To freeze or not to freeze

A motion-capture approach to detecting deceit

Sophie Van Der Zee[1], Ronald Poppe[2], Paul J. Taylor[3 4], & Ross Anderson[1]

[1] Computer Laboratory, University of Cambridge, Cambridge, UK

[2] Information and Computing Sciences, Utrecht University, Utrecht, The Netherlands

[3] Psychology, Lancaster University, Lancaster, UK

[4] Psychology, University of Twente, Enschede, The Netherlands

Address correspondence to: sophie.van-der-zee@cl.cam.ac.uk

***Abstract*— We present a new robust signal for detecting deception: full body motion. Previous work on detecting deception from body movement has relied either on human judges or on specific gestures (such as fidgeting or gaze aversion) that are coded or rated by humans. The results are characterized by inconsistent and often contradictory findings, with small-stakes lies under lab conditions detected at rates only slightly better than guessing. Building on previous work that uses automatic analysis of facial videos and rhythmic body movements to diagnose stress, we set out to see whether a full body motion capture suit, which records the position, velocity and orientation of 23 points in the subject's body, could yield a better signal of deception. Interviewees of South Asian (*n* = 60) or White British culture (*n* = 30) were required to either tell the truth or lie about two experienced tasks while being interviewed by somebody from their own (*n* = 60) or different culture (*n* = 30). We discovered that full body motion – the sum of joint displacements – was indicative of lying approximately 75% of the time. Furthermore, movement was guilt-related, and occurred independently of anxiety, cognitive load and cultural background. Further analyses indicate that including individual limb data in our full body motion measurements, in combination with appropriate questioning strategies, can increase its discriminatory power to around 82%. This culture-sensitive study provides an objective and inclusive view on how people actually behave when lying. It appears that full body motion can be a robust nonverbal indicator of deceit, and suggests that lying does not cause people to freeze. However, should full body motion capture become a routine investigative technique, liars might freeze in order not to give themselves away; but this in itself should be a telltale.**

Keywords—deception; movement; nonverbal cues; interview; motion-capture; automatic analysis

I. Introduction

Although nonverbal cues to deception have been studied for decades, the current literature is characterized by inconsistent and often contradictory findings. For example, both leg movements and head movements have been found to both decrease [1, 2] and increase [3, 4] when lying. In an effort to clarify these mixed results, a number of researchers have provided meta-analyses [5, 6, 7, 8]. These concluded that the majority of cues (about 75%) of studied cues that researchers thought were related to deceit, and that they measured in their deception experiments, were not actually related to deceit (e.g. gaze aversion and postural shifts). Of those found to be related, the relationship between the cue and lying was typically weak [6, 9]. For example, DePaulo et al. [6] found that amongst nonverbal cues, only illustrators (movements that accompany or emphasize speech; d = -.14), general fidgeting (d =.16) and chin raising (d = .25) were significantly related to deception. In practice, this means that real-life differences between truth tellers and liars are more subtle and less clear than is stated in police interview manuals and indeed believed by the common public [4, 6].

Researchers have therefore sought to identify moderators of cue saliency. Zuckerman et al. [5] for example, argued that the type and magnitude of deceptive behavior is dependent on three factors: the extent to which liars experience arousal and emotions such as guilt, fear and delight [10]; the extent to which they experience cognitive load as a result of difficulties constructing and maintaining the lie [5, 9]; and how able they are to control their "lying behavior" [11]. Each of these three factors has been found to influence a liar's behavior in different and sometimes contradictory ways [5, 8, 12]. Emotions like guilt and fear have been found to decrease the production of illustrator gestures [13], while the increased physiological arousal caused by fear may increase self-adaptors and fidgeting [5]. Similarly, compared to truth telling, the excitement experienced when lying can increase the occurrence of body movements like smiling and illustrators [8], while cognitive load can reduce hand movement [14], foot and leg movement [7], overall body animation [8], and eye blinks [15]. As a consequence, either an increase or a decrease in specific behaviors can be a sign of lying (e.g., an increase in fidgeting caused by lie-related nervousness, or a decrease due to increased cognitive load or attempted behavioral control).

However, while researchers have gone to great lengths to increase the salience of cues within their studies, comparatively little effort has been made to improve the sensitivity with which nonverbal behavior itself is measured. As with most signal detection problems, effective progress within the field is made by both reducing the 'noise' surrounding the signal (i.e., by increasing its salience within the context) and by improving the efficiency with which one can measure the signal itself [16]. So far, most nonverbal deception research has derived its data by having researchers manually code videos, typically using a classification scheme [17]. There are several problems here. First, manual coding requires the researcher to decide what cues to code beforehand. This top-down research approach is useful, but it can curtail the detection of novel and lesser-known cues. Second, because manual coding is time-consuming, it creates a trade-off between the amount of data collected and the number of coded actions [17]. In other words, there is a limit to the diversity of behavior a research team can practically code, which again limits the chances of finding cues that are related to deceit. Third, manual coding is subjective and can cause reliability issues [18], which can lead to both false alarms and missed positives (i.e., cues going undetected). Fourth, manual coding in deception research is often expressed binomially (e.g., head movement: yes or no), and on rare occasions includes the duration of a movement [19]. The magnitude and direction of the movement are typically not taken into account, despite evidence that such differences carry the 'meaning' of the movement [20]. Fifth, researchers usually focus their coding on large movements, so smaller may go undetected.

All five of these issues may be tackled by replacing manual coding with an automatic measurement of nonverbal behavior. This can be done in many ways, such as the automatic coding of video footage [21] or the analysis of recorded motion capture data [17]. Automatic coding of video data does not require interpretation and is therefore more objective than manual coding. However, automatic video coding is typically based on 2D representations of behaviors that are 3D in real life, and this has been shown to impair the resulting analysis [21], and, additionally, video quality issues can significantly impair the robustness of automatic video-based analyses [22].

By contrast, full body motion-capture systems deliver rich, 3D data. For example, an Xsens full body suit contains 17 inertial sensors that register movement up to 120 times per second in three dimensions for 23 joints. The suit registers both local and global position data, so the experimenter knows how the subject's limbs move with respect to each other and to the floor; with this information it is possible to generate a 3D representation of the subject. Because there is no human in the analysis loop, it can be more objective and may be less likely to miss or misidentify cues. Recently, motion-capture research has assisted in the diagnosis of post-traumatic stress disorder (PTSD); Mahmoud et al. [23] have shown that a Kinect can be used to measure behaviors that are indicative of PTSD, such as rhythmic fidgeting and rocking.

Early results from automatic analyses of nonverbal behavior to detect deceit are promising. Using a video-based automatic analysis of deceptive facial expressions, Bartlett, et al. [24] were able to identify deceit with 85% accuracy using machine learning, while humans in their experiment did not perform better than 55%. This study demonstrates that some behaviors indicative of deception are difficult to pick up for humans, but can be robustly identified using automatic analyses. Similarly, Meservy et al. [22] were able to correctly identify deceit with 71% accuracy based on a neural network analysis of facial expressions and gestures; and blob analyses have been used to automatically classify deception-related behaviors such as agitation and behavioral control [25]. Although these studies provide an objective and inclusive measure of specific types of deceptive behavior, they are often limited to examining facial expressions [24, 26] or specific body parts such as face and hand movement [22, 25, 27]. This is a limitation because several manually coded studies have found that other aspects of body movement, for example foot, leg, and head movements, may also be indicative of deception [1, 2, 3, 4].

A. *Current study*

To take the optimal inclusive approach to investigating nonverbal indicators to deceit, in the current study we chose to implement an automatic analysis based on motion capture data, as it allows for a full body analysis. However, a sensitive analysis is more effective if there is little noise in the data. One factor that may cause noise in a behavioral data set is the cultural background of participants. Although no culture-specific cues to deceit have been identified so far, there is evidence that cultural background affects behavior in general (regardless of deceit) [28]. For example, even when being truthful, Surinamese participants naturally showed more behaviors that are related to deception (e.g., gaze aversion, speech disturbances and higher tone pitch) compared to Dutch participants [29]. These potential differences in baseline behavior between people from different cultural backgrounds led us to include cultural background as an independent variable in this study. This allows us to investigate objectively and inclusively how interviewees move different limbs when lying and when telling the truth, and if this movement differs between cultures.

To examine the impact of lying on nonverbal behavior, we conducted an experiment in which we compared full body behavior of interviewees telling truths or lies. The interview comprised of two tasks to investigate whether interview techniques that have previously shown to magnify behavioral differences between truth tellers and liars [4] have a similar enhancing effect on full body movement. We measured full body movement using Xsens MVN motion-capture suits. To achieve a culture-sensitive analysis of lying behavior, we compared the behavior of interviewees with a low-context cultural background (i.e., from a predominantly individualistic society) with interviewees with a high-context cultural background (i.e., from a predominantly collectivistic society) [30, 31, 32]. We did so in both within-cultural and cross-cultural interviews. Because theoretical models (i.e., the emotional, cognitive load and attempted behavioral control approaches) [5] and empirical research have demonstrated that movement can both increase and decrease when lying [1, 2, 3, 4], we refrained from postulating directive hypotheses.

II. METHOD

A. *Participants*

One hundred-and-eighty students and employees from Lancaster University (*M* Age = 22.43 yrs, Range 18 – 84, Males = 80) volunteered to participate as either an 'interviewee' or 'interviewer.' The dataset comprised of 18 male pairs, 28 female pairs and 44 mixed pairs. The experiment took approximately 70 minutes, and both interviewees (n = 90) and interviewers (n = 90) were paid £7.50 for their participation. Cultural background was an independent variable of interest in this study. Accordingly, we divided participants into low-context and high-context communicators based on their self-reported country of birth [31], and we combined them into three kinds of interviewer-interviewee pairs: British interviewer and interviewee (30 pairs); South Asian interviewer and interviewee (30 pairs); and British interviewer and South Asian interviewee (30 pairs). The latter cross-cultural pair was included because the nature of interactions between low-context interviewers and high-context suspects is relevant for law enforcement practice in predominantly low-context countries such as the UK and the US.

B. Measuring Absolute Movement

Absolute movement was measured using two full body Xsens motion-capture suits. For each person, we obtained the 3D positions of 23 joints in the body, which we normalized for global position in space using the processing described in [17]. The distance between poses of subsequent frames was then calculated as the sum of the differences of all joints. Absolute movement was measured as the mean value of pose differences over time. For body parts (i.e., arms, legs, head and body), we took into account only the differences for a subset of the joints. Before calculating the absolute movement per body part, we aligned the subset of joints on the body part root (i.e., shoulder, hip, neck or pelvis, respectively). This effectively eliminates the movement due to movement in other parts of the body. For example, leaning forward affects the shoulder locations. By aligning on the shoulder, only the movement of the (upper and lower) arm can be measured. We additionally created an aggregated measure of full body movement based on movement in all joints.

C. Materials

Post-interview questionnaire. On completing the interview, interviewers and interviewees completed a post-interview questionnaire that required them to respond to a series of statements, using a Likert scale ranging from 'not at all' (1) to 'very much' (7). The statements comprised a measure of cultural background and stereotype threat. They also asked participants to indicate how difficult they found their assignment, and how they felt after the interview on a range of emotions (i.e., frightened, anxious, fearful, nervous, guilty, regretful, repentant, penitential, happy, cheerful, pleased and enthusiastic).

Cultural background. To ensure the communication preferences of participants is consistent with our high-/low-context assignment based on country of birth, participants completed a 22-item cultural scale [31] derived from the 71-item scale by [30]. The 22-items captured participants attitudes towards indirect communication (3 items, e.g., "I catch on to what others mean even when they do not say it directly"), sensitivity for maintaining social harmony (5 items, e.g., "I often bend the truth if the truth would hurt someone"), humbleness in communication (8 items, e.g., "I am modest when I communicate with others"), and persuasion and multitasking (6 items, e.g., "I do not like to engage in several activities at the same time). One item in the original scale (Humbleness: "I listen very carefully to people when they talk") was excluded from analysis because it had an unduly detrimental impact on the scale's internal consistency (22 items, α = .65). The remaining 21-item scale showed acceptable consistency (α = .71).

Stereotype threat. To better understand the impact of cultural background on interviewees' experiences and feelings, especially when interacting cross-culturally, participants completed a 4-item stereotype threat measure. Stereotype threat is a situational predicament in which one can feel at risk of confirming negative stereotypes others may hold on their social group, and experiencing this threat can cause behavioral changes [33]. The 4-item measure asked participants: i) People sometimes make judgments about my honesty based on my ethnic group; ii) People sometimes make judgments about my trustworthiness based upon my ethnic group; iii) People sometimes think I am not a truthful person based on my ethnic group; and, iv) People sometimes think my behavior is suspicious based on my ethnic group. The internal consistency of the stereotype threat measure on our data was high (α = .93).

D. Procedure

The experiment comprised a pre-interview and an interview stage. The pre-interview stage required interviewees to complete two tasks (e.g., playing a computer game and handling a missing £5 note), while the interviewer received instructions about the interview. On completion of these tasks, the interviewee and interviewer were led to the interview room where they were each fitted a motion capture suit. Half of the interviewees were instructed to respond truthfully to the questions of the interviewer, while the other half were instructed to lie. Interviewees, regardless of veracity condition, were told that their name would be put in a prize draw for £50 if they managed to convince the interviewer that they were being truthful about both topics. This incentive was implemented to increase the stakes and to encourage participant motivation.

Pre-interview. After giving informed consent, interviewees received instructions about the two pre-interview tasks. These instructions differed depending on the veracity condition. The first task required participants to play a computer game called 'Never End' for seven minutes. Interviewees in the truth condition played the game for this time, while interviewees in the lie condition did not play the game. Instead, they studied an information sheet about the game, which provided them with details that enabled them to fabricate a story about playing the game. This enabled interviewees in both the truth and the lie condition to describe during the interview how they played the computer game, although only the participants in the truth condition actually had.

The second task involved handling a lost wallet that contained £5. In the truth condition, participants were asked to bring the wallet to the lost-and-found box, while, in the lie condition, participants were asked to remove the £5 note from the wallet and hide it somewhere on their body. These participants were instructed to put the wallet back where they found it instead of bringing it to lost and found. During the interview, interviewees in the lie condition were instructed to hide the fact that they had hidden the £5 note, and to pretend that they brought the wallet to the lost-and-found box.

Interview. After 12 minutes, the experimenter returned to the lab and checked that the interviewee had followed the instructions correctly. She then removed all remaining evidence (e.g., the wallet in the lie condition) and invited the interviewer into the room. She helped both interviewer and interviewee into Xsens motion-capture suits and invited them to sit on one of two chairs that were positioned facing one another. To ensure participants had an unobstructed view of the other's behavior, no table was situated between them.

Interviewers had previously been informed (while the interviewee was carrying out the pre-interview tasks) that they were to ask a set of pre-made questions about the computer game 'Never End' and about a missing £5 note. The questions about the missing £5 note were asked in normal order, while the questions about the computer game of 'Never End' involved interviewees to recall the event in reverse order. The latter was done to increase the cognitive load experienced by liars, which had previously been shown to magnify behavioral differences between truth tellers and liars [4]. Reverse order questions about the game 'Never End' were: (1) Please tell me how your game ended; (2) At what level did your game end? (3) What was your total score? (4) How was the score calculated? (5) For what item did you get the most points? (6) What happened when you went through an exit? (7) How many times did your character die? (8) How did your character usually die? (9) Please tell me about the lay-out of the game: any specific colors, effects or sounds? (10) Please tell me about the commands; (11) What is the main aim of this game? (12) Please tell me how your game started; and (13) Please tell me how you felt when playing the game. Normal order questions about the missing £5 note were: (1) Did you take the £5 while you were here playing 'Never End'? (2) Please explain what you were doing while you were in this room from start to finish. Include all details please; (3) So this means you went out of the room? (4) How long was the walk to the room where the lost property box was located? (5) Did you see anyone in the hallway while you were walking to the lost property box? (6) If so, how did he/she look like? (7) When you arrived in the room, how many items were in the lost property box? (8) Could you describe these items for me please? (9) What was written on the box? (10) What was next to the box? (11) Describe the room the lost property box was in; (12) Where did you put the wallet in the box, in relation to the other items? (13) How long were you gone from this room? (14) How do you feel about this money gone missing? and (15) Lastly, I will ask you again: did you take the £5? They were instructed that their task was to decide, for each topic, whether or not they thought the interviewee was being truthful. If their judgments were correct, their name would be put in a prize draw for £50. Thus, after setting up the equipment, the experimenter handed the interviewer his or her first set of questions and then retreated to monitor the incoming data. The participants spoke for 2.5 minutes about the computer game 'Never End', followed by 2.5 minutes about the missing £5 note. Interviews were cut off after 2.5 minutes to keep the length of the interaction consistent.

III. Results

A. Cultural Background Check

The 21-item cultural scale provided the opportunity to compare the culture-specific communication preferences and beliefs of the participants to their self-declared ethnicity [31]. An analysis of the average response over the 21 items revealed that participants classified as high-context scored higher on this scale (M = 5.06, SD = .56) than participants classified as low-context (M = 4.85, SD = .52), $t(178) = -2.61$, $p = .010$, suggesting that the initial division based on country of birth was acceptable. To reinforce this assessment, we examined participants' average Stereotype threat score as a function of their assigned culture, since those from high-context cultures typically report feeling greater stereotype threat [33]. Participants who were classified as high-context (M = 3.33, SD = 1.73) reported experiencing more stereotype threat than participants who were classified as low-context, (M = 1.88, SD = 1.02), $t(178) = -6.84$, $p < .001$. A follow-up Cultural background x Veracity ANOVA on average Stereotype threat score indicate that stereotype threat perceptions were not moderated by veracity condition, $F(1, 176) = 2.59$, $p = .110$, $\eta^2_p = .01$.

B. Mood Check

To examine the relationship between cultural group and participants' self-reported emotional experiences, we conducted a 2 (Veracity condition: truth and lie) x 3 (Culture condition: low-context, high-context and mixed) MANOVA with reported feelings of being Frightened, Anxious, Fearful, Nervous, Guilty, Regretful, Repentant, Penitential, Happy, Cheerful, Pleased and Enthusiastic as the Dependent Variables. We have reverse-scored the positive emotions (Happy – Unhappy, Cheerful – Cheerless, Pleased – Displeased, and Enthusiastic – Unenthusiastic), in order for all emotions to be scored in the same direction (i.e., the higher the more negative). Interviewees' emotional experience varied as a function of both Veracity, $F(12, 73) = 3.81$, $p < .001$, $\eta^2_p = .39$, and Culture, $F(24, 148) = 1.61$, $p = .046$, $\eta^2_p = .21$. Fig. 1 illustrates the effect of Veracity on self-reported emotions. As can be seen from Fig. 1, compared to participants who told the truth, participants who lied reported feeling more Frightened, $F(1, 84) = 4.99$, $p = .028$, $\eta^2_p = .06$, more Anxious, $F(1, 184) = 4.61$, $p = .035$, $\eta^2_p = .05$, more Fearful, $F(1, 84) = 7.93$, $p = .006$, $\eta^2_p = .09$, more Guilty, $F(1, 84) = 28.56$, $p < .001$, $\eta^2_p = .25$, more Regretful, $F(1, 84) = 9.68$, $p = .003$, $\eta^2_p = .10$, more Penitential, $F(1, 84) = 16.56$, $p < .001$, $\eta^2_p = .17$, more Unhappy, $F(1, 84) = 11.17$, $p = .001$, $\eta^2_p = .12$, more Cheerless, $F(1, 84) = 13.41$, $p < .001$, $\eta^2_p = .14$, and more Displeased, $F(1, 84) = 16.81$, $p < .001$, $\eta^2_p = .17$. There were no differences across Veracity for feeling Nervous, $F(1, 84) = 3.91$, $p = .051$, $\eta^2_p = .05$, feeling Repentant, $F(1, 84) = .82$, $p = .369$, $\eta^2_p = .01$ and feeling Unenthusiastic, $F(1, 84) = 2.07$, $p = .154$, $\eta^2_p = .02$.

Fig. 2 illustrates the direction of the significant effects that Culture has on emotion experience when telling truths or lies. Culture condition affected feelings of Nervousness, $F(2, 84) = 6.02$, $p = .004$, $\eta^2_p = .13$, with interviewees in the low-context condition (M = 4.47, SD = 2.28) feeling more nervous than high-context interviewees in both the high-context (M = 3.20, SD = 1.92), $p = .019$, and mixed condition (M = 3.03, SD = 1.73), $p = .006$; feelings of Unhappiness, $F(2, 84) = 3.63$, $p = .031$, $\eta^2_p = .08$, with interviewees in the low-context condition (M = 3.67, SD = 1.49) reporting feeling unhappier than interviewees in the high-context condition (M = 2.73, SD =1.39), $p = .026$; feelings of Cheerlessness, $F(2, 84) = 4.34$, $p = .016$, $\eta^2_p = .09$, with interviewees in the low-context condition (M = 3.70, SD = 1.54), reporting feeling more cheerless than

Fig. 1. The effect of veracity on a range of self-reported emotions (error bars = 95% CI).

interviewees in the high-context condition (*M* = 2.77, *SD* = 1.31), *p* = .021, and feeling Unenthusiastic, *F*(2, 84) = 5.52, *p* = .006, η^2_p = .12, with interviewees in the low-context condition (*M* = 3.33, *SD* = 1.21) reporting feeling more unenthusiastic than interviewees in the high-context condition (*M* = 2.33, *SD* = 1.24), *p* = .008 and the mixed condition, (*M* = 2.53, *SD* = 1.28), *p* = .044. Culture condition did not affect feeling Frightened, *F*(2, 84) = .83, *p* = .439, η^2_p = .02, feeling Anxious, *F*(2, 184) = .13, *p* = .874, η^2_p < .01, feeling Fearful, *F*(2, 84) = .61, *p* = .545, η^2_p = .01, feeling Guilty, *F*(2, 84) = 2.57, *p* = .083, η^2_p = .06, feeling Regretful, *F*(2, 84) = 1.29, *p* = .281, η^2_p = .03, feeling Repentant, *F*(2, 84) = 1.70, *p* = .189, η^2_p = .04, feeling Penitential, *F*(2, 84) = .86, *p* = .427, η^2_p = .02, and feeling Displeased, *F*(2, 84) = 1.66, *p* = .197, η^2_p = .04.

C. Full Body Motion

To examine whether truth tellers and liars show different nonverbal movement, and to test whether or not this movement was moderated by cultural context, we examined absolute movement (i.e., displayed as centimeters per second) as a function of Veracity condition and Culture condition. Fig. 3 shows the full body motion data as a function of Veracity and Task across Culture conditions. A 2 (Veracity condition) x 3 (Culture condition) x 2 (Task) mixed ANOVA with Task as the repeated measure and full body movement as the dependent variable revealed main effects for both Task, *F*(1, 84) = 36.66, *p* < .001, η^2_p = .30, and Veracity condition, *F*(1, 84) = 29.41, *p* < .001, η^2_p = .26, which were subsumed in a Task x Veracity interaction, *F*(1, 84) = 17.99, *p* < .001, η^2_p = .18. Although in general liars (*M* = 9.87, *SD* = 5.32) moved more than truth tellers (*M* = 5.97, *SD* = 3.67), how much interviewees moved was dependent on what topic they were discussing. While truth tellers moved similar amounts during both the computer game ‘Never End’ task (*M* = 6.07, *SD* = 3.54) and the missing £5 note task (*M* = 5.87, *SD* = 3.81), liars moved much more when being interviewed about the computer game ‘Never End’ (*M* = 11.70, *SD* = 5.95), compared to the missing £5 note (*M* = 8.03, *SD* = 4.69). However, Culture did not affect full body movement, *F*(2, 84) = .50, *p* = .609, η^2_p = .01.

When examining the movement data into more detail, we found that the full body movement (i.e., an interaction effect of Task and Veracity condition) was replicated at the level of individual limbs. A series of equivalent ANOVAs revealed significant interaction effects between Task and Veracity for the left arm, *F*(1, 84) = 9.46, *p* = .003, η^2_p = .10, right arm, *F*(1, 84) = 21.78, *p* < .001, η^2_p = .21, left leg, *F*(1, 84) = 6.47, *p* = .013, η^2_p = .07, right leg, *F*(1, 84) = 9.68, *p* = .003, η^2_p = .10, head *F*(1, 84) = 21.83, *p* < .001, η^2_p = .21, and torso, *F*(1, 84) = 17.48, *p* < .001, η^2_p = .17. Interestingly, Culture condition only affected the amount that interviewees moved their legs, and did not affect any other body parts. In general, Interviewees in the low-context condition (*M* = 1.20, *SD* = 1.69) moved their right leg more compared to interviewees in the high-context (*M* = .59, *SD* = .49) and mixed condition (*M* = .57, *SD* = .42), *F*(2, 84) = 3.67, *p* = .030, η^2_p = .08.

D. Detecting Deception on an Individual Level

To measure how discriminative full body movement would be when applied on an individual level, we calculated how much truthful interviewees moved in total when discussing the game ‘Never End’ (*M* = 6.07, *SD* = 3.54) and the missing £5 note (*M* = 5.87, *SD* = 3.81), and how much deceptive interviewees moved in total when discussing the game ‘Never End’ (*M* = 11.70, *SD* = 5.95) and the missing £5 note (*M* = 8.03, *SD* = 4.69). Subsequently, we ran a binary logistic regression to calculate the predictive value of full body movement when discussing the game ‘Never End’, and full body movement when discussing the missing £5 note on predicting deception. A test of the full model against a constant-only model was statistically significant, indicating that the full body movement predictors reliably distinguished between truth tellers and liars, X^2 (2) = 35.19, *p* < .001, Nagelkerke R^2 = .43. Overall, we correctly classified 74.4% (truths: 80.0%, lies: 68.9%) of the interviewees as either being truthful or deceptive based on their full body movement. We ran a second binary logistic regression to calculate if incorporating individual limb movement increases the predictive validity of our model. We included absolute movement values of both arms, both legs, the head and the body during the interview about the game ‘Never End’ and the interview about the missing £5 note to predict if the participant was lying or being truthful. Again, a test of the full model against a constant-only model was statistically significant, indicating that the individual limb movement predictors, as a set, reliably distinguished between truth tellers and liars, X^2 (12) = 48.45, *p* < .001, Nagelkerke R^2 = .56. Overall, we correctly classified 82.2% (truths: 88.9%, lies: 75.6%) of the interviewees as either being truthful or deceptive based on the combined movement in their individual limbs.

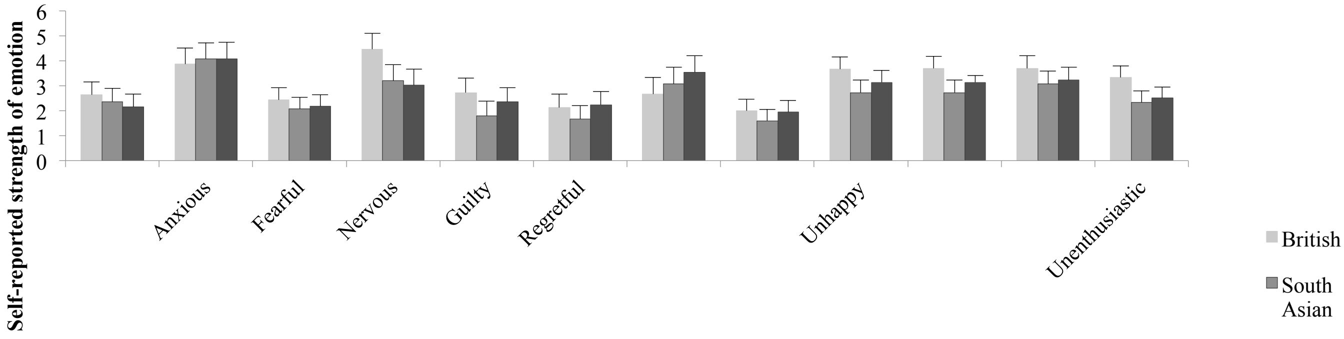

Fig. 2. The effect of culture on a range of self-reported emotions (error bars = 95% CI).

E. *Influence of Emotions and Cognitive Load on Motion*

To measure whether self-reported difficulty, implemented as a measure of experienced cognitive load, affects movement, we calculated correlations between difficulty and the interviewee's full body movement when answering questions about the game 'Never End' and when answering questions about the missing £5 note. Although liars (M = 3.30, SD = 1.63) did report finding their assignment more difficult than truth tellers (M = 2.02, SD = 1.21), $F(1, 84) = 21.31$, $p < .001$, $\eta^2_p = .20$, this increase in difficulty did neither affect full body movement during the game 'Never End', $r = .089$, $n = 90$, $p = .404$, nor during the missing £5 note, $r = .038$, $n = 90$, $p = .724$. To investigate if any specific limbs were affected by cognitive load, we calculated a correlation matrix of self-reported difficulty on absolute movement in individual limbs when being interviewed about both topics. However, movement in none of the limbs was correlated with self-reported difficulty during any of the tasks.

To measure if mood has an impact on how people move, we calculated correlations between self-reported mood and the interviewee's full body movement when answering questions about the game 'Never End' and when answering questions about the missing £5 note. The results indicated that feeling Guilty and feeling Penitential were positively correlated with full body movement, but only when answering reverse order questions about the Game 'Never End'. In other words, interviewees that indicated feeling guilty moved more than interviewees who reported feeling less guilty, $r = .247$, $n = 90$, $p = .019$. Similarly, interviewees that indicated feeling penitential moved more than interviewees who reported feeling less penitential, $r = .260$, $n = 90$, $p = .013$. None of the other self-reported emotions were correlated with full body movement. To investigate these effects into more detail, we created a correlation matrix of self-reported mood on absolute movement in individual limbs when being interviewed about both topics, displayed in Table 1. Interestingly, anxiety-related emotion such as feeling Frightened, Anxious, Fearful and Nervous had no effect on movement in any of the interviewee's limbs. Neither did the pleasant emotions, such as feeling Happy, Cheerful, Pleased and Enthusiastic. The self-reported emotions that did affect movement were all guilt related: feeling Penitential, Guilty and Regretful, although feeling Regretful and Repentant were not correlated with any type of movement. Although the guilt-related emotions affected movement in the interviewee's arms, head and body, they did not affect leg movement. Also, the guilt-related emotions had a larger effect on the interviewee's movement during the reverse order questioning about the game 'Never End', compared to when discussing the missing £5 note.

IV. Discussion

We started this paper by noting that research on nonverbal indicators of deceit has reported inconsistent and even contradictory results [8], and that the identified cues often have a weak relationship [5, 6, 7, 8]. We set out to investigate whether this lack of reliable nonverbal cues can be remedied by more sensitive measurements. We used full body motion capture suits to automatically measure whether people move any of their body parts and compared total body motion when lying with when being truthful. We did not hypothesize a direction of the results because both on a theoretical and on a practical level, mixed findings have been reported. With an effect size of .26, our results indicate that a reliable nonverbal indicator of deceit is full body motion; a measure that includes not just discrete, large and easily coded movements, but also the many smaller movements that people make. An examination of full body motion showed that people who lied moved more than people who spoke the truth. Based on one aggregated full body motion measure, we could correctly classify 74.4% (truths: 80.0%, lies: 68.9%) of all interactions, and based on movement in the individual limbs, we could even

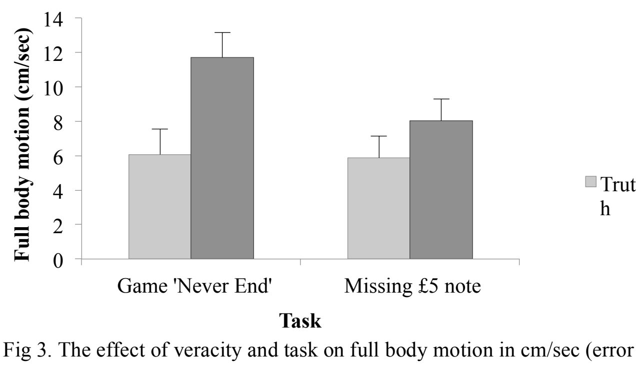

Fig 3. The effect of veracity and task on full body motion in cm/sec (error bars = 95% CI).

Table 1. Correlations between self-reported emotions and movement per limb, divided by task. * indicates the correlation is significant at the .05 level, and ** indicates the correlation is significant at the .01 level (2-sided).

	Game 'Never End'						*Missing £5 note*					
Emotions	*Left arm*	*Right arm*	*Left leg*	*Right leg*	*Head*	*Body*	*Left arm*	*Right arm*	*Left leg*	*Right leg*	*Head*	*Body*
Frightened	.026	-.034	.021	.034	.012	.057	-.049	.055	.002	-.031	.027	.029
Anxious	.101	.112	-.025	-.049	.109	.118	.046	.077	-.001	-.071	.039	.058
Fearful	.064	.029	.105	.074	.108	.174	.033	.088	.137	.036	.101	.160
Nervous	.006	-.023	.120	.158	-.060	.060	-.071	.074	.093	.113	-.083	.004
Guilty	.243*	.240*	.198	.108	.244*	.373**	.151	.220*	.164	.025	.133	.252*
Regretful	.154	.127	.090	.003	.155	.224*	.146	.158	.070	-.057	.133	.211*
Repentant	.124	.103	-.051	-.135	.086	.118	.059	.021	-.038	-.148	-.013	.016
Penitential	.278**	.262*	.010	-.017	.234*	.280**	.160	.201	.004	-.077	.151	.197
Unhappy	.057	.134	.087	.056	.100	.181	.031	.132	.067	.012	.015	.121
Cheerless	.069	.086	.092	.078	.041	.150	-.039	.053	.062	.034	-.068	.069
Displeased	.175	.120	.077	.035	.097	.177	.094	.015	.061	-.008	.001	.112
Unenthusiastic	-.112	-.118	.122	.142	-.093	.085	-.078	-.014	.121	.124	-.119	.072

further increase our correct classification to 82.2% (truths: 88.9%, lies: 75.6%). Compared to an average detection rate of around 54% in similar experimental settings when humans attempt to detect deceit [34], 82.2% is a solid improvement. The discriminatory power of absolute movement might be even further boosted by interviewing technique, because in line with our expectations [4], behavioral differences between truth tellers and liars were larger when interviewed in reverse order (i.e., the game 'Never End'), compared to being interviewed in normal order (i.e., the missing £5 note). In short, interviewing techniques can enlarge behavioral differences between truth tellers and liars [4, 35]. In addition to identifying a new method to reliably identifying deception based on absolute movement, this study also supports the argument by Vrij and Granhag [36] to focus on the type of questions being asked (i.e., interviewing techniques) when conducting deception research.

Although previous research has demonstrated baseline differences between people from low-context (i.e., British) and high-context cultures (i.e., South Asian), we did not find any such differences in full body movement. A closer examination of the individual limbs revealed that only leg movement was influenced by cultural background. Regardless of lying, British interviewees moved their right legs more than South Asian interviewees during within-culture and cross-culture interviews. Because this effect only occurred in one limb, and none of the other culture-sensitive studies on nonverbal indicators of deceit have included leg movement because usually only upper-body movement in included in such analyses, we cannot assess the stability of this finding without further research. On the one hand, leg movement might be a culture-dependent movement that has not been previously identified due to lack of measurement. On the other hand, culture-sensitive movements that have been previously identified, such as smiles, hand and arm movement, trunk movement, and self-touches were occurring more often in high-context individuals than in low-context individuals [29]. In this line of thought it would be expected that high-context individuals display more leg movement than low-context individuals, instead of less. Regardless, our results confirm the lack of culture-specific cues to deceit, and even provide evidence that there are little differences in baseline behavior between British and South Asian individuals.

Currently, the lack of identified reliable nonverbal indicators of deceit is being explained by the moderating function of the emotional, cognitive-load and attempted behavioral-control approach. To test how these different approaches actually impact movement, and if they serve as moderators, we asked participants to self-report how difficult they found their assignment, and how they were feeling on a range of emotions. The results indicate that although liars reported finding their assignment more difficult than truth tellers, self-reported difficulty (implemented as a measure of cognitive load) was not correlated with movement in any of the limbs during either of the tasks. Although in the literature, increased cognitive load has been associated with a reduction in movement [8, 14], this assumption was not supported by our data. Similarly, emotions elicited by the act of lying have been proposed to explain some of the variation in nonverbal cues to deceit. For example, lying-related fear is believed to increase physical arousal, self-adaptors, fidgeting and eye-blinks [5], while guilt might be related to gaze aversion [8]. Our findings are in line with the existing literature from the perspective that lying can increase experiences cognitive load and affect emotions, as liars reported finding the interview more difficult than truth tellers, and liars also reported feeling more negative than truth tellers. However, correlation analyses indicated that difficulty and anxiety related emotions did not influence nonverbal behavior. Instead, behavioral changes in liars were caused by guilt-related emotions.

There are two possible explanations for the lack of correlations between cognitive load and movement, and between the anxiety-related emotions and movement in our data. One possibility is that the nature of the lies that participants were instructed to tell in this study affected the emotional and cognitive processes that can be elicited by lying. The lies were low-stake in the sense that participants would not be punished if they failed to convince the interviewer of their honesty. We did try to increase the stakes in several ways. First, we implemented a task with criminal intent. Participants in the lie condition had to steal a £5 note and hide this information from the interviewer by providing a false alibi. Second, we offered interviewees the chance of winning £50 if they managed to convince the interviewer of their innocence. And although participants did not move differently depending on how difficult they found the assignment, liars did report finding the assignment much harder than truth tellers. Lying also negatively affected people's self-reported emotions: liars reported feeling more frightened, anxious, fearful, guilty, regretful, and penitential than truth tellers, and reported feeling less happy, cheerful and pleased than truth tellers. The lie-induced differences in self-reported mood and experienced difficulty indicate that it is unlikely that the lack of correlation between cognitive load and movement, and between emotions and movement is caused by the nature of the lie. A second explanation for the lack of correlations with anxiety-related emotions and cognitive load is that it is not impossible that it was caused by the use of motion-capture equipment to measure movement instead of using manual coding. To investigate if the same data can lead to different conclusions based on the type of coding used (i.e., differences between manual coding and automatic coding based on motion-capture data), more research on this topic should be conducted in the future. This future research could indicate whether our results can be explained by methodological choices, or whether the assumption that cognitive load and anxiety related emotions cause liars to behave differently might need to be reconsidered.

A. Limitations and Future Research

To increase generalizability we implemented two types of lies, including a more controversial one as the missing £5 note topic required denying stealing money. The stakes were still relatively low, as there were no consequences to being caught. Raising the stakes, and for example investigating lies that are told voluntarily instead of on instruction [37], will provide a better insight into the generalizability of our results.

A tactic that has proven to be beneficial in the detection of deceit is the manipulation of interview technique. Here, we implemented the reverse order questioning technique, which has shown to magnify differences between truth tellers and liars by imposing extra cognitive load on the liar [4]. However, other interview techniques have also proven to elicit more, and more blatant, cues to deceit [35], such as asking unanticipated questions [38] and playing devil's advocate [39]. Experimenting on such techniques using motion-capture systems looks like a promising line of research.

A second limitation is the possible hindrance of natural movement by the motion-capture suits. We did what we could to minimize the effect by giving all participants time to get used to the suit by starting the interview with a baseline, neutral conversation. In future research this potential issue can be solved by using two or more video cameras to create a point cloud model of the subject's body or by using millimeter-wave radar to measure total movement directly. Such techniques could allow unobtrusive surveillance of subjects in a police interview room or other operational interrogation environment. Remotely measuring nonverbal behavior in an accurate and objective manner will further help bridging the current gap between theory and practice in improving ways to detect deception [40].

V. Acknowledgment

The authors would like to thank Lieke Rotman, MSc, and Prof. Ellen Giebels for the important role they played in the design and data collection of this study. Dr. Sophie Van Der Zee was supported by the UK Engineering and Physical Sciences Research Council. This publication was supported by the Dutch national program COMMIT.

VI. References

[1] Buller, D. B., Burgoon, J. K., White, C. H., & Ebesu, A. S. (1994). Interpersonal deception VII: Behavioral profiles of falsification, equivocation, and concealment. *Journal of Language and Social Psychology, 13,* 366-395.

[2] Levine, T. R., Feeley, T. H., McCornack, S. A., Hughes, M., & Harms, C. M. (2005). Testing the effects of nonverbal behavior training on accuracy in deception detection with the inclusion of a bogus control group. *Western Journal of Communication, 69,* 203-217.

[3] Davis, M., Markus, K. A., Walters, S. B., Vorus, N., & Connors, B. (2005). Behavioral cues to deception vs. topic incriminating potential in criminal confessions. *Law and Human Behavior, 29*, 683-704.

[4] Vrij, A., Mann, S., Fisher, R., Leal, S., Milne, B., & Bull, R. (2008). Increasing cognitive load to facilitate lie detection: The benefit of recalling an event in reverse order. *Law and Human behavior, 32,* 253-265.

[5] Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. In L. Berkowitz (Ed.), *Advances in experimental social psychology*, volume 14, (pp. 1-57). New York, US: Academic Press.

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

[7] Sporer, S. L., & Schwandt, B. (2007). Moderators of nonverbal indicators of deception: A meta-analytic synthesis. *Psychology, Public Policy, and Law, 13,* 1-34.

[8] Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities*. Chichester, England: Wiley.

[9] Vrij, A., Granhag, P. A., & Porter, S. (2010). Pitfalls and opportunities in nonverbal and verbal lie detection. *Psychological Science in the Public Interest, 11,* 89–121.

[10] Ekman, P. (1989). Why lies fail and what behaviors betray a lie. In J. C. Yuille (Ed.), *Credibility assessment* (pp. 71-82). Dordrecht, the Netherlands: Kluwer.

[11] Burgoon, J. K., & Buller, D. B. (1994). Interpersonal deception: iii. Effects of deceit on perceived communication and nonverbal behavior dynamics. *Journal of Nonverbal Behavior, 18,* 155-184.

[12] Akehurst, L., & Vrij, A. (1999). Creating suspects in police interviews. *Journal of Applied Social Psychology, 29,* 192-210.

[13] Ekman, P. (1988). Lying and nonverbal behavior: Theoretical issues and new findings. *Journal of Nonverbal Behavior, 12*, 163-176.

[14] Ekman, P., & Friesen, W. V. (1972). Hand movements. *Journal of Communication, 22,* 353-374.

[15] Bagley, J., & Manelis, L. (1979). Effect of awareness on an indicator of cognitive load. *Perceptual and Motor Skills, 49,* 591-594.

[16] Taylor, P. J., Bennell, C., & Snook, B. (2002). Problems of classification in investigative psychology. In K. Jajuga, A. Sokolowski, & H.-H., Bock (Eds.), *Classification, clustering, and data analysis: Recent advances and applications* (pp. 479-487). Verlag, Heidelberg: Springer.

[17] Poppe, R. W., Van Der Zee, S., Taylor, P. J., & Heylen, D. K. J. (2014). AMAB: Automated Measurement and Analysis of Body Motion. *behavior Research Methods, 46,* 625-633.

[18] Scherer, K. R., & Ekman, P. (1982). Methodological issues in studying nonverbal behavior. In K. R. Scherer, & P. Ekman (Eds.), *Handbook of methods in nonverbal behavior research* (pp. 1-44). New York: Cambridge University Press.

[19] Hillman, J., Vrij, A., & Mann, S. (2011). Um… They were wearing…: The effect of deception on specific hand gestures. *Legal and Criminological Psychology, 17,* 336-345.

[20] McNeill, D. (1985). So you think gestures are nonverbal? *Psychological Review, 92,* 350-371.

[21] Metaxas, D., & Zhang, S. (2013). A review of motion analysis methods for human nonverbal communication computing. *Image and Vision Computing, 31,* 421-433.

[22] Meservy, T. O., Jensen, M. L., Kruse, J., Burgoon, J. K., Nunamaker, J. F. Jr., Twitchell, D. P., Tsechpenakis, G., & Metaxas, D. N. (2005). Deception detection through automatic, unobtrusive analysis of nonverbal analysis. *Intelligent Systems, IEEE, 20,* 36-43.

[23] Mahmoud, M., Morency, L-P.., Robinson, P. (2013). Automatic Multimodal Descriptors of Rhythmic Body Movement. *Proceedings of the 15th ACM on International conference on multimodal interaction (ICMI 2013), Sydney, 2013, 429-436.*

[24] Bartlett, M. S., Littlewort, G. C., Frank, M. G., & Lee, K. (2014). Automatic decoding of facial movements reveals deceptive pain expressions. *Current Biology, 24,* 738-743.

[25] Lu, S., Tsechpenakis, G., & Metaxas, D. N., Jensen, M. L., Kruse, J. (2005). *Blob analysis of the head and hands: A method for deception detection.* Presented at Hawaii International Conference on System Science (HICSS'05), Hawaii, 2005.

[26] Yu, X., Zhang, S., Yan, Y., Dunbar, N., Jensen, M. L., Burgoon, J. K., Metaxas, D. N. (2013). *Automated analysis of interactional synchrony using robust facial tracking and expression recognition.* 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition, Shanghai, 2013.

[27] Yu, X., Zhang, S., Yan, Z., Yang, F., Huang, J., Dunbar, N., Jensen, M., Burgoon, J. K., & Metaxas, D. N. (2013*). Is interactional dissynchrony a clue to deception: Insights from automated analysis of nonverbal visual cues.* Presented at Hawaii International Conference on System Science (HICSS'13), Hawaii, 2013.

[28] Taylor, P. J., Tomblin, S., Conchie, S., & Van Der Zee, S. (2014). Cross-cultural deception detection. In Vrij, A., Granhag, P. A., & Verschuere, B. (Eds.), *Deception detection.* Chichester: Wiley Blackwell.

[29] Vrij, A., & Winkel, F. W. (1991). Cultural patterns in Dutch and Surinam nonverbal behavior: An analysis of simulated police/citizen encounters. *Journal of Nonverbal behavior, 15,* 169-184.

[30] Adair, W. L., Buchan, N., & Chen, X. P. (2009, June). *Communication and social interaction style (CSIS): Reconceptualization and scale development.* Paper presented at the Association for International Business conference, San Diego, CA.

[31] Beune, K., Giebels, E., Adair, W. L., Fennis, B. M., & Van der Zee, K. I. (2011). Strategic sequences in police interviews and the importance of order and cultural fit. *Criminal Justice and behavior, 38,* 934-954.

[32] Hall, E. T. (1976). *Beyond Culture.* Garden City, New York: Anchor Press/Doubleday.

[33] Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52,* 613-629.

[34] Bond, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgements. *Personality and Social Psychology Review, 10,* 214-234.

[35] Vrij, A., Granhag, P. A., Mann, S., & Leal, S. (2011). Outsmarting the liars: Toward a cognitive lie detection approach. *Current directions in Psychological Science, 20,* 28-32.

[36] Vrij, A., & Granhag, P. A. (2012). Eliciting cues to deception and truth: What matters are the questions asked. *Journal of Applied Research in Memory and Cognition, 1*, 110-117.

[37] Duran, N. D., Dale, R., Kello, C. T., Street, C. N. H., & Richardson, D. C. (2013). Exploring the movement dynamics of deception. *Frontiers in Psychology, 4*, 1-16.

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

[39] Leal, S., Vrij, A., Mann, S., & Fisher, R. (2010). Detecting true and false opinions: The devil's advocate approach as a lie detection aid. *Acta Psychologica, 134,* 323–329.

[40] Kassin, S. M. (2012). Paradigm shift in the study of human lie-detection: Bridging the gap between science and practice. *Journal of Applied Research in Memory and Cognition, 1*, 118-119.