0.20$. Interestingly, as can be seen in Table 3, participants rated lying clips as more trustworthy than truth-telling clips (mean lying = 25.81, and mean truth-telling = 24.95).

The analysis of the second family of contrasts examining whether trustworthiness ratings varied depending on the culture or language condition of the clip showed that participants' ratings were significantly different as a function of culture, F (1, 69) = 20.66, $p < .025$, and that the effect size was medium, $d = 0.43$, and language, F (1, 69) = 83.77, $p < .025$, with a large effect size, $d = 1.00$. These results suggest that observers were likely to perceive Colombian video clips as less trustworthy than Australian clips (mean for Australian clips = 26.66 and mean for Colombian clips = 24.74), and also likely to perceive Colombian – English clips as more untrustworthy than the Colombian – Spanish clips (mean = 22.51 and mean = 26.98, respectively).

The analysis of the third family of contrasts examined whether participants' trustworthiness ratings of truthful and deceptive clips varied depending on the culture or language condition of the clip. They showed that participants' ratings were significantly different as a function of a veracity x culture interaction, F (1, 69) = 55.27, $p < .025$, and that the effect size was large $d = 0.68$, but not as a function of a veracity x language interaction, F (1, 69) = 0.72, $p > .025$. These results indicated an interesting pattern with truth-telling Australian clips being judged as less trustworthy than lying Australian clips, but Colombian lying clips showing the opposite pattern and were perceived as more untrustworthy than Colombian truth-telling clips.

The main effect of audio condition was not statistically significant, F (1, 69) = 1.79, $p > .025$ ($d = 0.19$), nor were audio condition's interactions with the clip condition contrasts statistically significant (Fs ≤ 3.52, ps > .025), indicating that participants' trustworthiness ratings did not differ as a function of the presence or absence of audio ($d \leq 0.18$).

Confidence ratings

For each video clip participants were also asked to rate on a 5 point rating scale (1 = *not very confident*; 5 = *extremely confident*) how confident they were of their lie/truth judgments. To create a scalar measure of participants' judgment about the credibility of each clip, these confidence scores for each clip were combined with participants' dichotomous truth/lie judgments as follows: (i) dichotomous credibility judgments were coded so that +1 = *lying* and -1 = *telling the truth*; (ii) credibility ratings were transformed by subtracting one from each rating, to produce a 0-4 scale; and (iii) these two values were multiplied, resulting in a nine-point scale in which -4 indicated *extremely confident the person is telling the truth*, +4 indicated *extremely confident the person is lying*, and zero (0) indicated that the person was not very confident of whichever dichotomous judgment was chosen.

The same three families of contrasts used in the trustworthiness analyses were used to examine: (i) whether scalar credibility ratings for lies versus truths differed across all clip type conditions, (ii) whether scalar credibility ratings differed as a function of culture or language, and (iii) whether the difference between scalar credibility ratings for lies and truths varied as a function of a culture or language. As with the previous analyses, Bonferroni adjustments were made to keep the family wise Type I error rate at .05 such that contrasts and interactions the alpha level was kept at .025. The mean and standard deviations for confidence ratings across clip type, audio and veracity conditions are reported in Table 4. Higher positive scores indicate that the participants were confident that the person depicted in the video clip was definitely lying, negative scores indicate that the participants were confident that the person was definitely telling the truth, and scores closer to zero indicate that the participants were not very confident whether the person was either lying or telling the truth.

As can be seen in Table 4, overall, participants were not particularly confident of their judgments. However, the first contrast analysis showed that participants' confidence ratings varied as a function of veracity F (1, 69) = 8.94, p <.05, with participants being generally more confident that actors in truth-telling clips were telling the truth. The effect size for this difference was small (d = 0.28).

The second analysis showed that participants' ratings differed as a function of culture F (1, 69) = 27.58, p < 0.25 (d = 0.48). As can be seen in Table 4, participants were generally more confident that the actors in the Australian clips were telling the truth compared to Colombian actors, thus providing further evidence of the truth bias seen in the response bias measure and trustworthiness scale. The second analysis also showed that the participants' ratings differed depending on the language spoken in the Colombian clips F (1, 69) = 65.32, p < 0.25 (d = 0.96). The participants were more confident that Colombians speaking in English were lying, whereas they were more confident that Colombians speaking in Spanish were telling the truth .The large effect size indicates that language was a major influence on these perceptions.

Table 4

Mean and standard deviations of confidence ratings across clips, audio and veracity conditions

	Clips		
Condition	Australian clips	Colombian – Spanish clips	Colombian – English clips
Audio			
Lying	-1.63 (1.57)	-0.75 (1.46)	0.60 (1.68)
Truth-telling	-0.26 (1.34)	-1.24 (1.27)	0.29 (1.27)
No Audio			
Lying	-1.56 (1.33)	-0.43 (1.13)	0.44 (1.50)
Truth -telling	0.29 (1.12)	-0.76 (1.30)	0.69 (1.25)

Note: Positive scores indicate confidence that the actor was lying and negative scores indicate confidence that the actor was telling the truth.

The third family of contrast analyses indicated that participants' confidence ratings for truth and lies significantly varied as a function of a veracity x culture interaction F (1, 69) = 56.11, $p < 0.25$ ($d = 0.67$), but not as a function of a language x culture interaction. The veracity x culture interaction indicates that Australian participants were generally less confident of their judgments for both truth-telling and lying clips. However, when judging truth-telling Colombian clips participants were more confident that the actors were telling the truth.

The main effect of audio was not statistically significant F (1, 69) = 2.50, $p > 0.25$ ($d = 0.20$), nor were the audio condition's interactions with the clip condition contrasts statistically significant (Fs ≤ 2.43, ps $> .025$), indicating that participants' confidence ratings did not differ as a function of the presence or absence of audio ($d \leq 0.3$).

DISCUSSION

The present study sought to investigate whether (i) participants' ability to discriminate between liars and truth-tellers (i.e., discriminability) differed as a function of culture and language, and (ii) whether participants' general tendency to assign truthful or deceptive judgments (i.e., response bias) was influenced by culture and language.

Discriminability

Consistent with the literature (e.g., Bond & DePaulo, 2006; Vrij, 2008), the findings of the signal detection analysis indicated that participants' discriminability was poor across all video clips. Overall accuracy rates were slightly below the level of chance with an average correct classification rate of 47.3% across all clips. Consistent with previous literature, this poor level of performance further highlights the unreliable nature of lie detection (e.g., Bond & DePaulo, 2006; Ekman, O' Sullivan & Frank, 1999; Vrij, 2008).

The signal detection analysis also showed that culture had a significant impact on lie detection performance. It was found that individuals were particularly poor at discriminating lies and truths from their compatriots – i.e., Australian video clips compared to Colombian clips. Specifically, it was found that participants' discrimination ability for Australian clips was significantly below the level of chance, but remained close to chance levels for Colombian clips. Interestingly, the results indicated that Australian clips depicting truth-tellers were more likely to be perceived as deceptive compared to Australian clips depicting liars. This tendency to judge Australian truth-telling clips as deceptive increased the number of false alarms and, therefore, resulted in overall poor lie detection accuracy for Australian clips. This finding was unexpected and raised the possibility that a particular clip (or several clips) in the Australian truth-telling condition may have skewed the results by appearing particularly suspicious. However, this does not appear to be the case. As described in the method section of this article, the video clips included in this task were carefully selected to be equivalent on a number of factors, i.e., attractiveness, friendliness, interestingness and likeability, all thought to possibly confound judgments. In light of the unexpected results for Australian clips, individual responses to those video clips were inspected again, and the clips were checked for possible peculiarities in the communicator's behavior. There was nothing obvious or peculiar noted about any person's behavior that could have confounded the results. Moreover, examination of participants' responses to those clips did not suggest that any particular clips were overwhelmingly regarded as highly suspicious or more credible than the others. Therefore there were no grounds to indicate that one or more clips needed to be excluded from the study. Further research needs to be conducted to explore this unexpected result.

The analysis of the second contrast indicated that an individual's capacity to discriminate between liars and truth-tellers was not influenced by language. In other words, participants' ability to discriminate between liars and truth-tellers was not significantly different when video clips depicted Colombians speaking in English compared to when the clips depicted Colombians speaking in Spanish. It was also found that audio did not have a significant impact on lie detection performance. Both these findings suggest that presence or absence of verbal cues did not seem to affect the observers' capacity to discriminate between liars and truth-tellers. More importantly, it suggests that the presence of verbal cues made no contribution to either improving or worsening performance (i.e., discriminability). These results indicate that the observers may have relied on behavioral cues rather than the verbal content of the clips to make their judgments of credibility. This is consistent with the literature on stereotypical cues to deception, which has shown that observers often associate deception with a number of behavioral cues (e.g., gaze aversion, body movements) and rely to a much lesser extent on verbal content when making judgments of credibility (e.g., Global Deception Research Team, 2006; Granhag, Andersson, Strömwall & Hartwig, 2004; Vrij, 2008).

Response bias

The hypothesis that observers would ascribe more deception judgments to Colombian video clips than Australian video clips was partially supported. The analysis in-

dicated that, overall, observers were more likely to ascribe truthful judgments to Australian clips than Colombian clips. This general tendency, which is often referred as truth bias, is consistent with the literature on lie detection, particularly by laypersons, within cultures contexts (Bond & Atoum, 2000; Bond & DePaulo, 2006; Bond & Rao, 2004; Levine, Kim, Park & Hughes, 2006; Levine, Park & McCornack, 1999). However, the difference in response bias across cultures was in the direction that suggests a tendency to greater suspicion of people from another culture – i.e., Colombian clips, in particular, those speaking in a second language. Therefore, even though the literature and the current findings for the Australian clips showed that there is a general tendency to make more truthful than untruthful judgments, this truth bias was attenuated and the tendency to suspect deception heightened for people of another culture.

Furthermore, as expected, observers were more likely to ascribe deceptive judgments to clips depicting Colombians speaking in English compared to Colombians speaking in Spanish. This also suggests that own-language favoritism was not observed in Australians' judgments of Colombian video clips. Interestingly, the results showed that whilst observers had a tendency to judge more Colombian – English clips as dishonest (i.e., lie bias) they also had a tendency to judge more Colombian – Spanish clips as truthful (i.e., truth bias).

There are two plausible explanations for a difference in response bias based on which language participants were speaking. The first explanation is based on verbal content; in particular, it is possible that observers would be less suspicious when they are aware that communicators are speaking in their first language, which the observer cannot understand. According to Bond and Atoum (2000) hearing someone speak in an unfamiliar language encourages observers to acknowledge their ignorance of the speaker's culture, and thus increases the tendency to give the communicator the benefit of the doubt. Coupled with a cross-cultural bias in deception judgments, this notion is consistent with the finding of a general tendency to be more suspicious of Colombians and with a lower suspicion of Spanish-speaking than English-speaking Colombians. However, the language difference was not dependent on the participants being able to hear that the person was speaking in their first language (i.e., the interaction effect with audio was not significant), which suggests that an explanation that relies on participants' access to verbal content cannot account for these findings.

The second explanation is that there are nonverbal behavioral differences associated with speaking in a second language, perhaps resulting from additional cognitive load or nervousness associated with speaking in a language in which one does not have great fluency. These behavioral differences may have prompted observers to believe that the communicator was being deceptive, and thus caused participants to ascribe more deceptive judgments to Colombian – English clips than Colombian – Spanish clips. The findings of the present study are consistent with this explanation. As described in the method section above, the 24 video clips used as stimulus materials in the present study were part of the principal's investigators doctoral dissertation. As part of that research program a behavioral analysis of these video clips on14 nonverbal and paraverbal behaviors (e.g., head

nods, smiling, eye contact, speech errors, etc.) was conducted (Castillo, 2011). The analysis showed that there were substantial behavioral differences between Colombians who were speaking Spanish and those who were speaking English. Specifically, it was found that Colombian participants made more functional hand and arm movements and their response latency was considerably shorter when speaking in their first language (Spanish) than their second language (English). These differences would have been available as cues to judgment irrespective of the participant's audio condition. Therefore, the difference in response bias appears likely to have been the result of perceptions of nonverbal behavior and not verbal content or language familiarity.

A mediation analysis could potentially provide a clearer indication of whether, in fact, behavioral differences in second language use results in a lie bias. However, the nature of the design of the present study did not allow for such an analysis to be conducted. Appropriately designed future research can explore this mediating relationship.

Ratings of credibility

In addition to the dichotomous credibility judgments, a number of continuous scales were used to assess participants' evaluations of the actors in the clips. These measurements were used in order to assess whether dichotomous judgments were not sufficiently fine-grained to identify significant cultural and language effects in the present study. However, these measures produced a similar general pattern of results to the dichotomous judgments, such that overall, Colombian clips were regarded as more suspicious than Australian clips, and within Colombian clips those depicting people speaking in their second language were regarded as more suspicious than those depicting people speaking in their first language. These measures also tended to show the unexpected result in which Australian truth-tellers tended to be regarded as more suspicious than Australian liars.

In conclusion, the clear difficulties associated with accurately distinguishing truthful from deceptive messages in a cross-cultural context and the potential for biases that result from culturally and linguistically based behavioral differences have important implications for many social, legal, business and national security settings. For instance, the tendency toward bias in cross-cultural judgments of deception could contribute to miscarriages of justice in which immigrants, asylum seekers or foreign visitors are wrongly suspected of deception because their behavioral pattern when speaking in a second language is indicative of cognitive load and thus misinterpreted as an attempt to hide the truth. The findings of the present study, however, should be taken as an initial demonstration of the potential for bias in cross-cultural judgments of deception. Future research should explore the impact of these cross-cultural factors in more applied and naturalistic settings.

REFERENCES

Bird, K., Hadzi-Pavlovic, D., & Isaac, A. (2000). PSY: Statistical program for contrast analysis. Sydney: School of Psychology, University of New South Wales.

Bond, C. F., Jr., & Atoum, A. O. (2000). International deception. *Personality and Social Psychology Bulletin, 26*(3), 385-395.

Bond, C. F., Jr., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10*(3), 214-234.

Bond, C. F., Jr., & Rao, S. R. (2004). Lies travel: Mendacity in a mobile world. In P. A. Granhag & L. A. Strömwall (Eds.), *The detection of deception in forensic contexts* (pp. 127-147). UK: Cambridge University Press.

Bond, C. F., Omar, A., Pitre, U., Lashley, B. R., Skaggs, L. M., & Kirk, C. (1992). Fishy-looking liars: Deception judgment from expectancy violation. *Journal of Personality and Social Psychology, 63*(6), 969-977.

Burgoon, J. K., & Ebesu Hubbard, A. S. (2005). Cross-cultural and intercultural applications of expectancy violations theory and interaction adaptation theory. In W. B. Gudykunst (Ed.), *Theorizing about intercultural communication* (pp. 149-171). Thousand Oaks, CA: Sage.

Castillo, P. A., (2011). *Cultural and cross-cultural factors in judgments of credibility* (Unpublished doctoral dissertation). Charles Sturt University. Australia.

Cheng, K. H., & Broadhurst, R. (2005). The detection of deception: The effects of first and second language on lie detection ability. *Psychiatry, Psychology and Law, 12*(1), 107-118.

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.

Ekman, P. (1972). Universals and cultural differences in facial expressions of emotion. In J. Cole (Ed.), *Nebraska symposium of motivation* (pp. 207-283). Lincoln, NE: University of Nebraska Press.

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology, 17*(2), 124-129.

Ekman, P., O' Sullivan, M., & Frank, M. G. (1999). A few can catch a liar. *Psychological Science, 10*(3), 263-266.

Frumkin, L. (2007). Influences of accent and ethnic background on perceptions of eyewitness testimony. *Psychology, Crime & Law, 13*(3), 317-331.

Global Deception Research Team. (2006). A world of lies. *Journal of Cross-Cultural Psychology, 37*(1), 60-74.

Granhag, P. A., Andersson, L. O., Strömwall, L. A., & Hartwig, M. (2004). Imprisoned knowledge: Criminals' beliefs about deception. *Legal and Criminological Psychology, 9*(1), 103-119.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. Oxford, England: John Wiley.

Inquisit 3.0 (2007) [computer software]. Seattle, WA: Millisecond Software LLC.

Levine, R., Anders, L., Banas, J., Baum, L., Endo, K., Hu, A., & Wong, N.C.H. (2000). Norms, expectations and deception: A norm violation model of veracity judgments. *Communication Monographs, 67*(2), 123-137.

Levine, T. R., Kim, R. K., Park, H. S., & Hughes, M. (2006). Deception detection accuracy is a predictable linear function of message veracity base-rate: A formal test of Park and Levine's probability model. *Communication Monographs, 73*(3), 243-260.

Levine, T. R., Park, H. S., & McCornack, S. A. (1999). Accuracy in detecting truths and lies: Documenting the "veracity effect." *Communication Monographs, 66*(2), 125-144.

Macmillan, N. A., & Creelman, C. (1990). Response bias: Characteristics of detection theory, threshold theory, and "nonparametric" indexes. *Psychological Bulletin, 107*(3), 401-413.

Matsumoto, D. (1989). Cultural influences on the perception of emotion. *Journal of Cross-Cultural Psychology, 20*(1), 92-105.

Matsumoto, D. (1991). Cultural influences on facial expressions of emotion. *The Southern Communication Journal, 56*(2), 128-137.

Matsumoto, D., Consolacion, T., Yamada, H., Suzuki, R., Franklin, B., Paul, S., . . . Uchida, H. (2002). American-Japanese cultural differences in judgments of emotional expressions of different intensities. *Cognition & Emotion, 16*(6), 721-747.

Matsumoto, D., & Kudoh, T. (1987). Cultural similarities and differences in the semantic dimensions of body postures. *Journal of Nonverbal Behavior, 11*(3), 166-179.

Matsumoto, D., Seung Hee, Y., & 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.

McCroskey, J. C., & Teven, J. J. (1999). Goodwill: A reexamination of the construct and its measurement. *Communication Monographs, 66*(1), 90-103.

Spencer-Rodgers, J., & McGovern, T. (2002). Attitudes toward the culturally different: The role of intercultural communication barriers, affective responses, consensual stereotypes, and perceived threat. *International Journal of Intercultural Relations, 26*(6), 609-631.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments & Computers, 31*(1), 137-149.

Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities* (2nd ed.). Chichester, West Sussex, Hoboken, N.J.: John Wiley & Sons.

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

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*(3), 169-184.

Vrij, A., & Winkel, F. W. (1992). Crosscultural police-citizen interactions: The influence of race, beliefs, and nonverbal communication on impression formation. *Journal of Applied Social Psychology, 22*(19), 1546-1559.

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*(2), 284-295.

Wilcox, R. R. (1987). New designs in the analysis of variance. *Annual Review of Psychology, 38*(1), 29-60.

Received: 12/2013
Accepted: 4/2014

Castillo, P.A., Tyson, G., & Mallard, D. (2014). An investigation of accuracy and bias in cross-cultural lie detection [Electronic Version]. *Applied Psychology in Criminal Justice, 10*(1), 66-82.