# NONVERBAL AND PARALINGUISTIC BEHAVIOURS DURING THE ADULT ATTACHMENT INTERVIEW: THE CONSTRUCTION OF A NOVEL CODING SYSTEM

LAURA GOLDSTEIN

A THESIS SUBMITTED TO THE FACULTY OF GRADUATE STUDIES IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF ARTS

GRADUATE PROGRAM IN CLINICAL DEVELOPMENTAL PSYCHOLOGY
YORK UNIVERSITY
TORONTO, ONTARIO

June 2019

## Abstract

The current study outlines the development, construction, and reliability of a novel coding tool for the Adult Attachment Interview and was designed to address perceived deficits in both research and clinical practice. Just as nonverbal behavior finds its roots in ethology, so too does attachment theory. While the Ainsworth Strange Situation Protocol adhered to the ethological roots of attachment theory by observing nonverbal behavior, the Adult Attachment Interview is traditionally scored based on verbal content alone. At present, there does not exist a reliable taxonomy of nonverbal behaviors associated with adult attachment style. As the first stage of a larger stepwise research program, the current study describes the process of manual development by selecting relevant factors through a recursive process of literature review and direct video observation, coding process, training, and feedback, and presents preliminary reliability estimates and agreement statistics for both individual behaviors and larger aggregate behavioral categories. These preliminary results showed great promise for the newly developed coding tool, allowing the investigators to identify 1) reliable behavioral categories and individual behaviors, 2) behavioral categories and individual behaviors that demonstrated sensitivity to training and feedback, and 3) individual behaviors and categories that require further remediation and investigation in future studies.

# Dedication

*For my parents – Bruce and Susan Goldstein*

*For my partner – Ilja Pfaff*

*Because blood isn't always thicker than water*

# Acknowledgements

First and foremost, I would like to share my gratitude with those who will likely never read this manuscript, but who should nonetheless be acknowledged: Thank you to all our parents who participated in the joint EFFT/AAI study throughout the years. Thank you for trusting us with your most profound sorrows and your greatest joys. Thank you for making this possible.

Thank you to the wonderful professionals who consulted on the early stages of this project: Leslie Atkinson, Christine Shea, Dana Shai, Elisabeth Fivaz, and Diane Philipp. And to my wonderful coders who dedicated, and continue to dedicate, so much time to this study: Ava Pouyandeh, Aaron Palachi, Cassy Wyers, Elina Gama Fila, Camille Cresencia-Mills, Caterina Parmegiani, Flora Roudbarani, and Bianca Ugiccioni - I am profoundly grateful for all your hard work.

To John Eastwood, for your incredibly insightful comments and guidance. You intuitively knew what direction this study should take, and what direction it should continue to take. I am so thankful that you were a part of this study. Rob Cribbie. My god. I don't know what I would have done without you. My profound thanks for sticking with me through the tumultuous process figuring out the statistical nuances of this study. To Mirisse, for the countless growthful opportunities that you have provided for me. Thank you for your light, guidance, and mentorship– both clinically and personally. You are the reason I am where I am.

To the brilliant, kind, patient, hilarious, and inspirational Robert Muller. Thank you for giving me the bragging rights of having the best supervisor in the world. You are the walking and talking embodiment of what an academic supervisor should be. Although never made explicit, I believe that you carefully take the time to get to know all of your students in order to understand how best to guide them. It is felt and appreciated by all of us. I am so grateful to have the opportunity to work under the guidance of someone who values clinical practice just as much as research. You are a two-in-one supervisor – something that I believe is rare, and possibly lacking, in this field. I didn't belong anywhere else. Thank you for allowing me to join your lab-family.

Speaking of our lab family: What makes the Trauma and Attachment Lab special isn't the research we do. It's the people who do it. Leah, Sheila, Sara, and Kris. Though we are all so different, there is a common thread that runs through all of us. It's almost magical and unapologetically weird. Sheila and Leah – thank you both for your constant support and encouragement, and for always making me smile! Sara. You are such a powerhouse force of nature and I feel so lucky to have hit the ground running with your guidance. I feel as if we are constantly on the same wavelength and I look forward to continuing to learn and laugh with you. Kristina. I have no words that can do justice to the impact you have had. Life is better with you in it. You never cease to inspire me – not only with your accomplishments, drive, and dedication to your work, friends, and family – but your incredible spirit completely moves me. You deserve all the good and the love and the light and the laughter that this world has to offer. Thank you for your constant support, contributions, encouragement, magical words, and love. And thank you for existing. You are simply the *prima facie* best (that's three now)! Finally, Bianca – I could not have done this without you. Thank you so much for all your hard word, friendship, and near-neurotic dedication to this project. You blew us all away and were absolutely the best student a girl could dream of!

I also want to acknowledge my former professors and mentors at Seneca College and, particularly, Brian Nakata, Johnathan Singer, Adam Phillips, and Naomi Kestenbaum. I got my start at this wonderful institution and learned the skills I needed to succeed in academia under the guidance of truly outstanding educators. Thank you for believing in me. To the York University faculty and staff, thank you for supporting me and making me feel welcome from day-one. Barb, Freda, and Lori, your support, advice, and guidance is so appreciated and valued!

Of course, to my amazing and brilliant cohort family – what a ride we've had together! Ben, Bianca, Elizabeth, Khad, Kyla, Rachael, Sam, and Tracy, and our adopted cohort members – Jenkin and Holly – Thank you all for being the best group of people in the world to share this experience with. I feel so lucky and I have so much love for each of you. It's a rare thing to find yourself in such a supportive and cohesive group where we thrive as a unit, celebrate each others' wins, and support each other when things get tough. It hasn't always been easy, but it is what it is! And we did it together.

Thank you to all my friends who encouraged me through this process and checked in for signs of life through the tail end of this project. My connection to you all is what keeps me going. You are my light at the end of the tunnel because I'll finally have time to see you all again! Special thanks to George for lapping me academically so that I can finally learn from your experiences. To Jessica for being the most hilarious and grounding presence in my life. And especially to Debi – you're the reason I did this in the first place.

Finally, thank you to my family. Momma and Papa Pfaff – thank you for your endless support, check-ins and spaghetti. Thank you to my father for being my role model for hard work and dedication and to my mother for being my role model for compassion and care. And my endless gratitude for my partner, Ilja, for your constant support, (occasional annoyance), and seemingly bottomless patience, understanding, and love. I love you all so much!

# Table of Contents

## List of Tables

# Introduction

The Adult Attachment Interview (AAI: Main & Goldwyn, 1984b) is a structured 60-90-minute interview designed to activate an individual's attachment system (Steele & Steele, 2008). In addition to being used in clinical practice (Steele et al., 2008), the AAI is a widely used research measure used to determine the attachment style of adults (Cassidy & Shaver, 1999). While this measure is considered the "gold standard" in identifying adult attachment, it is not without its limitations: it requires extensive training to become a reliable coder and is time- and labour-intensive to transcribe and score the interviews. Although modelled after the Strange Situation Procedure which is evaluated based on nonverbal content of infants' behaviours (Cassidy et al., 1999), the AAI is coded entirely based on the verbal content of adults' answers. Given that 60-65% of communication is nonverbal, (Burgoon, Guerrero, & Floyd, 2016) it is likely that the standard coding of the AAI, which omits any nonverbal signalling, misses valuable clinical information.

Identifying attachment style provides clinicians with clues to possible history, motives, and goals of an individual and is implicated in relational patterns and cognitions about the self and others (Cassidy et al., 1999). Most important, there is strong evidence that parental attachment styles are predictors of child attachment styles, creating a cycle of intergenerational attachment transmission (Steele et at., 2008). If the transmitted attachment style is maladaptive, there is potential for major implications in child attachment difficulties and developmental trajectory. As attachment is a critical entry point for prevention and early intervention, the current study centers on creating a reliable method of coding the AAI based on nonverbal and paralinguistic behaviour.

# Background Literature

## Attachment Theory

Attachment Theory focuses on the importance of early relationships with primary caregivers, which helps to develop the foundation for socialization and development (Laible & Thompson, 2007). Often regarded as the father of Attachment Theory, John Bowlby was influenced by the work of Charles Darwin and viewed attachment as an evolutionary imperative insofar as he believed that humans are biologically driven to seek proximity to their primary caregivers to ensure survival in the face of real or perceived threats, and that the child will always adapt to their caregiving context to obtain or maintain this proximity (Ainsworth & Bowlby, 1991; Bretherton, 1992). He also noted that children adapt to the caregiving context in which they find themselves in order to obtain or maintain proximity even when faced with suboptimal caregiving (Bretherton, 1992). Thus, attachment styles are patterns of relational responding, shaped by early caregiving experiences. Similarly, internal working models are formed through these early experiences and are mental representations, or frameworks, of how relationships should function and what to expect from them (Grossmann, Grossmann, & Waters, 2005; Main, 1983; Main & Goldwyn, 1984a). Thus, these internal working models result in the development of an individual's attachment style, which dictates, in part, how they will interact with significant or "important" others moving forward.

Bowlby suggested all humans have an adaptive attachment behavioural system designed to seek and maintain proximity to their primary caregivers(Bowlby, 1982; Cassidy, & Shaver, 2008; Mikulincer, Shaver, & Pereg, 2003). Proximity seeking is the primary strategy employed in this system and when attachment figures are available and responsive to this strategy, this leads to optimal functioning wherein the child experiences a sense of attachment security (Ainsworth et al., 1978; Bowlby, 1973; Cassidy & Shaver, 2008; Mikulincer et al., 2003). This sense of security alleviates distress in the individual and allows a person to increase their repertoire of internal resources which assist in coping, problem-solving, achieving adequate adjustment, and sustaining comfortable and supportive relationships (Cassidy & Shaver, 2008). When primary caregivers are not present, or when they are present but not responsive, the proximity seeking mechanism fails to relieve distress and achieve attachment security, and secondary attachment strategies become activated (Ainsworth et al., 1978; Cassidy, & Shaver, 2008; Mikulincer et al., 2003). Management of attachment-system activation and the alleviation of pain and frustration become the goals of these secondary attachment strategies in place of distress-regulation. In order to achieve these goals strategies of hyperactivation or de-activation are set in motion (Cassidy & Shaver, 2008; Mikulincer et al., 2003).

Hyperactivating strategies, employed when proximity seeking is possible, keep the attachment system chronically activated and on alert for threats, separations, and betrayals. This can result in an overdependence on relationships with a focus on minimizing distance from attachment figures (Cassidy & Shaver, 2008; Mikulincer et al., 2003). Deactivating strategies, employed when proximity seeking is not possible, keep the attachment system deactivated to avoid further instances of frustration due to unavailable primary caregivers. This strategy can result in patterns of detachment, apathy, and  distance from others (Cassidy & Shaver, 2008; Mikulincer, et al., 2003).

Building on Bowlby's work, Mary Ainsworth developed the Strange Situation Procedure (SSP), through which she was able to identify three patterns of attachment, resulting in the classifications of secure, anxious-ambivalent, and anxious-avoidant, the latter two representing attachment insecurity (Ainsworth, 1985a; Ainsworth, 1985b; Main & Solomon, 1990). Subsequently, as researchers began to observe that not all infant attachment behaviour fit neatly into these three categories, the fourth category of "disorganized" was added into the model (Main & Solomon, 1990). Secure attachments with caregivers provide children with a *secure base* from which to learn and explore their world and is thus the foundation for optimal development. Consistent attuned responses from the primary caregivers allow children to develop the belief that they can attain help, safety, and a sense of security from those closest to them in times of need. This sense of security influences the quality of all attachment relationships, in addition to social, cognitive, and emotional development and physiological functioning (Easterbrooks & Goldberg, 1990; Morrisset, Barnard, Greenberg, Booth, & Spiecker, 1990; Perry et al., 1995).

The Strange Situation Procedure paved the way for the development of the Adult Attachment Interview (AAI). The AAI was modelled after the SSP and became a leading tool used to measure adult attachment patterns (George et al., 1996; Steele & Steele, 2008). The various classifications of adult attachment are separated by secure and insecure attachment patterns (Ainsworth et al., 1978). Autonomous attachment is indicated as secure whereas dismissing and preoccupied attachment are indicated as insecure. The unresolved/disorganized pattern is a secondary classification, which can co-occur with any of the other three classifications, and commonly co-occurs with insecure attachment patterns (Greenberg et al., 1993; George et al., 1996). The specific predictive pairings of the adult classification and strange situation classification are autonomous-secure, dismissing-insecure avoidant, preoccupied-insecure anxious/ambivalent and unresolved/disorganized-disorganized (Benoit & Parker, 1994). Understanding a parent's attachment style may provide insight into how they will behave towards their child as well as how the child will act toward their parent. This is important information to have when understanding attachment relations and its impact on an individuals' wellbeing.

**Attachment disruptions.** Though perhaps not *prima facie* evident, there exists a strong link between attachment and relational trauma. According to the Dynamic Maturation Model (DMM: Crittenden, 2000), attachment responses are adaptive within their original caregiving contexts, but can become maladaptive when applied indiscriminately to all relationships. A child may scream to gain proximity and attention from a neglectful parent, but should this strategy continue to be applied widely to all subsequent relationships (e.g., screaming for attention), it becomes maladaptive.

Insecure attachment styles of children and adults result in the use of secondary attachment strategies when interacting with important others. Attachment theorists regard secondary attachment strategies as risk factors that can affect an individual's ability to cope during stressful situations and can lead to poor emotion regulation and adjustment (Cassidy & Shaver, 2008). Hyperactivating strategies are related to preoccupied-insecure ambivalent attachment style pairings and have been linked to attachment anxiety wherein the individual becomes overly distressed, anxious, and unable to effectively regulate their emotions (Cassidy & Shaver, 2008; Mikulincer & Shaver, 2003). Previous research has suggested that individuals who exhibit hyperactivating strategies present with lower levels of wellbeing and suffer from a variety of mental health issues such as depression, anxiety, personality disorders, eating disorders, substance abuse disorders, and conduct disorders (Cassidy & Shaver, 2008; Mikulincer & Shaver, 2003; Mikulincer & Shaver, 2012).

Similarly, deactivating strategies may lead to attachment avoidance and are associated with the dismissing-insecure avoidant pairing (Mikulincer & Shaver, 2003; Mikulincer et al., 2003). These individuals may build defenses in order to block and distance themselves from their emotions, resulting in an emotion avoidance defensive strategy which may lead to difficulty with adverse life events, in addition to potential social isolation and hostility (Cassidy & Shaver, 2008; Mikulincer et al., 2003). Although deactivating strategies may facilitate the dampening of conscious emotional experiencing and subsequent outward reaction, this avoidance has the potential to take a physiological toll on the individual, potentially leading to physical illness and sleeping difficulties, among other somatic concerns (Cassidy & Shaver, 2008; Waller, Scheidt & Hartmann, 2004). Previous studies have found that these individuals may suffer from various personality disorders, depression, anger, substance abuse and conduct disorders (Cassidy & Shaver, 2008; Mikulincer & Shaver, 2003; Mikulincer & Shaver, 2012).

Experiences of neglectful, dangerous, and inconsistent caregiving lead the child to rely on distorted cognition and affective responses which are inherently traumatizing as they leave the child with a lack of a secure base to turn to when there is a real or perceived threat (Purnell, 2010). These attachment disruptions can be understood as relational or attachment traumas, as the child grows up believing that they are “bad” (unworthy, unimportant, unwanted) and that those around them cannot be trusted to provide safely and security when needed (Ford et al., 2000; Foroughe & Muller, 2012; Zlotnick, Zakriski, Shea, & Costello, 1996). These attachment disruptions can lead to developmental delays, negative impacts on physiological functioning and brain structure, and socioemotional difficulties both concurrently and subsequent to the disruption (Gander & Buchheim, 2015, Green & Goldwyn, 2002; van IJzendoorn, Schuengel, & Bakermans-Kranenburg, 1999). Although many attachment strategies are effective in the moment, they can lead to poor outcomes in the long-term.

Trauma Theory emphasizes that trauma is understood not as the event itself, but rather the impact the event has on the individual (Storr, lalongo, Anthony, & Breslau, 2007). Attachment traumas can be understood within the context of Trauma Theory, in that traumatic experiences overwhelm the individual’s sense of self, resulting in an inability to create meaning through the integration of emotional responses, cognitions, and situational factors (Herman, 1997). Trauma can be classified as either Type I (simple and discrete) or Type II (chronic, ongoing) (Terr, 1991), though this provides a relatively narrow definition of the experiences. Recently, researchers have begun to identify what they call “little t” traumas, which do not necessarily require physical threats to be present, as has been the case historically, but are conceived as “ego-threats” such as abandonment, neglect, or emotional abuse (Herman, 1997; Lyons-Ruth, Dutra, Schuder, & Bianchi, 2006). Although these “little t” traumas are not as overt in their etiology or symptomatology, they are the most common type of childhood trauma and often involve relational disturbances within the family context. Lyons-Ruth and colleagues (2006) explain that childhood relational traumas do not often stand out as salient, nor do they often result from physical abuse. Rather, they are a consequence of negative affective signals and/or lack of parental availability and responsivity. In this way, we can understand these intrafamilial traumas as being truly relational and attachment based.

**Impact.** There is a consistent pattern in the literature that links relational traumas to increased risk for adverse physiological, psychological, and developmental outcomes. Roughly one third of Canadian adults report some form of childhood trauma (Afifi et al., 2014) and research suggests that there is a strong link between trauma histories and mental health difficulties (Breslau & Kessler, 2001). Interpersonal (relational) trauma results in greater psychological harm than most other events, including events causing physical harm, due to the violations or betrayals committed by a "trusted-other" (Kessler et al., 1995).

The concepts of equifinality and multifinality posit that there are simultaneously many different paths that may lead to the same outcome and that the impact of a given event may manifest differently for different individuals, respectively (Cicchetti et al., 1995). Thus, the implications for both scientists and clinicians are that it becomes quite difficult to: 1) identify "hidden" relational traumas and 2) predict how those traumas may affect diverse areas of functioning. Indeed, taking a biopsychosocial stance to relational trauma (Sameroff, 2010), we discover even more specific ways in which symptoms can manifest within different domains of functioning. Physiological responses to trauma include increased amygdala activity which activates the sympathetic nervous system, and results in the flight-or-flight response and increased cortisol production, inhibiting fear-reducing processes in the cortical areas of the brain. This process is directly related to hyperarousal and the classically conditioned generalization of the fear response (Baranowsky & Gentry, 2015). We may also find difficulties with executive functioning, as these elevated cortisol levels interfere with processes in the prefrontal cortex and the hippocampal region (Bomyea & Lang, 2015).

The psychological consequences of trauma have been well documented and include somatic difficulties, mood dysregulation, substance abuse, suicidality, and dissociative symptoms (McFarlane, 2010; Chu & DePrince, 2006). Similarly, relational trauma prompts paradoxical social behaviours whereby an individual both seeks out, and withdraws from, close relationships due to ruptures in his/her sense of self and trust in others (Herman, 1997). This instability invariably elicits a sense of loss, loneliness, and disconnection in areas of social and emotional functioning. These findings paint a complex picture of trauma symptoms when investigating if, when, and how an event will affect an individual, and how these symptoms may manifest. The picture becomes even more complex when considering that both trauma and attachment difficulties do not begin and end with the individual, but are transmitted intergenerationally (van IJzendoorn, 1995).

Attachment styles born from relational trauma are highly predictive of both the attachment style of a person's caregiver, as well as the subsequent attachment style of their child through a sequence of intergenerational transmission of attachment patterns (Behrens, Hesse, & Main, 2007; Tarabulsy et al., 2005). This link likely results from a history of relational trauma within the parent influencing attachment insecurity in the child. For example, an avoidant attachment response may become triggered when that child becomes a parent him/herself, as the parent-child relationship demands closeness, intimacy, vulnerability, and interdependence. These demands can be difficult for a parent with avoidant attachment tendencies due to their avoidance of closeness and, no matter how well-meaning, these tendencies have the potential to cause attachment disruptions in the child (Foroughe & Muller, 2012).

**Nonverbal Behaviour**

Nonverbal behaviours may provide important clues to possible attachment disruptions as they display a wealth of affective information within the context of therapy, often outside of the patient's awareness (Gabbard, 2007). Attachment difficulties manifest in many different ways and stem from differing early experiences. As nonverbal behaviour is an extension of personal communication style (Burgoon et al., 2016), differential experiences may manifest differently through nonverbal behaviours and enactments within the therapeutic context.

Though there exist many standards of classification for coding nonverbal behaviour, this current study borrows facets from Shea's (1998) model which characterizes 3 areas: Kinesics, proxemics, and paralanguage. Kinesics refers to how the body moves an includes gestures, eye movements, body movements, and facial expressions. Proxemics refers to the use of space (how close you are to another person), and paralanguage includes pauses, verbal fillers, and non-responses. The clinical importance of nonverbal behaviour has been noted, linking together Shea's kinesic elements to domains investigated on a mental status exam (Foley & Gentile, 2010).

As with attachment theory, nonverbal communication finds its roots in evolutionary theory, such that certain nonverbal and emotional tendencies provided survival advantages (Patterson, 2003). Scientific literature provides a wide breadth of information regarding the history, functions, and interpretation of nonverbal communication. However, germane to the current discussion, there exists a subset of studies on the functions of nonverbal behaviour in emotional expression and deception.

**Emotional expression.** According to Ekman's Basic Emotions Approach (Ekman, 1971; Izard, 1977; Tomkins, 1962), all humans experience and express emotions similarly using six universal human emotions: joy, anger, sadness, fear, surprise, and disgust-contempt. Ekman also posited that much of emotional expression is spontaneous and nonverbal (Ekman, 1971). Further, it has been suggested that each universal emotion has an associated action tendency allowing individuals to adapt to their environment, which supports the idea of emotions being a motivating force for survival (Lazarus, 1991). Building on Ekman's early work, the Neocultural Perspective posits "display rules" that moderate the expression of emotion in different social situations, based on cultural norms (Casey & Fuller, 1994). At first, emotions are expressed freely but as children grow older, they learn that not all affective expressions are appropriate in all situations. The result is that emotional expression can sometimes become "blocked" by display rules such that all emotional expression becomes a combination of what we truly feel and what we feel comfortable sharing with the world. Additionally, Darwin (1998/1872) believed that emotions were evolutionarily adaptive and described the chain of events as such: "Actions, which were at first voluntary, soon became habitual and […] hereditary, and may then be performed even in opposition to the will" (p. 356). Darwin's statement implies that, regardless of intention, emotional expression will "leak out", despite the display rules in place.

**Deception.** Although we often conceptualize deception as a verbal act (saying something that is false), many clues to deception can be found within nonverbal behaviours (Burgoon et al., 2016). Research has demonstrated that deception is not always done consciously, but that individuals may unconsciously disguise their true feelings if they are too difficult to be processed consciously (Burgoon et al., 2016). The leakage hypothesis states that physiological, emotional, and cognitive processes are involved in deception and that these processes will produce unconscious external displays of individuals' true feelings, as the more emotionally heightened one becomes, the more difficult it is to conceal unwanted emotions (Ekman, 2009).

As we have seen, trauma and attachment difficulties often remain covert and sometimes outside of awareness. If true emotions will likely leak out in the presence of emotional activation, it is important that clinicians have a method of identifying, not only the emotions that the individual wishes to portray, but the emotions and nonverbal action tendencies that "leak" in those moments. As the AAI serves to activate the attachment system, it makes sense that there be a method to investigate nonverbal behaviours within this context.

## The Current Study

The intergenerational transmission, overlapping symptoms, and heterogeneous symptomatology pose significant obstacles in the assessment and treatment of attachment and trauma-related difficulties, especially when considering that emotionally difficult experiences are often suppressed. It is critical to identify these difficulties because if the symptom is addressed without acknowledging the underlying difficulties, it is likely that the patient will not improve or will relapse (Baranowski & Gentry, 2015). Moreover, the longer an individual continues down a maladaptive path, the more difficult it becomes to return to optimal or normal development (Cicchetti & Cohen, 1995). The intergenerational transmission of these difficulties provides clinicians with a unique opportunity and critical entry point for early intervention/prevention of child mental health difficulties by first identifying these patterns in the parent.

This current study represents the first phase in a larger stepwise research program designed to validate a novel nonverbal coding instrument against the standard coding system of the AAI. The result of this research program will serve to better align the AAI with the foundations of attachment theory, increase the AAI’s utility in clinical practice, and build upon previous studies in this area by addressing methodological gaps.

As previously stated, though the AAI was modelled after the Mary Ainsworth’s Strange Situation Protocol (SSP: Ainsworth & Bell, 1970), it does not incorporate an examination of adult nonverbal behaviour analogous to the examination of nonverbal infant behaviour in the SSP. While the determination of infant attachment style is necessarily contingent on an analysis of nonverbal behaviour due to a lack of meaningful linguistic behaviour at this stage of development, Ainsworth and Bell also took great care to align infant attachment behaviours with the ethological-evolutionary foundation of attachment theory by examining separation and reunion behaviours directly (Ainsworth & Bell, 1970). Additionally, the biological function of infant-mother attachment is highlighted in their original paper and is in line with Bowlby’s proposition of attachment as a behavioural system, which is borrowed from ethology (Mikulincer & Shaver, 2003). While the AAI indeed claims strong psychometric properties (George, Kaplan, & Main, 1996), the target of observation (relying on verbal content alone) deviates from the behavioural systems underpinnings of both the SSP and from attachment theory itself. In fact, while the AAI has several alternatives to its original coding system (Reflective Functioning, and the AAI Q-Sort), none of these adaptations involve assessing nonverbal behaviour (Fonagy et al, 1998; Kobak, 1993). This research program merges the adaptive behavioural facets of nonverbal behaviour and emotion, and physiological responses to trauma, with the ethological foundations of the SSP and attachment theory by creating a novel coding instrument for the AAI based on nonverbal behaviour and taking into consideration the biological function of these behaviours.

Although numerous researchers have highlighted the importance of attending to nonverbal behaviours within a clinical context for both the client and the therapist (Ciarochi, Robb & Godsell, 2005; D'Agostino & Bylund, 2014; Paniagua, 2004; Philippot, Feldman, & Coats, 2003; Ramseyer & Tschacher, 2011) very little that has been done to create a systematic way of identifying these behaviours within a clinical context. Few methods have been developed to examine the nonverbal behaviours of specific clinical populations to assess therapeutic outcomes, such as the Nonverbal Interaction Coding System for patients diagnosed with Bipolar Disorder (Simoneau, Miklowitz, & Saleem, 1998) and the Ethological Coding System for patients diagnosed with Schizophrenia (Troisi, 1999). However, no such coding system has been developed to identify attachment related difficulties. Further, the drawbacks inherent in self report methods (Cervone & Pervin, 2013; Heppner, Wampold, Owen, Wang, & Thompson, 2016) and the prevalence of deception in the psychotherapy room (Blanchard & Faber, 2016; Carlson & Kottler, 2011; Gediman & Lieberman, 1996) suggest that identifying nonverbal cues to deception and attachment related difficulties is a worthwhile endeavour.

There exists a paucity of research specifically investigating nonverbal behaviour within the context of the AAI. Roisman and colleagues (2004) examined individual differences in emotion regulation during the AAI using measures of physiological responses and facial expressions and found that securely attached adults demonstrated higher rates of emotional cohesion and insecurely attached adults demonstrated higher rates of discrepancy between their "lived" and "told" stories. However, the only nonverbal indicators investigated in this study were facial expressions, to the exclusion of all other kinesic behaviours. Karlsson (2005) created a tool to investigate both facial expression and kinesic behaviour during the AAI and found that he could reliably distinguish between secure and insecure attachment styles. The limitation of this study was that he did not code the clinical sessions for nonverbal events, rather provided his overall impressions, and did not enlist a second coder for reliability. Lambert (2012) examined congruence versus incongruence of body language when the AAI was administered and found that one could categorize attachment style based on congruence vs. incongruence in nonverbal behaviours in relation to the individual's narrative. Once again, this study did not involve a measure of inter-rater reliability. In addition, none of these studies included a baseline period through which they could compare nonverbal behavior observed during periods of attachment activation and nonverbal behavior observed during periods of neutrality.

The current study addresses the limitations of prior work in three important ways: 1) by establishing a nonverbal baseline to assess nonverbal behaviour in periods of neutrality versus periods of emotional activation, 2) by establishing inter-rater reliability of the coding instrument, and 3) by incorporating kinesics, proxemics, and paralanguage, in line with the nonverbal framework presented above.

**Research Questions, Goals, & Proposed Processes**

With a dialectical constructivist approach in mind (see Greenberg & Angus, 2004 for a review), the current phase of this research program underwent careful planning and consideration, with three general exploratory questions guiding this work: (1) Is it possible to better align the AAI with the theoretical/ethological foundations of the SSP based on nonverbal behaviour? (2) Can specific nonverbal behaviours during the AAI be comprehensively identified, catalogued, and aligned with the extant literature in the area of nonverbal behaviour? (3) Can the aforementioned nonverbal behaviours be reliably identified and coded by observing the video-recordings?

In order to empirically address these questions, the goals of the present study were twofold. The first goal was to identify the common nonverbal behaviours exhibited during the AAI and develop a systematic coding manual for identifying nonverbal behaviours during this interview. In keeping with a dialectical constructivist framework (Greenberg & Angus, 2004), The Adult Attachment Interview Nonverbal Behaviour Manual (AAI-NVB) evolved from the recursive processes of video observation and extensive review of the nonverbal literature. Building upon the gaps in previous studies of this nature (Karlsson, 2005; Lambert, 2012; Roisman, 2004), it was equally important to incorporate the theoretical frameworks and categorizations of nonverbal behaviour documented in previous literature. The rationale for this decision was made in order to assure that the nonverbal behaviours observed in our video recordings could be organized in such a way that they would fit into these pre-existing, empirically validated categories. Finally, the observation process and manual development itself was guided a series of 5 steps outlined by Floyd, Baucom, Godfrey, and Palmer (1998). These steps relevant to the first goal of this study included (1) Cataloguing relevant behaviour, (2) Selecting a unit of observation, (3) Creating coding categories, and (4) Developing a coding manual[1]. Later iterations of this manual were subject to change based on the results of the data analysis and feedback from the coding team.

In order to address the question of whether these nonverbal behaviours can be reliably identified, the second goal of the current study was to establish inter-rater agreement for the nonverbal behaviours outlined in the coding manual. The coding process was guided by Harrigan's 2013 chapter titled *Methodology: Coding and studying nonverbal behaviour* (Harrigan, 2013) to ensure that common, but frequently overlooked difficulties in coding observational data were circumvented to the best of our abilities. A team of coders were trained on the coding instrument and were all tasked with coding the same set of videos, one at a time. This process was adopted to allow this author to review the data output from each coder, analyse the data received, and provide feedback and remediation to each coder on targeted areas of individual improvement. This process will continue until optimal levels of agreement are reached. Once these levels are attained, each coder will be dispatched to complete their coding independently, and the resulting data will be used for the next phase of this project.

## Methods

The current study uses Adult Attachment Interview video data collected as part of a larger study on Emotion Focused Family Therapy and adult attachment. The study took place in Toronto, Ontario through the Family Psychology Clinic (formerly known as the Kindercare Pediatrics Psychology Clinic) and the Trauma & Attachment Lab at York University. Recruitment began in May 2016 and continued until data collection was completed in September 2018. Ethics approval was obtained by the Human Participants Review Subcommittee at York University.

[1] The fifth step outlined in this guide is concerning the establishment of inter-rater agreement, which is relevant to the second goal of this current study.

**Participants**

A sample of 416 self-referring caregivers seeking treatment for their children were recruited from the above-mentioned pediatric psychology clinic in Toronto. A broad range of child presenting concerns were reported at the time of registration including various psychiatric disorders and mental health difficulties. There were no restrictions with regard to the age of the parents or children, and the only exclusion criteria was active psychosis. Participants were randomized into two conditions. One set of caregivers received the Adult Attachment Interview and the other did not. For the purposes of this study, only participants in the Adult Attachment Interview condition were included. The interviews were videotaped, and video data was selected based on the procedure listed below. For the purposes of establishing coder reliability, videos were, and will continue to be, selected at random until coder reliability is established. Once this occurs, the remainder of the videos will be used for the next phase of this research program, which is currently outside the scope of the current study. Of the videos selected for the larger research program (n=42), below is a brief demographic description of the participants.

Generally, participants came from diverse socio-economic and educational backgrounds. The sample consists of 28 females and 14 males with an average age of 43. Participants of North American birth represented 70% of the sample, with the remaining participants born abroad in such counties as South Africa, Brazil, UK, Prague, Russia, and China. Of those born abroad, the average age at which participants immigrated to Canada was 15 years-of-age. Verbal, written, and audio-visual consent was obtained from each participant, with the option of declining video recording and opting out of the study at any time. The current study includes 3 subjects randomly selected from the subset of 42 participants. Two of these subjects were male, one female, with an average age of 36 years. All three subjects were born in Canada.

**Instrument**

The Adult Attachment Interview (AAI: George et al., 1996; Main & Goldwyn's, 1984) is a semi-structured clinical interview, lasting between 60 and 90 minutes, which determines attachment style in adults. (George et al., 1996; Steele & Steele, 2008) The interview focuses on early attachment experiences with caregivers and the influence this has had on their current state of mind and life experiences. Participants are asked to describe their childhood experiences with their parents or any other significant attachment figures as well as any considerable losses they may have experienced (George et al., 1996).

The AAI is also used widely as a tool in clinical practice in order to reach such goals as: facilitating therapeutic alliance, setting the agenda, revealing losses and traumatic experiences, identifying reliance on defensive processes, understanding the influence early relationship patterns have on adults mind and behaviour, as an aid in lawful matters, observation of reflective functioning, use in the supervision and training of clinicians, and assisting in assessing therapeutic outcomes (George et al., 1996; Steele & Steele, 2008). For the purposes of this study, AAIs were administered, and video recorded in a private room at community psychology clinic by a trained AAI administrator. Three of the twenty-one questions on the AAI will be used for the purpose of coding. These three questions were selected because of their ability to activate the attachment system, as they focus on the participant providing descriptions of their childhood relationship with their primary caregivers. This is accomplished by having participants provide 5 adjectives to describe their childhood relationship with each of their caregivers and then describing memories associated with those particular adjectives. Question 3 specifically, which asks the participant to provide 5 adjectives to describe their childhood relationship with a focus on their mother, has been found to be the most activating question, thus this question was ideal to code as it would likely provide a wide array nonverbal behaviour useful for classifying attachment (Karlsson, 2005). A two-day training for administrators was conducted on the clinical applications of the AAI and how to administer the interview to caregivers within a private practice setting. A clinical psychologist trained on the AAI conducted the two-day training. Hesse (1999) established that the psychometric properties of the AAI were satisfactory. The average inter-judge reliability was 82% with a mean Kappa of .71 based on an informal survey of available AAI publications before 1999.

## Procedure

**Video selection.** The first round of video selection was completed largely based on the quality of the videos obtained. Lower quality videos, or videos in which the full body of the participant was not in constant view were excluded, and the remaining videos were separated on the basis of whether the participant was sitting on a chair or on a couch. Participants who were seated on a couch were selected for the current research program, as it would allow for a variable to be created based on seating location (e.g., closer to the exit, closer to the window, or seated in the middle of the couch).

**Coding team.** Coders were recruited by way of referral or email communication through York University. Eight coders were recruited and underwent multiple rounds of training. These individuals have similar educational backgrounds: 4 are currently completing their undergraduate degrees in psychology, 3 have completed their undergraduate degree in psychology, and 1 has completed a master's degree in psychology. The age of the coders ranges from 21 to 29 years old with 80% female and 20% male. Coders of Canadian birth represent 70% of the group, with the remaining 30% born outside of Canada, but have been residing in Canada for most of their lives.

**Training.** Individual and group training will be provided to all coders. These coders received a manual outlining the various factors and coding protocol, with each factor discussed in detail and examples of factors provided. Each time analyses for inter-rater reliability (IRR) are completed, thorough feedback is provided to the coders based on the individual results of the coders as well as the global results of the overall coding team. Further training will take place depending on the strengths and weaknesses of the coders and additional videos will be coded. This process will be repeated until acceptable levels of IRR are achieved.

**Coding procedure.** Videos were selected using a randomizer tool and the supervisor and/or project manager assigned deadlines for these videos. The videos themselves are provided to coders via an encrypted USB key and the time codes for the sections of the videos to be coded are emailed to them. The Adult Attachment Interview length ranges from 30 minutes to 2 hours and coders only code questions 2, 3, and 4, in addition to the baseline. Two different events are coded: point events and state events. Point events refer to behaviours that occur at one point in time whereas state events have a duration, and therefore must be coded twice – once when the behaviour begins, and once when it ends. The following sections describe the baseline period, in addition to the three AAI questions that have been selected for coding in this study.

***Baseline.*** The baseline period refers to the section of the video that encompasses everything between the time the participant sits down on the couch until the interviewer begins the AAI administration. The baseline includes questions and comments involving innocuous topics like parking, traffic, and weather. The interviewer also asks about any chronic pain conditions and/or temperature in the room during the baseline in order to control for environmental variables that may confound nonverbal behaviour during the interview (Foley & Gentile, 2010). The use of the baseline will allow for the comparison of a participant's typical nonverbal behaviour and their nonverbal behaviour when their attachment system is activated during subsequent stages of this study.

***Question 2:*** This question asks participants retrieve some of their earliest memories of their relationship with their parents (George, Kaplan, & Main, 1996).

*I'd like you to try to describe your relationship with your parents as a young child, if you could start from as far back as you can remember.*

***Question 3:*** This question involves two parts; one focused on descriptive words to describe the relationship with their parent and the other asks the participant for a memory which accompanies each descriptive word. This question provides an overall impression of the relationship with the specific parent discussed and it generally sets the tone for how the interview will proceed (Steele & Steele, 2008)

*Now I'd like to ask you to choose five adjectives or words that reflect your childhood relationship with your mother/father starting from as far back as you can remember, but say, age 5 to 12 is fine. I know this may take a bit of time, so go ahead and think for a minute...I'll write each word down as you give them to me, then I'll ask you why you chose them.*

*Okay, that's great. The first word you used to describe your relationship with your mother/father was _____(OR: you used the phrase _____ to describe your relationship with your mother/father).*

***Question 4:*** This question is identical to Question 3, however the focus of the descriptive words and memories is targeted to the other caregiver.

*Now I'd like to ask you to choose five adjectives or words that reflect your childhood relationship with your mother/father starting from as far back as you can remember, but say, age 5 to 12 is fine. I know this may take a bit of time, so go ahead and think for a minute...I'll write each word down as you give them to me, then I'll ask you why you chose them.*

*Okay, that's great. The first word you used to describe your relationship with your mother/father was _____(OR: you used the phrase _____ to describe your relationship with your mother/father).*

Additionally, qualitative scales will be completed and used to assess the overall impressions of each video in order to maintain consistency between the current study and previous investigations into nonverbal behaviour (DePaulo et al., 2003). These scales are completed online by each coder and include items such as overall involvement, engagement, and logical structure.

**Coding phases.** The coding of the videos occurs in various phases. Phase 1 consists of coding 5 videos based on 17 preliminary factors, which were devised from anecdotal clinical evidence. The factors were then reframed based on further literature review and consultation. To date in Phase 2 of the current study, three subjects have been coded based on sequentially revised versions of the AAI-NVB Manual. Videos were coded one at a time, with a period of analysis and feedback between each coding phase.

### Materials

Phase 1 coding was completed using manual coding sheets. The coders inputted the various timepoints at which an event occurred, frequency of the event, and the duration (in seconds) when coding state events. Beginning with Phase 2 coding and onward, coding was completed using the Behavioural Observation Research Interactive Software (BORIS: Olivier & Marco, 2016), which allowed coders to code with minimal keystrokes and automatically timestamps the coding output for each factor coded. The data was then exported into an MS Excel spreadsheet for cleaning and analysis.

Sony video cameras were used to video record the AAIs with participants. The cameras were set up on a tripod to get a full body image of the participant in a private room in the clinic. Once the interviews were completed the videos were saved on an encrypted drive as well as on a password protected Hewlett-Packard computer.

### Ethical Considerations

This research has been approved by the Human Participants Review Subcommittee at York University. Verbal, written, and audiovisual consent was obtained from each participant, and all were given the option to withdraw from the study or opt out of the video recording. All hard-copy data are held in a secure filing cabinet within a locked office and all digital data is stored on an encrypted and password protected drive. Due to the personal nature of the AAI itself, some participants find it difficult or distressing to speak about some of their early childhood experiences. As such, all AAI administrators are trained in a distress protocol and risk assessment, and there was always a licenced clinical psychologist on site in the event that further debriefing is required.

## Manual Development

The Adult Attachment Interview Nonverbal Behaviour Manual (AAI-NVB) originally evolved from anecdotal clinical evidence, video observation, and an extensive review of the nonverbal behaviour literature. The literature review ensured that the nonverbal behaviours observed in the video recordings could be organized such that they would fit into pre-existing, empirically validated factors and categories. While assigning meaning to specific nonverbal behaviors is outside the current scope of this study, it was important to make certain that the behaviors selected for this manual had been previously investigated in some manner such that, in future studies, it would be possible to compare our findings with previous findings in the literature.

The manual development period was a discovery-oriented process that included systematic, intensive observation and description rather than direct hypothesis testing. The observation process and manual development itself was guided by a series of 5 steps outlined by Floyd, Baucom, Godfrey, and Palmer (1998): (1) Cataloguing relevant behaviour, (2) Selecting a unit of observation, (3) Creating coding categories, (4) Developing a coding manual and (5) Inter-rater reliability. The 5 steps outlined below occur in a cyclical manner and later iterations of the manual are subject to change based on the results of the data analysis and the feedback from the coding team. The objective of developing the AAI-NVB Manual was to create an observer-based manual that allowed for the identification of nonverbal behaviours that would enable researchers to reliably code various nonverbal behaviours from video recorded AAI's.

**Step 1: Cataloguing Relevant Behaviours**

During the initial step of developing a novel coding system, investigators begin by developing an exhaustive list of relevant behaviours. This process can be completed in two ways: The first involving taking initial observations of any relevant behaviours and creating an ethogram using a theta statistic to increase the probability that the behavioural repertoire has been adequately sampled, and the second involving listing any relevant behaviours gleaned from research, theory, and experience. The first method is typically reserved for observational research involving animals. The second method, more common to observational research involving humans, requires the subsequent step of observing the human subjects and creating behavioural categories based on the information gathered from the previous research phase. Floyd and colleagues (1998) also encourage the widening, tightening, creation of new categories based on these observations.

As stated above, the first goal of this study was to identify the common nonverbal behaviours exhibited during the AAI and develop and systematic manual for coding these behaviours. In light of the exploratory nature of this goal, it was important that the initial stages of this project – the identification of nonverbal behaviour in video-recorded AAI sessions – evolved from intensive and systematic observation and description, rather than hypothesis testing. It was imperative to observe the video recordings directly to see what emerged from the AAI sessions. However, it was also important for this process to be guided by pre-existing theoretical frameworks found within the scientific literature. Therefore, by following the recommendations of Floyd and colleagues (1998), and in keeping in line with the goals of the current study, the relevant behaviours were catalogued through consultation with nonverbal and attachment experts, reviewing literature for relevant theories and behaviours, and careful observation of video-recorded AAIs.

In the pilot phase (Phase 1) of this project, the behaviours selected for coding were based solely on the anecdotal clinical evidence from two prominent attachment-informed psychologists. These original behaviours were: *Latency, Pauses, Lack of Memory, Generalizations, Body Shifts, Verbal Diversions, Laughter, Verbal Fillers, Humor, Protective Objects, Dismissiveness, Incongruence, Dissociation, Self Soothing, Gaze Aversion,* and *Micro Expressions.* The observation of these nonverbal behaviours during the AAI marked the foundation of this study. However, after a pilot round of coding and an examination of the nonverbal literature, it was determined that many of these original behaviours were not specific enough to provide any meaningful information regarding the subject's internal state. For example, the *Body Shifting* behaviour encompassed all manner of body movements, and was therefore more suitable as a potential category, rather than a stand-alone behaviour. Thus, in order to create an exhaustive and scientifically supported taxonomy of nonverbal behaviours observed during the AAI, we began the task of validating the existing behaviours and adding to our taxonomy through the process of literature review, consultation, and observation.

**Theoretical constructs.** This process began by first becoming familiar with the scientific landscape of nonverbal behaviour. Burgoon and colleagues (2016) outline the myriad of functions for nonverbal behaviour and the many ways that these behaviours have been interpreted in the scientific community, from cultural interpretations, studies of friendships and intimate relationships, emotional intelligence, message production and processing, to social cognition. Of particular interest was a subset of knowledge that focused on the function of nonverbal behaviour as it pertains to emotional expression, deception, and even more specifically, attachment.

Researchers have distinguished between emotional experiences and emotional expression (Burgoon et al., 2016) such that experiencing refers to the internal state of the individual and expression refers to the external and interpersonal display of emotion. According the Basic Emotions Approach (Ekman, 1971; Izard, 1977; Tomkins, 1962), all humans experience and express emotions similarly, and most emotional expression is spontaneous and nonverbal. Darwin (1998/1872) believed that emotions were evolutionarily adaptive and described the chain of events as such: “Actions, which were at first voluntary, soon became habitual and […] hereditary, and may then be performed even in opposition to the will” (p.356). Building on this work, the Neocultural Perspective posits “display rules” that moderate the expression of emotion in different social situations based on cultural norms. At first, emotions are expressed freely but, as children grow older, they learn that not all affective expressions are appropriate in all situations (Casey & Fuller, 1994). The result is that emotional expression can sometimes become “blocked” by display rules, resulting in all emotional expression becoming a combination of what individuals are truly experiencing internally, and what they are comfortable sharing externally. It was for precisely this reason that we felt it necessary to incorporate nonverbal communication and deception into our review of the literature.

Ekman and Friesen (1969b) makes an important distinction between self- and alter-deception. While alter-deception is the attempt to conceal information from another individual, self-deception is the attempt to conceal information from the self and the ego is therefore the target of the deception. However, he is careful to note that, with regard to nonverbal cues to deception, there does not exist a fundamental difference between self- and alter- deception as the internal mechanisms are the same during both processes. According to Ekman, these processes can consist of either inhibition, wherein the individual omits certain nonverbal messages, and simulation, wherein the individual fills the gaps left by the omitted messages by replacing them with a false representation of their internal experience. This simulation also creates a barrier against the breakthrough of the inhibited behaviour, otherwise known as “nonverbal leakage”.

Researchers posit several theories regarding nonverbal behaviour and deception that serve to guide research endeavors. Zuckerman (1981) states that research should focus on the types of behavioural cues that are associated with thoughts, feelings, of physiological processes that are likely to occur when an individual is being deceptive. This may include generalized arousal, feelings of guilt and fear, increased cognitive processing, and attempted control.

Similarly, Ekman and Friesen (1696b) describes that "thinking and feeling cues" should be the focus of research on deception detection. Other orientations focus on a more interpersonal landscape, such as the Communications Perspective by Buller and Burgoon (1996) which examines interpersonal processes for signs of deception. These researchers believe that individuals who are being deceptive will look for signs of suspiciousness in those around them and change their behaviour accordingly. They claim that these patters of behaviour change will be based on expectations, motivations, goals, and relationships, and therefore patterns of deception will vary from person to person. Additionally, they hypothesize that deceit is more likely to be detected if the underlying motivation is for instrumental purposes, rather than for relational purposes or identity protection. In a similar fashion, DePaulo (1992) posits the Self-Presentational Perspective, wherein individuals are most frequently deceptive about their feelings, preferences, attitudes, and opinions, and that the rewards sought from these deceptions are typically of a psychological nature. At this point in the study, we aimed to gain a better understanding of the theoretical landscape rather than subscribe to one particular orientation. Additionally, these theoretical perspectives were instrumental in guiding the consideration of potential relevant nonverbal behaviours implicated in deception.

With regard to additional considerations and methodology in detecting deception, Vrij (2008) outlines three reasons why deception is likely to go undetected. The first involves motivational factors wherein individuals are not always motivated to detect lies because they may not know what to do if the truth is revealed. Secondly, the individual engaging in deception may employ the use of countermeasures that make the deception more difficult to detect. Thirdly, despite the copious amount of research that has been conducted in this area, there is limited evidence to suggest that there are universal deception signals and it is unlikely that different individuals will demonstrate the same patterns of behaviour while deceiving. However, Vrij identifies a number of universal factors that are thought to influence deception cues. These include emotional reactions and arousal state, cognitive effort, and attempted behavioural control. For example, Ekman and Friesen (1969b) states that deception is typically associated with guilt, fear, or delight. These emotional states may elicit physiological responses that may "leak" out through nonverbal displays.  Likewise, as deception requires a more complex set of cognitive steps and increased cognitive involvement, such as monitoring their own reactions as well as the reactions of the other person. Additionally, Vrij (2008) points out that, despite the individual's best efforts, some behaviours are beyond control as they are intrinsically linked to strong emotions which are often difficult to suppress. Despite gender, cultural, and individual differences, there does not appear to be theoretical reasons or empirical evidence to suggest that deception behaviours would vary across these domains. Finally, it is recommended to consider target behaviours in combination, rather than individually, and to compare target behaviours to an individual's baseline (Vrij, 2008). The ultimate aim of deception detection is to notice behavioural changes as they occur and then attempt to make meaning of these changes.

The final theoretical realm that was targeted in this review was the intersection between nonverbal behaviour and adult attachment. There exists a paucity of published work surrounding nonverbal behaviour and attachment, however there is some evidence to suggest that those with secure attachments demonstrate moderate levels of nonverbal involvement, as compared to higher levels of nonverbal involvement by individuals with preoccupied attachment and lower levels of nonverbal involvement demonstrated by those with avoidant attachment (LePoire, Shepard, & Duggan, 1999). These researchers also found that the attachment style of their romantic partner had a moderating effect on this outcome. With regard to the quality of nonverbal involvement, Becker-Stoll and colleagues (2001) found that those with secure attachment exhibit more frequent open and positive nonverbal displays, while those with a dismissive style exhibited more frequent communication inhibiting behaviours. Specifically, females displayed more passive dominance while males displayed more anger. These results mirror those found by George and Main (1979) which demonstrated that abused females were more likely to display more passive-aggressive nonverbal behaviours whereas abused males displayed higher instances of nonverbal anger and aggression. Finally, Dozier and Kobak (1992) found that individuals employing deactivating strategies demonstrated increased skin conductance, which supported the authors' notion that individuals engaging in deactivation experience heightened conflict or inhibition during periods of attachment threat.

**Selecting behaviours.** While a review of the scientific literature regarding the theory of nonverbal behaviour aided in the selection of behaviours that may be particularly relevant to our study, the constrains of the study itself dictated which classes of nonverbal behaviours were to be excluded on the basis of feasibility. For example, in the *Handbook of Nonverbal Behaviour*, Burgoon and colleagues (2016) outline the major categories of nonverbal codes as being Kinesics (including gaze behaviour and facial expression), Vocalics, Haptics, Proxemics, Chronemics, and Artifacts. However, the present study *prima facie* excludes the study of touch and space (Haptics and Proxemics)[2] due to the fact that the subject remains seated throughout the interview and there is no expectation that the subject would have cause to physically touch the interviewer. Similarly, movements that happen *in the moment* were of primary interest, therefore chronemics and artifacts (time and objects) were not immediately relevant. More specifically, a number of ocular and vocal behaviours were not feasible for investigation in the present study due to lacking the sophisticated equipment needed to examine these areas. However, several ocular and vocal behaviours were examined in the present study, with particular attention paid to paralinguistic behaviour. While numerous sources were employed in this selection process (see Reference section of the AAI-NVB Manual in Appendix F), certain seminal studies, discussed below, were highly relevant and influential for our manual development.

While the work of Paul Ekman provided additional examples of behaviours to exclude on the basis of relevance and feasibility, his research was imperative with regard to the process of selecting relevant behaviours due to the immense amount of work that he has done in the field of nonverbal behaviour, emotion, and deception. Ekman and Friesen's 1969 paper titled *The Repertoire of Nonverbal Behaviour* provided an excellent framework through which to observe and catalogue relevant behaviours, as he outlines five Kinesic categories. Firstly, *Emblems* describe behaviours that are most closely associated with verbal communication and are typically exhibited with awareness and intentionality. These behaviours have a direct verbal translation, such as "yes" for a head nod. However, emblematic slips can exist by way of the behaviour directly contradicting the simultaneous verbal message being delivered. Emblems are usually exchanged when verbal channels of communications are prevented (e.g., when someone else is speaking). In the present study, we selected emblems that were most frequently observed and most frequently cited in the literature. For example, *Head Nodding, Head Shaking,* and *Shrugging* were all included in the AAI-NVB Manual.

[2] Of note, it could be argued that a number of behaviors in the present taxonomy could be categorized into the haptic or proxemic categories. Self-Soothing, for example, is indeed related to touch behavior. However, we feel that these behaviors are better categorized through Ekman's Kinesic categories in his repertoire of nonverbal behavior (1969a). For example, this type of self-touch may also be referred to as a self-adaptor. When the subject is seated, the use of space (leaning forward/leaning back) may also be referred to as an alter-adaptor. Predominately, the study of haptics and proxemics refer to non-seated movements within the realm of intimate behavior, which is not relevant to this particular study.

Ekman’s second kinesic category is *Illustrators* and these behaviours are also closely tied to speech. However, rather than having a direct translation, these movements serve to illustrate and enhance what is being said verbally. Ekman outlines 6 types of illustrators: 1) *Batons* which accent or emphasize a particular word or phrase, or provide tempo, 2) *Ideographs* which sketch the path or direction of a particular thought, similar to “acting” out a cognitive itinerary, 3) *Dietic Movements* which point to a particular object that is physically present, 4) *Spatial Movements* which depict a spatial relationship, 5) *Kinetographs* which are movements that physically depict a bodily action, and 6) *Pictographs*, whereby the subject traces or draws a picture of a referent using their body. Though in the present study, we adopted *Illustrators* as one of our kinesic codes, we did not require coders to differentiate between type of illustrator observed.

The kinesic category of *Affect Display* is featured most prominently in the nonverbal literature on emotion and became the area of study that Dr. Ekman is best known for (Ekman, 1997; Ekman, 1999; Ekman, 2003; Ekman, 2016; Ekman & Friesen, 1971; Ekman, Davidson, Richard, & Wallace, 2005; Keltner & Ekman, 2000; Keltner & Ekman, 2003). Developed from Tomkins (1962) and Darwinian (Darwin, 1872/1998) evolutionary theory, *Affect Displays* are universal, but the evoking stimuli may be culture specific and regulated by culture-bound display rules (Ekman, 1999). The face is the primary site for affective displays and Darwin maintained that the affect displayed evolved from functional activities associated with facial movements (1872/1998). Ekman notes that the difference between macro and micro expressions is determined by the duration, and further explains that micro expressions are created as a direct result of display rules insofar as, when an emotion is felt, the resulting facial expression is either masked (with another affect display) or neutralized based on what is socially appropriate in the moment (Ekman et al., 1969a). Ekman also stated that “the face is the best liar, with the exception of micro expressions (Ekman et al., 1969b). Within the AAI-NVB Manual, each universal micro-expression is included, along with categories for masking and neutralized expressions. Additionally, we have also included a category for incongruent macro-expressions.

*Regulators* represent Ekman’s fourth kinesic category and describe acts which maintain and regulate the back-and-forth nature of speaking and listening. Similar to illustrators, they are closely related to conversation. However, while *Illustrators* are displayed during moment-to-moment fluctuations in conversation, *regulators* are related to the conversational flow and pacing of the exchange. These behaviours do not tend to convey a substantial amount of communicative information and, as such, they have been excluded from the present study. However, *regulators* are described in the AAI-NVB Manual for the purpose of differentiating them from other behaviours.

Finally, the fifth kinesic category of *Adaptors* was highly implicated in the present study. Ekman describes *Adaptors* to represent adaptive efforts to satisfy bodily needs. These behaviours are categorized into self-, alter-, and object-adaptors, which are socially learned and maintained through habit. These tend to be associated with drive states when first learned, or emotional expectancies and interpersonal interactions later in life. Ekman posits that when an *Adaptor* is present, there exists something in the environment that has triggered a more primal behaviour. These behaviours occur outside of awareness and are displayed as only partial aspects of the full behaviour. For example, when seated with legs crossed, a slight swinging of the leg may be interpreted as an alter adaptor whereas the completed or "full" version of the behaviour may be a more violent or aggressive kick. *Self-Adaptors* are originally learned mastery or management of problems or needs. Some *adaptors* are learned in order to facilitate or block sensory input, such as sound or speech, perform ingestive or excretive functions, perform autoerotic functions, or engage in grooming behaviour. Although these partial behaviours are frequently seen in public during conversation, more fulsome versions of the behaviour may be observed in private or intimate settings, or in the case of personality disorganization. Similarly, alter adaptors are behaviours that are learned early in interpersonal contact. For example, the acts of giving or receiving something from another individual, attaching or protecting the self from attach, establishment if intimacy or affection, flight, or withdrawal. Many of the behaviours found within the AAI-NVB Manual, particularly those involving the hands, limbs, and torso, (e.g., *Leaning Forward, Leaning Back, Leaning Side, Hand Scratch/Pinch, Wiping Mouth, Covering Mouth, Self Soothing, Arms Crossed)* were adopted from Ekman's *Adaptor* model. Though researchers must be extremely cautious when attempting to interpret nonverbal behaviour (Burgoon et al., 2016; Floyd et al, 1998; Harrigan et al., 2013; Shea, 1998) these *Adaptor* behaviours provide the potential for meaningful exploration of subjects' internal states during future studies and were therefore included in this preliminary phase.

As above, caution must also be taken when interpreting nonverbal behaviour within the context of deception. However, there do exist certain nonverbal behaviours that have been well documented within the deception literature and while there is some degree of academic disagreement regarding how exactly to interpret the results of these studies (DePaulo et al., 2003; Ekman, 1971; Vrij, 2008), many of these behaviours were considered within the context of this project at this early stage by virtue of the fact that these behaviours were previously investigated by other researchers. Within this context, pupil dilation, vocal uncertainty, upward head tilt, speech errors, pursed lips, body orientation, blinking pauses, fidgeting, gaze behaviour, illustrators, and hand movements have all been highly implicated as deception cues DePaulo et al., 2003; Ekman et al.,1969b; Vrij, 2008). Apart from pupil dilation, each of these behaviours were included in the AAI-NVB Manual. More specifically, Ekman, discusses the importance of observing hand movements and he advises that hands commit lies of omission rather than lies of commission, indicating that a decrease in hand movement may be relevant to a change in the individual's internal state (Ekman et al., 1969b). He also states that eye contact, leg movements may be implicated as cues for deception. Most importantly, both Ekman and Shea (Ekman, 1969a; Shea, 1998) carefully outline the importance displayed nonverbal behaviours that contradict other bodily movements or tandem verbal information. Therefore, we felt that *Incongruence* was a particularly important factor to include in the present study. Though incongruence itself is not a stand-alone behaviour, it was used as a primary modifier for many of the behaviours in the AAI-NVB Manual, as described in Step 3: *Creating Code Categories*.

Although the present study does not seek to investigate deception explicitly, one of the most valuable resources during this stage of the process was a meta-analysis conducted by DePaulo and colleagues (2003), which outlined over 150 "Deception Cues" which have been previously investigated by researchers. Through careful consideration of each of these factors, we were able to identify a number of important behaviours to be included in the AAI-NVB Manual. Specifically, two studies cited in the DePaulo meta-analysis were of particular interest to our paralinguistic behaviour categories. The first, published in 1995 by Zaparniuk, Yuille, and Taylor, examined the Criteria-Based Content Analysis (CBCA: Steller & Köhnken, 1989), which is a subscale of the Statement Validity Analysis interview (SVA: Raskin & Esplin, 1991)The CBCA assesses the presence or absence of features that often characterize veracity, and of the 19 items on this scale, three were believed to be especially relevant to an attachment context. The first, lack of memory, was one of the original behaviours identified during Phase 1. The second and third, spontaneous corrections and pardoning the perpetrator, were also added to the AAI-NVB Manual. As this interview is typically administered during the process of police interrogation, we shifted the label and definition from “pardoning the perpetrator” to *Pardoning Other* and added an additional behaviour: *Pardoning Self*. The second study of interest was published in 1965 by Kasl and Mahl suggested that utterances such as “ah”, “um”, and similar verbal fillers are distinct from other forms of speech disturbance. Although *Verbal Fillers* were considered as a relevant behaviour in Phase 1 of this study, we failed to consider additional verbal peculiarities such as word or phrase repetition, sentence changes, and word omissions. Kasl and Mahl (1965) list seven forms of speech disturbance in addition to verbal fillers, and we included each of these under the definition of the newly added *Speech Errors* behaviour.

An additional step in selecting relevant behaviours consisted of consultation with nonverbal and attachment experts in the fields of psychology, criminology, anthropology, and business. During this process, we were able to speak with researchers and clinical professionals from North America, Asia, and Europe, some of whom were authors cited in the preceding sections. While the majority of these individuals were able to support the validity of our catalogued behaviours, others suggested the addition of reflexive behaviours such as *Yawn, Sigh, Hard Swallow, Cough, Clear Throat, Sniff,* and *Deep Inhale.* It was also suggested that we return to the original Strange Situation Protocol (SSP: Ainsworth & Bell, 1970) to confirm that our catalogued behaviours in the AAI-NVB Manual resembled those used in this seminal attachment study. Indeed, behaviours such as locomotion, body movement, posture, hand movements, visual regard, and oral behaviours were well represented within the AAI-NVB manual. Others, such as location, contact, and crying, while relevant to an infant attachment context, were not as applicable to the present adult attachment context.

As a final step in the process of cataloging relevant behaviours, and as recommended by Floyd and colleagues (1998) two additional rounds of observation were made using the collected AAI video data. The first round of observation consisted of observing a cross-section of 10 randomly selected videos and noting down any behaviours that were not included in the original Phase 1 behaviours. Finally, to ensure that the behavioural repertoire was adequately sampled, an estimate of the quality of sample coverage was estimated using a theta statistic (Floyd et al., 1998). After creating a behavioural ethogram in the BORIS (Olivier & Marco, 2016) software using a rough draft of the AAI-NVB Manual as a guide, this author randomly selected three AAI videos to code. The resulting data was compiled and analyzed by counting the number of individual behaviours in the catalogue and dividing this number by the total number of acts observed (in this case, using the mean total frequency across the three subjects), and subtracting this number from 1. According to Floyd and colleagues (Floyd et al., 1998) as the value of theta approaches 1, the probability of encountering additional behaviours approaches 0. The resulting theta value was $\theta = 0.94$.

**Step 2: Selecting a Unit of Observation**

The second step outlined by Floyd and colleagues (1998) involves selecting a unit of observation and should be interpreted in two important ways. The first describes an evaluation of the behaviour with regard to either frequency (event) or duration (state). The investigator must ask whether the duration of a particular behaviour is relevant to the construct they are hoping to describe with their coding tool, or if a behavioural frequency is sufficient for their needs. If a behaviour is judged to require a duration, it is known as a state and if frequency is sufficient, this is known as an event.[3] The second interpretation is with regard to the sampling technique used. Specifically, whether the observation and coding will involve observing an entire event and then recording each instance of the behaviour (event sampling), or if the coders will be alternating between observing and coding in shorter time intervals (time sampling).

It is recommended that the decisions surrounding identifying event types and sampling types should be guided by, among other things, study goals and practical considerations (Floyd, 1998). The practicality of duration coding was of minimal consequence, as the BORIS coding software easily accommodates the calculation of duration. For our purposes, it was important that data be collected in a way that would allow for maximal information regarding reliability and agreement estimates, as well as for future studies focusing on validation and exploratory analyses. The units of measurement for the present coding system consisted mainly of state units. However, both state and event units were collected for behaviours such as *Protective Objects*, *Self Soothing*, *Pauses*, *Latency*, *Dissociation*, and all leg behaviour and arm crossing. The frequency of these behaviours was coded in the same manner as the frequency-only behaviours, in addition to a duration modifier added to the coding manual and software.

Floyd and colleagues (1998) caution against the unnecessary coding of state units, as the data can be cumbersome and does not often add value to the coding tool. The decision surrounding which state units to record was based upon how meaningful these data would be.

[3] In the BORIS coding software, events and states are referred to as “point events” and “state events”, respectively.

For example, if a *Yawn* behaviour was identified, data regarding the duration of this yawn was not though to contribute any additional information regarding the subject's internal state. In other words, the frequency of yawning behaviour was believed to be sufficient for our purposes. Overall, the majority of the behaviours with duration modifiers represent potential body-blocking behaviours, as the duration of any body-blocking behaviour represents information over-and-above what might be gleaned from frequency data alone. Additionally, behaviours with frequency-only data were typically too short to warrant a duration modifier to begin with.

With regard to the sampling procedure, and in keeping with our goal of maximizing our collected data, we used a hybrid of event and time sampling. To begin with, rather than coding the AAI in its entirety, we opted to code only questions 2, 3, and 4, plus a baseline period. During this selected interval, the coding process most closely resembled an event-sampling procedure, as the entire segments were coded for every behaviour. However, after data conversion and entry, the database itself most closely resembled a time-sampling procedure, as each behaviour was represented as being either present or absent within each five-second interval of the selected segments (questions two through four, plus baseline). The decision to represent the data using five-second intervals was made primarily to assist in the evaluation of agreement and reliability, and for training purposes. For example, it was important to determine that if Coder 1 noted five instances of *Yawning* and Coder 2 also noted five instances of *Yawning* that these observations were the same five yawns, rather than there being 10 yawns in total, coded at different five-second intervals. Finally, in another attachment-based coding study, Mann and colleagues (1991) reported that their target attachment behaviours tended to occur in bouts and were not accurately captured using time-sampling protocols.

**Step 3: Creating Code Categories**

Floyd and colleagues (1998) advise that, when developing behavioural categories, these categories must be both mutually exclusive and exhaustive. As such, when organizing the selected relevant behaviours into their respective discrete categories, there should be no overlap between these categories, and each behaviour should be accounted for within these categories. If possible, it may also be helpful for some categories to be arranged in a hierarchical manner to assist coders in their decision-making process.

During Phase 1, the original behaviours were classified according to Shea's conceptualization of nonverbal behaviour and consisted of the organization of the extant behaviours into the three categories of *Kinesics, Proxemics,* and *Paralanguage* (Shea, 1998). However, the addition of a multitude of behaviours across Phase 2 necessitated a reimagining of categories with the aim of facilitating both the coding and analysis processes (Floyd et al., 1998)[4]. For the purposes of coding, the AAI-NVB Manual includes eight nine distinct behavioural categories: *Hands, Body, Head, Face, Eye Movement, Reflexive Behaviours, Paralanguage,* and *Affective Displays*. These categories were created in order to facilitate the coding process by directing the coders' attention to a specific area of the body or specific cluster of behaviour. Additionally, each behaviour was organized in a hierarchical fashion within these categories to further facilitate coding decisions. As the BORIS software (Olivier & Marco, 2016) was created by ethologists, the software is specifically designed to facilitate this hierarchical modelling organization.

For example, within the *Hands* category, *Illustrators, Hand to Ear, Hand to Nose, Hand Clench, Hand Stop, Hand Pick, Hands Together, Protective Object, Hand Hiding, Hand Dismiss, Shielding Eyes,* and *Obscene Gesture* are stand-alone point events with no modifiers. However, further differential decisions are required for the behavioural headings of Hands Apart (*Symmetrical* or *Asymmetrical),* Hand to Mouth (*Covering Mouth* or *Wiping Mouth),* Hand to Face (*Forehead* or *Chin),* Hand to Knee (*Above* or *Below),* Hand Shrug (*Incongruent* or *Not Incongruent*), and Self Soothe (*Repetitive Movement, Neck Covered,* or *Body Holding*). Additionally, *Self Soothe* is noted as a state event in the software, and therefore requires coding for both start-and endpoints. Some classes of behaviours require more than one differential decision to be made. Shrugging, for example, first requires the differentiation between *Full Shrug, Shrug Without Arms,* and *Asymmetric Shrug* and then requires a decision to be made surrounding whether the behaviours is *incongruent* or *not incongruent*.

**Manual refinement.** The development of the AAI-NVB Manual was a continual process which considered multiple rounds of coding and adjustment of categories and factors, particularly after Phase 2B. Most notably, the incongruence modifiers for the *Hand Stop, Hand Dismiss, Head Tilt Down and Away,* and *Obscene Gesture* behaviours were removed. We found that incongruence is most useful when the parent behaviour contained a clear meaning as with

[4] For theoretical purposes, one may consider Shea's nonverbal conceptualization (Shea, 1998) using Ekman's Repertoire of Nonverbal Behavior (Ekman & Friesen, 1969a) to categorize the Kinesic movements. However, this was not thought to be useful for the purposes of coding and analysis.

cases like *Head Nod, Head Shake,* and shrugging behaviours. Otherwise, the incongruence was too difficult to establish and could not be reliably coded. Indeed, Floyd and colleagues recommend that codes be relatively elemental rather than inferential (Floyd et al., 1998). For example, in order to code an incongruent modifier for the *Head Tilt Down and Away* behaviour, one would first have to establish the meaning of that parent behaviour in order to determine if the behaviour was incongruent or not. As we are presently not concerned with, nor are we able to definitively establish the meaning of nonverbal behaviour, we only retained incongruent modifiers for behaviours which, in and of themselves, contain an independent meaning. The exception to this rule was the incongruent modifier for *Obscene Gesture*. We concluded that, within the context of a clinical interview, the presence of an *Obscene Gesture* would be incongruent on its face. Additionally, if used within the context of story telling "…and I told him to hit the road!" then the obscene gesture would instead be coded as an *Illustrator*. The decision to remove both *Scoff* and *Sarcasm* was made on the basis that 1) no true instances of these behaviours were found during Phase 2A or Phase 2B, 2) the presence of the *Sarcasm* and *Scoff* behaviours were common confounds for *Laughter, Humor*, and *Laughter, Incongruent*, and 3) there was little empirical support for these behaviours in the nonverbal literature.

Three new behaviors were added in Phase 2C (*Fidget*, *Gaze Aversion,* and *Arms Crossed)* in order to represent the merging of *Body Fidget* and *Object Fidget*, *Eyes Looking Away* and *Eyes Searching,* and *One Arm Crossed* and *Two Arms* Crossed, respectively. Once again, a review of the literature could not effectively determine that the original two behaviours were functionally different and the cost of decreased reliability due to coding error outweighed any potential usefulness of keeping these behaviours separate. Additional behaviours were also added to the AAI-NVB Manual during Phase 2C for added specificity. The affective categories of Micro Expressions and Incongruent Expressions became more detailed by adding variables specific to the affective displays observed, and a number of leg behaviours were also included to increase specificity within the *Limbs* category. Finally, updates were made to AAI-NVB Manual descriptions and examples for *Hands Apart Symmetrical* and *Asymmetrical, Hand Covering Mouth* and *Hand Wiping Mouth, Hand Touching Forehead* or *Chin, Hand Touching Ear, Hand Clench*, *Hand Stop, Hand Dismiss, Hand Touching Knees (Above and Below) Fidget, Head Tilt, Verbal Diversions, Lack of Memory*, and *Dismissiveness.*

**Step 4: Developing a Coding Manual**[5]

Once the relevant behaviours, units of interest, and categories have been defined, the next step recommended by Floyd and colleagues (1998) is to develop a coding manual. This manual should provide a list of all codes, a descriptive definition for each code, examples of behaviours that represent the different codes, and examples of differential decisions. The APA guidelines (APA, 1985) also suggest that providing information regarding the theoretical underpinnings of the coding tool may also be useful.

In the present study, the development of the coding manual was a cyclical process. As each video was coded and reviewed, codes and categories were revised on the basis of the agreement and reliability results, as well as observations from the coders and primary investigators. Each updated iteration of the manual was subsequently disseminated to the coding team prior to being assigned a new video to code. The coding manual begins a theoretical background, which describes the foundations of attachment theory, clinical applications of the AAI, and a rationale for the current study, and continues by describing key points in nonverbal behaviour theory.  The third section provides an in-depth description of the coding protocol, and includes sections on confidentiality, suggestions regarding coding strategies, and a step-by-step overview of the coding software[6]. The longest and final section includes, as Floyd and colleagues (1998) recommended, a list of all behaviours and their respective modifiers, and a break-down of each behaviour, a description of that behaviour, examples that match the behaviour, and differential examples that would not be considered as "correct" for that particular behaviour.

Additionally, each behavioural section included a statement regarding to which nonverbal category the behaviour belonged, as well as to which of Ekman's kinesic categories the behaviour belonged. Finally, each behavioural category provided an indication of which modifiers, if any, were relevant to that particular behaviour.  Each behaviour was organized into one of 9 larger behavioural sections: *Head, Body, Hands, Face, Affect Displays, Eye Movement, Reflexive*, *Paralanguage I*, and *Paralanguage II*.

---

[5] Floyd and colleagues (2000) refer to a "codebook" rather than a coding manual.
[6] Coders were additionally provided with the BORIS software manual, however the AAI-NVB manual contains only descriptions the key functions necessary to complete the coding for this specific project.

### Step 5: Inter-Rater Agreement

Agreement in Phase 1 was informally assessed via group consensus and Phase 2 agreement and reliability were investigated through a series of statistical analyses. Additionally, in Phase 2, reliability and agreement were investigated on three separate occasions and the results of these investigations served to inform subsequent changes to the codes and manual. Floyd and colleagues (1998) provide guidelines and recommendations regarding the training of coders for the purposes of establishing adequate reliability of the newly developed coding tool, many of which are addressed in the subsequent Coding Process and Analysis sections.

## Coding Process

An integral part of the five-step process noted above included coding video recorded AAIs based on the behaviours compiled within the coding manual. The coding with these videos not only served to investigate coder reliability, but also assisted in the continual process of cataloguing relevant behaviours, validating units of measurement, further developing the manual, and providing feedback to the coding team. The coding process confirmed the existence of the anecdotal behaviours noted by clinicians, as well as validated behaviours noted in the literature, and guided the cataloguing of new behaviours that were observed by coders during the coding process. As alluded to above, the coding process in Phase 1 of this study differed substantially from the coding process in Phase 2 and are described later in this section. The following considerations were made as per Harrigan's suggestions outlined in *The New Handbook of Methods in Nonverbal Behaviour Research* (Harrigan, Rosenthal, Scherer, & Scherer, 2008).

### Methodological Considerations

As per Harrigan and colleagues (2008) there exist several conceptual considerations which affect the coding process and strategy. With regard to feasibility, although there exists a vast array of possible body movements, especially giving consideration to speed, frequency, interactive qualities, and individual variably, there exists a finite number of body parts that can move independently of one another. As noted by Harrigan and colleagues (2008), this fact helps to reduce the intricacies of coding. For example, the upper arm cannot move independently of the lower arm. We leveraged this fact in creating the behavioural categories seen in Phase 2, which additionally assisted in the organization of the coding manual and behavioural coding ethogram, as described below.

Further, of all possible behaviours, individuals are typically limited by cultural conventions and would not, for example, gesticulate widely while another person is talking. This consideration served two functions for the purposes of this study: 1) limiting the frequency at which behaviours were coded and, 2) allowing for easier identification of anomalous behaviours of interest, as they are not typically performed and tend to “stand out” when they are. The final consideration regarding feasibility is the notion that many movements, especially those consisting of postural shifts, occur in tandem. Thus, they do not require separate overlapping codes.

**Phase 1: Pilot**

In Phase 1, anecdotal comments from attachment-informed clinicians were considered when coding the AAI videos. These clinicians habitually utilized the AAI within their clinical practices and noted some common nonverbal occurrences. As stated above, these behaviours included *Latency, Pauses, Lack of Memory, Generalizations, Body Shifts, Verbal Diversions, Laughter, Verbal Fillers, Humor, Protective Objects, Dismissiveness, Incongruence, Dissociation, Self Soothing, Gaze Aversion,* and *Micro Expressions.* In-Person training during this phase included observation of video-recorded examples for each behaviour with one of the principle investigators and in-depth discussions and examples regarding the various presentations of these identified behaviours. This training phase also included an introduction to the AAI and its applications for clinical practice.

The identified Phase 1 behaviours were coded manually using a MS Word template developed by the one of the primary investigators and two graduate students. For each behaviour, space was provided to indicate at which point the behaviour was observed (time), and spaces for start and end points for behaviours that were temporally based (*Self Soothing, Dissociation, Pauses*, *Latency,* and *Protective Objects*). A comments section was also included, and coders were encouraged to take note of any commonly observed behaviours that were not included in the main coding system and any contextual factors observed during the coded behaviours. The resulting data were not analyzed. Rather, the coders participated in a group feedback session during which each coded instance of each behaviour was reviewed, and any disagreements were resolved by all coders reviewing the subject’s video together. This coding/feedback process was repeated using three subjects during this phase.

**Phase 2**

As previously discussed, the coding manual for Phase 2 of this study was the product of careful review of the relevant nonverbal literature, consultation with attachment and nonverbal experts, and feedback and observations from the pilot phase (Phase 1). The most obvious drawback to the Phase 1 coding process was found to be the coding technique itself. Given the drastic increase of target behaviours in the updated Phase 2 manual, it was unreasonable and inefficient to have observations be made manually (i.e., manually stopping and starting the videos; typing in the timepoints; calculating duration). Through additional research, consultation, and testing various software interfaces, the Behavioural Observation Research Interactive Software (BORIS: Olivier & Marco, 2016) was decidedly ideal for our purposes. This software allows for the pre-programming of a behavioural ethogram – an inventory of behaviours and modifiers of interest – which structured to map onto the behavioural categories found within the coding manual. A copy of this ethogram can be found in Appendix A. The BORIS software also allows videos to be played within the interface of the software itself and provides coders with the ability to record observations by pressing the pre-programmed keystrokes. When pressing a series of keys that corresponded to the observed behaviour (e.g., "D" for the *Head* category, "N" for *Nodding* behaviour, and then "I" to indicate the presence of an incongruent head nod), the software automatically paused the video in order to select various behaviours and modifiers from the menu and recoded the timepoint at which the coded behaviour took place. Once coding was complete, the software provides several export options and compiles all the observations together into one file for review.

Subsequent to the programming and software testing, a subset of three videos were randomly selected from the AAI video library for Phase 2 coding. In this Phase, training was delivered in two modules and guided by Floyd, Baucom, Godfrey, and Palmer's (1998) training recommendations. The first module was similar to the training in Phase 1, wherein all coders met with the lead graduate student and were presented with a copy of the coding manual. Examples of each behaviour were presented, along with a rationale for each coding decision, and discussions were had surrounding potential behavioural confounds.  For example, coders were instructed on methods to differentiate between *Head Tilt Down* and *Head Nodding*. Further discussion and questions were encouraged during this Phase of training. Coder then returned for another training session in which they were tested on the manual through both written and experiential exercises. The second module of this training phase was to deliver instruction and experiential training on the newly adopted coding software. As a group, coders and trainers downloaded and installed the software and ethogram on their respective computers, saved within a password-protected and encrypted partitioned drive, and each of the necessary functions of the software were tested by each coder using a test video.

Three rounds of coding training were completed in Phase 2, which represent Phase 2A, Phase 2B and Phase 2C of this study. During each round, coders were each assigned the same video, ethogram, and assignment document, which outlined the target time intervals in the video, due dates, and any special instructions. For example, during one round of training, the subject needed to take an urgent phone call during one of the target intervals. Coders were given the start and end timepoints of this phone call and were instructed not to code during this period. Once each coder had completed the assigned video, observations were exported and send to the lead graduate student for review.

Due to the importance placed on establishing reliability of this coding tool, the analysis and feedback process was an integral step toward this goal. Floyd and colleagues (1998) recommend that investigators should evaluate the most precise unit of observation possible in order to provide additional training, monitoring, and corrective feedback to the coding team. As such, a series of steps were conducted, beginning with the reorganization of the coding forms received from the coding team. The resulting data was then entered into the main database, which was structured to represent every possible five-second time interval of interest for the video under review. Total frequencies for each behaviour, and for each coder, were also calculated for subsequent analysis. Once all the data from a particular phase was entered and cleaned, the data was exported into .csv format and imported into the statistical software. At this point, syntax was written to parse out 8 columns of data (one per coder) for each target behaviour. These matrices were subsequently exported into MS excel format (one matrix per behaviour), and each of the five-second intervals along the y axis were evaluated for the presence or absence of a coded observation. Rows with zero observations were deleted, and the remaining rows of data were each validated by the lead graduate student by looking at each coded instance within the spreadsheet and comparing it to that particular timepoint within the video itself. Eight feedback documents (one for each coder) were created, which provided the “positive” instances of each behaviour along with any missed targets and false positives coded.

In order to deliver corrective feedback, each coder was first provided with their respective feedback documents and instructed to review each instance of missed and false positive behaviour, along with a brief comment regarding what they observed about *their own* coding for each behaviour. All members of the team then came together in a group feedback session, where all instances of missed behaviours and false positives were reviewed cumulatively, allowing each member of the team could learn from each other. Global feedback was also provided, which included general observations made by the graduate student and coding team. At the end of each feedback session, a new video, ethogram, and updated manual were provided to each coder. This coding and feedback processes were repeated three times during the course of this current study.

## Analysis & Results

Germane to the second goal of this project, we conducted statistical analyses in order to ascertain whether the nonverbal behaviours identified in the AAI-NVB Manual could be reliably coded. The following analyses were conducted on the data collected during Phase 2 (A, B,& C).

As Cohen's Kappa is limited to two observers, common alternatives are to employ Fleiss' Kappa or intraclass correlation (ICC) for cases of continuous data as an estimate of inter-rater reliability for coders making independent ratings (Hallgren, 2012). Though Fleiss' Kappa and ICC were selected to compute the primary reliability estimates for this study, it has been noted by other researchers that reliability and agreement, although frequently used interchangeably, do not effectively measure the same construct (Hallgren, 2012; Viera & Garrett, 2005). While agreement refers to stability of scores within a single observation, reliability refers to consistency between ratings (Liao, Hunt, & Chen, 2010). Consequently, because reliability is calculated using between-observation variance and within-observation variance, there can exist cases of simultaneous high levels of agreement and low reliability estimates when between-observation variance is low (i.e., when the target behaviours are rare or low frequency, such as coughing). The resulting reliability coefficient can be profoundly affected by the prevalence of the target, indicating that reliability and agreement can, but do not always, occur together (Liao, Hunt, & Chen, 2010; Viera & Garrett, 2005).

Indeed, this "Agreement-Reliability Paradox" has been well documented (Cicchetti & Feinstein, 1990; Falotico & Quatto, 2015; Feinstein & Cicchetti, 1990; Gwet, 2008; Karstad et al., 2018) and, as such, a decision was made to include Percentage Agreement values along with the reliability estimates to better represent the data. Moreover, due to the number of target variables, complexity of the target behaviours, and the prevalence of common behavioural confounds (i.e., behaviours that look similar but are functionally different), it was also important to acknowledge cases of perfect agreement resulting from non-occurrence, which cannot be represented by a reliability estimate (Bryington, Palmer, & Watkins, 2002). The strategy of employing multiple measurement and estimate strategies is common within observational studies (Brouwer, Reneman, Dijkstra, Groothoff, Schellekens, & Göeken, 2003; Karstad, et al., 2018. Menz, Fotoohabadi, Wee, & Spink, 2012; Reneman, Brouwer, Meinema, Dijkstra, Geertzen, & Groothoff, 2004; Zander et al., 2016) and allows for a deeper exploration of the presented data.

Specifically, to obtain estimates of observer reliability on specific behaviours we employed a Fleiss' Kappa reliability estimate (Hallgren, 2012). This analysis was repeated on each individual behaviour across all three subjects in Phase 2. As mentioned in the previous Coding Process section, the data for each behaviour was entered into a database in a five-second interval format. For example, the behaviour *Illustrators* for Subject 1 included 210 distinct "observations" (210 five-second intervals), each indicating an occurrence (frequency) or non-occurrence (zero). Therefore, the reliability estimates for each behaviour were computed based on these 210 observations across 8 raters for Subject 1. Percentage Agreement values were computed on the same set of data for comparison.

Due to the importance placed on the feedback provided to coders throughout Phase 2, individual coder reliability estimates were computed by creating pairwise comparisons between the reliability estimates of each of the eight coders and obtaining a mean value for each coder. This method was repeated for each behaviour coded and allowed for additional insight into the individual strengths and weaknesses of each coder. The results from these analyses can be found in Appendix E.

Finally, due to the limitations of reliability for low-frequency observations and agreement, we sought to compute an additional reliability estimate for each behavioural group by collapsing each behaviour across observational points to obtain a total frequency for each coder, and comparing these total values with other total behavioural frequencies of the same functional category. This technique has been recommended by Harrigan and colleagues (2008) as well as Karstad (2018). Further, on the recommendation of Floyd and colleagues (1998), the data reduction process also considered grouping low-frequency behaviours with high-frequency behaviours. For these grouped behaviour analyses we employed an ICC two-way random effects, absolute agreement, multiple raters model for our fully crossed design (ICC 2, 8) (Koo & Li, 2016). These categories for the grouped analyses roughly corresponded to the behavioural categories found within the coding manual with some minor exceptions. Most notably, the paralanguage and affect categories were both divided in half, based on how similar the behaviours were to each other. For example, the affect category was divided into *Micro Expressions* and *Incongruent Expressions*. In addition, the incongruent and duration modifiers were combined to create their own categories, respectively. The rationale behind this decision was to examine the ease with which these modifiers could be reliability identified and coded.

While the recommended number of items within a category for ICC computation is 30 (Feng, 2015) and our categories ranged from 6-26 behaviours, this method was determined to be the most accurate estimate of reliability for the present data (Floyd et al., 1998). Values less than 0.00 are indicative of poor reliability, between 0.00 and 0.20 of slight reliability, between 0.21 and 0.40 of fair reliability, between 0.41 and 0.60 of moderate reliability, between 0.61 and 0.80 of substantial reliability, and between 0.81 and 1.00 of almost perfect reliability (Landis & Koch, 1977). All observations were generated by the BORIS software (Olivier & Marco, 2016), prepared using MS Excel and analyzed using the R software (R Core Team, 2017) using the computational package *irr* (Gamer, Lemon, & Fellows Puspendra Singh, 2012). The following results are grouped based on behavioural category and while the reliability estimates for the individual behaviors are integral to the manual refinement and feedback process, the most effective measure of instrument reliability are the reliability estimates of the aggregate categories.

**Hands**

The individual behaviours that comprise the *Hands* category are: *Illustrators, Hands Symmetrical, Hands Asymmetrical, Covering Mouth, Wiping Mouth, Hand to Eye, Hand to Forehead, Hand Touching Chin, Chin Resting in Hand, Hand to Ear, Hand to Nose, Hand Above Knee, Hand Below Knee, Hand Clench, Hand Shrug, Hand Stop, Hand Scratch/Pinch/Pick, Hands Together, Self Soothe, Self Soothe – Neck Covered, Self Soothe – Body Holding, Hands Hiding, Hand Dismiss, Shielding Eyes, Obscene Gesture, and Fidget Object.* Within the *Hands* category, the most frequently coded behaviour across all three subjects was *Illustrators*, while the lowest frequency behaviours were *Shielding Eyes* and *Obscene Gesture*. The frequency at which each *Hand* behaviour was coded during each Phase, along with mean frequency, SD, median, range, skewness, and kurtosis for each can be found within Appendix D.

The distribution of Fleiss' Kappa coefficient estimates for the *Hands* category, collapsed across subjects and behaviours, was found to center around 0.20, with a grand mean of 0.27, a median of 0.18, and a standard deviation of 0.21. Reliability estimates ranged from poor (-0.001) for the *Self Soothe: Neck Covered* behaviour to substantial (0.66) for the *Covering Mouth* behaviour, with a skewness of 0.50 and a kurtosis of -1.21. The distribution of percentage agreement values with a tolerance of zero for the *Hands* category, collapsed across subjects and behaviours, was found to center around 94, with a grand mean of 92.26, a median of 96.48, and a standard deviation of 11.35. Values ranged from 44.43 for *Illustrators* to 100.00 for the *Obscene Gestures* behaviour, with a skewness of -3.26 and a kurtosis of 12.89. At a tolerance of 1, the percentage agreement values for the *Hands* category, collapsed across subjects and behaviours, was found to center around 99, with a grand mean of 98.92, a median of 99.73, and a standard deviation of 4.93. Values ranged from 74.75 for *Illustrators* to 100.00 for the *Obscene Gesture, Shielding Eyes, Neck Covered, Hand to Ear, Hand to Forehead*, *Hand to Eye*, and *Hands Symmetrical* behaviours, with a skewness of -5.01 and a kurtosis of 25.36. See Table 1 for mean Fleiss' Kappa coefficient estimates and mean percentage agreement values for each individual behaviour in this category, collapsed across each Phase, and Table 2 for values specific to each individual behaviour within each Phase.

Table 1

*Estimated Reliability and Agreement Means for Each Behaviour in the Hands Category, Collapsed Across Three Subjects*

| | | | *PA (Tol =0)* | | *PA (Tol =1)* | |
|---|---|---|---|---|---|---|
| Behaviour | *k M* | *k SD* | *M* | *SD* | *M* | *SD* |
| Illustrators | 0.411 | 0.11 | 44.43 | 11.83 | 74.57 | 11.14 |
| Hands Symmetrical | 0.347 | 0.34 | 97.87 | 2.26 | 100.00 | 0.00 |
| Hands Asymmetrical | 0.166 | 0.09 | 87.43 | 12.12 | 98.50 | 1.82 |
| Covering Mouth | 0.657 | 0.13 | 90.90 | 11.48 | 98.70 | 1.47 |
| Wiping Mouth | 0.537 | 0.28 | 96.60 | 2.80 | 99.80 | 0.35 |
| Hand to Eye | 0.551 | 0.18 | 97.97 | 0.91 | 100.00 | 0.00 |
| Hand to Forehead | 0.314 | 0.27 | 98.33 | 0.23 | 100.00 | 0.00 |
| Hand Touching Chin | 0.1 | 0.10 | 98.47 | 0.55 | 99.63 | 0.32 |
| Chin Resting in Hand | 0.139 | 0.11 | 99.23 | 0.71 | 99.83 | 0.29 |

| Behaviour | *k M* | *k SD* | *PA (Tol =0)* | | *PA (Tol =1)* | |
|---|---|---|---|---|---|---|
| | | | *M* | *SD* | *M* | *SD* |
| Hand to Ear | 0.237 | 0.41 | 98.93 | 0.70 | 100.00 | 0.00 |
| Hand to Nose | 0.615 | 0.15 | 95.13 | 5.66 | 99.83 | 0.29 |
| Hand Above Knee | 0.179 | 0.09 | 96.83 | 3.16 | 99.70 | 0.27 |
| Hand Below Knee | 0.537 | 0.22 | 98.40 | 1.48 | 99.83 | 0.29 |
| Hand Clench | 0.5 | 0.25 | 97.10 | 1.10 | 99.67 | 0.58 |
| Hand Shrug | 0.238 | 0.16 | 87.33 | 2.27 | 98.87 | 1.10 |
| Hand Stop | 0.167 | 0.24 | 96.37 | 2.11 | 99.53 | 0.45 |
| Hand Scratch, Pick | 0.142 | 0.08 | 86.63 | 8.72 | 99.67 | 0.58 |
| Hands Together | 0.534 | 0.08 | 78.00 | 20.30 | 98.00 | 1.76 |
| Self Soothe | 0.106 | 0.11 | 84.10 | 20.77 | 99.37 | 1.10 |
| Neck Covered | -0.001 | | 99.67 | 0.58 | 100.00 | 0.00 |
| Body Holding | 0.069 | 0.10 | 95.10 | 8.14 | 99.67 | 0.58 |
| Hands Hiding | 0.073 | 0.07 | 87.47 | 12.77 | 99.40 | 0.66 |
| Hand Dismiss | 0.05 | 0.05 | 96.27 | 1.22 | 99.83 | 0.29 |
| Shielding Eyes | -0.001 | | 99.70 | 0.52 | 100.00 | 0.00 |
| Obscene Gesture | | | 100.00 | 0.00 | 100.00 | 0.00 |
| Fidget Object | 0.149 | 0.15 | 90.45 | 10.11 | 99.75 | 0.35 |

The ICC inter-rater reliability coefficients for the overall *Hands* category were estimated to range from ICC(2,8) = .705 to ICC(2,8) = .881 across all three subjects, with an estimated grand mean coefficient of ICC(2,8) = .795. The inter-rater reliability coefficient estimate for the *Hands* category was found to be substantial during Phase 2A, (ICC(2,8) = .800), with a 95% confidence interval ranging from 0.70 to 0.89, ($F(26,239) = 42.2$, $p < .001$). Similarly, during Phase 2B, the inter-rater reliability coefficient estimate for the *Hands* category was also found to be substantial, (ICC(2,8) = .705), with a 95% confidence interval ranging from 0.58 to 0.82, ($F(26,243) = 24.8$, $p < .001$). Estimates improved even further for the *Hands* category during Phase 2C and were found to be almost perfect, (ICC(2,8) = .881), with a 95% confidence interval ranging from 0.81 to 0.4, ($F(26,173) = 59.2$, $p < .001$). A summary of these overall intraclass correlation estimates can be found in Appendix C and a complete reporting of the individual behaviour estimates and values within this category can be found in Table 2.

Table 2

*Fleiss' Kappa Estimates and Percentage Agreement Values for the Behaviours in the Hands Category During Each Phase*

| Behaviour | *k* | *z* | *p* | Percentage Agreement *Tol* = 0 | Percentage Agreement *Tol* = 1 |
|---|---|---|---|---|---|
| Phase 2A: Subject 1 | | | | | |
| Illustrators | 0.501 | 56.00 | < .001 | 56.7 | 84.80 |
| Hands Symmetrical | 0.109 | 10.60 | < .001 | 98.1 | 100.00 |
| Hands Asymmetrical | 0.118 | 12.70 | < .001 | 93.3 | 99.50 |
| Covering Mouth | 0.794 | 80.20 | < .001 | 98.6 | 99.50 |
| Wiping Mouth | 0.678 | 65.90 | < .001 | 98.6 | 100.00 |
| Hand to Eye | 0.664 | 64.60 | < .001 | 99 | 100.00 |
| Hand to Forehead | 0.146 | 14.20 | < .001 | 98.6 | 100.00 |
| Hand Touching Chin | 0.108 | 12.40 | < .001 | 98.1 | 99.50 |
| Chin Resting in Hand | 0.065 | 7.23 | < .001 | 98.6 | 99.50 |
| Hand to Ear | -0.001 | -0.09 | 0.926 | 99 | 100.00 |
| Hand to Nose | 0.708 | 83.40 | < .001 | 98.6 | 99.50 |
| Hand Above Knee | 0.236 | 26.20 | < .001 | 93.3 | 99.50 |
| Hand Below Knee | 0.378 | 42.40 | < .001 | 96.7 | 99.50 |
| Hand Clench | 0.226 | 27.90 | < .001 | 97.1 | 99.00 |
| Hand Shrug | 0.172 | 19.40 | < .001 | 84.8 | 97.60 |
| Hand Stop | 0.058 | 6.67 | < .001 | 95.2 | 99.50 |
| Hand Scratch, Pick | 0.203 | 21.10 | < .001 | 77.1 | 99.00 |
| Hands Together | 0.474 | 48.60 | < .001 | 60 | 96.70 |
| Self Soothe | 0.22 | 22.6 | < .001 | 60.5 | 98.1 |
| Neck Covered | -0.001 | -0.093 | 0.926 | 99 | 100 |
| Body Holding | 0.139 | 15.1 | < .001 | 85.7 | 99 |
| Hands Hiding | 0.125 | 14.1 | < .001 | 89 | 99.5 |
| Hand Dismiss | 0.106 | 12.3 | < .001 | 95.2 | 99.5 |
| Shielding Eyes | - | - | - | 100 | 100 |
| Obscene Gesture | - | - | - | 100 | 100 |
| Fidget Object | 0.256 | 26.9 | < .001 | 83.3 | 99.5 |
| Phase 2B: Subject 2 | | | | | |
| Illustrators | 0.291 | 31.1 | < .001 | 33.1 | 62.7 |
| Hands Symmetrical | - | - | - | 100 | 100 |
| Hands Asymmetrical | 0.110 | 10.10 | < .001 | 73.5 | 96.4 |
| Covering Mouth | 0.627 | 58.7 | < .001 | 77.7 | 97 |
| Wiping Mouth | 0.218 | 19.5 | < .001 | 93.4 | 99.4 |
| Hand to Eye | 0.347 | 30 | < .001 | 97.6 | 100 |
| Hand to Forehead | 0.621 | 53.7 | < .001 | 98.2 | 100 |
| Hand Touching Chin | 0.194 | 18.1 | < .001 | 98.2 | 99.4 |
| Chin Resting in Hand | - | - | - | 100 | 100 |
| Hand to Ear | -0.002 | -0.156 | 0.876 | 98.2 | 100 |
| Hand to Nose | 0.693 | 59.9 | < .001 | 88.6 | 100 |
| Hand Above Knee | 0.221 | 19.1 | < .001 | 99.4 | 100 |
| Hand Below Knee | 0.443 | 38.3 | < .001 | 99.4 | 100 |
| Hand Clench | 0.725 | 62.7 | < .001 | 98.2 | 100 |

| Behaviour | *k* | *z* | *p* | Percentage Agreement | |
|---|---|---|---|---|---|
| | | | | *Tol* = 0 | *Tol* = 1 |
| Hand Shrug | 0.119 | 10.5 | < .001 | 88 | 99.4 |
| Hand Stop | -0.001 | -0.104 | 0.917 | 98.8 | 100 |
| Hand Scratch, Pick | 0.048 | 4.14 | < .001 | 88.6 | 100 |
| Hands Together | - | - | - | 100 | 100 |
| Self Soothe | 0.1 | 8.67 | < .001 | 92.2 | 100 |
| Neck Covered | - | - | - | 100 | 100 |
| Body Holding | - | - | - | 100 | 100 |
| Hands Hiding | -0.001 | -0.052 | 0.958 | 99.4 | 100 |
| Hand Dismiss | -0.002 | -0.209 | 0.835 | 97.6 | 100 |
| Shielding Eyes | - | - | - | 100 | 100 |
| Obscene Gesture | - | - | - | 100 | 100 |
| Fidget Object | 0.0416 | 3.59 | 0.000328 | 97.6 | 100 |
| Phase 2C: Subject 3 | | | | | |
| Illustrators | 0.442 | 43.30 | < .001 | 43.5 | 76.2 |
| Hands Symmetrical | 0.585 | 46.20 | < .001 | 95.50 | 100 |
| Hands Asymmetrical | 0.271 | 22.10 | < .001 | 95.5 | 99.6 |
| Covering Mouth | 0.549 | 47.30 | < .001 | 96.40 | 99.6 |
| Wiping Mouth | 0.715 | 56.50 | < .001 | 97.8 | 100 |
| Hand to Eye | 0.642 | 50.7 | < .001 | 97.3 | 100 |
| Hand to Forehead | 0.175 | 13.8 | < .001 | 98.2 | 100 |
| Hand Touching Chin | -0.001 | -0.089 | 0.929 | 99.1 | 100 |
| Chin Resting in Hand | 0.213 | 16.8 | < .001 | 99.1 | 100 |
| Hand to Ear | 0.713 | 56.4 | < .001 | 99.6 | 100 |
| Hand to Nose | 0.445 | 35.1 | < .001 | 98.2 | 100 |
| Hand Above Knee | 0.079 | 6.86 | < .001 | 97.8 | 99.6 |
| Hand Below Knee | 0.79 | 62.4 | < .001 | 99.1 | 100 |
| Hand Clench | 0.55 | 46.8 | < .001 | 96 | 100 |
| Hand Shrug | 0.424 | 34.3 | < .001 | 89.2 | 99.6 |
| Hand Stop | 0.445 | 39 | < .001 | 95.1 | 99.1 |
| Hand Scratch, Pick | 0.176 | 13.9 | < .001 | 94.2 | 100 |
| Hands Together | 0.593 | 49.4 | < .001 | 74 | 97.3 |
| Self Soothe | -0.001 | -0.044 | 0.965 | 99.6 | 100 |
| Neck Covered | - | - | - | 100 | 100 |
| Body Holding | -0.001 | -0.044 | 0.965 | 99.6 | 100 |
| Hands Hiding | 0.093 | 7.68 | < .001 | 74 | 98.7 |
| Hand Dismiss | 0.046 | 3.64 | < .001 | 96 | 100 |
| Shielding Eyes | -0.001 | -0.089 | 0.929 | 99.1 | 100 |
| Obscene Gesture | - | - | - | 100 | 100 |
| Fidget Object | | | | | |

*Note:* There were 210 five-second intervals during Phase 2A, 166 during Phase 2B, and 223 during Phase 2C.

**Body**

The individual behaviours that comprise the *Body* category are: *Lean Forward, Lean Back, Lean Side, Rotate Away, Protective Object, Body Fidget,* and *Fidgeting.* Within the *Body* category, the most frequently coded behaviours across all three subjects were Fidgeting behaviours (*Body Fidget and Fidgeting*) and the behaviours coded with the least frequency were *Leaning Side* and *Rotated Away.* The frequency at which each *Body* behaviour was coded during each Phase, along with mean frequency, SD, median, range, skewness, and kurtosis for each can be found within Appendix D.

The distribution of Fleiss' Kappa coefficient estimates for the *Body* category, collapsed across subjects and behaviours, was found to center around 0.14, with a grand mean of 0.16, a median of 0.12, and a standard deviation of 0.15. Reliability estimates ranged from slight (0.05) for the *Rotate Away* behaviour to moderate (0.49) for the *Fidgeting* behaviour, with a skewness of 2.11 and a kurtosis of 4.84. The distribution of percentage agreement values with a tolerance of zero for the *Body* category, collapsed across subjects and behaviours, was found to center around 90, with a grand mean of 87.28, a median of 94.50, and a standard deviation of 16.04. Values ranged from 54.30 for the *Fidgeting* behaviour to 98.47 for the *Rotated Away* behaviour, with a skewness of -1.88 and a kurtosis of 3.23. At a tolerance of 1, the percentage agreement values for the *Body* category, collapsed across subjects and behaviours, was found to center around 99, with a grand mean of 98.47, a median of 99.83, and a standard deviation of 2.54. Values ranged from 93.30 for the *Fidgeting* behaviour to 99.83 for the *Lean forward, Lean Back, Lean Side, Rotate Away,* and *Protective Objects* behaviours, with a skewness of -1.86 and a kurtosis of 2.93. See Table 3 for mean Fleiss' Kappa coefficient estimates and mean percentage agreement values for each individual behaviour in this category, collapsed across each Phase, and Table 4 for values specific to each individual behaviour within each Phase.

Table 3

*Estimated Reliability and Agreement Means for Each Behavior in the Body Category, Collapsed Across Three Subjects*

| Behaviour | *k M* | *k* SD | *PA* (Tol =0) | | *PA* (Tol =1) | |
|---|---|---|---|---|---|---|
| | | | *M* | SD | *M* | SD |
| Lean Forward | 0.183 | 0.13 | 95.93 | 2.76 | 99.83 | 0.29 |
| Lean Back | 0.147 | 0.07 | 94.50 | 4.22 | 99.83 | 0.29 |
| Lean Side | 0.071 | 0.08 | 93.27 | 4.40 | 99.83 | 0.29 |
| Rotate Away | 0.045 | 0.07 | 98.47 | 1.39 | 99.83 | 0.29 |
| Protective Object | 0.122 | 0.19 | 96.40 | 4.76 | 99.83 | 0.29 |
| Body Fidget | 0.073 | 0.02 | 78.10 | 8.77 | 96.85 | 0.21 |
| Fidgeting | 0.485 | | 54.30 | | 93.30 | |

The ICC inter-rater reliability coefficients for the overall *Body* category were estimated to range from ICC(2,8) = .146 to ICC(2,8) = .855 across all three subjects, with an estimated grand mean coefficient of ICC(2,8) = .350. The inter-rater reliability coefficient estimate for the *Body* category was found to be slight during Phase 2A, (ICC(2,8) = .146), with a 95% confidence interval ranging from -0.01 to 0.61, ($F(5,51) = 2.64$, $p = .344$). During Phase 2B, the inter-rater reliability coefficient estimate for the *Body* category was also found to be moderate, (ICC(2,8) = .048), with a 95% confidence interval ranging from -0.04 to 0.43, ($F(5,52) = 1.58$, $p = .183$). However, estimates improved for the *Body* category during Phase 2C and were found to be almost perfect, (ICC(2,8) = .855), with a 95% confidence interval ranging from 0.66 to 0.97, ($F(5,40) = 54.5$, $p < .001$). A summary of these overall intraclass correlation estimates can be found in Appendix C and a complete reporting of the individual behaviour estimates and values within this category can be found in Table 4.

Table 4:

*Fleiss' Kappa Estimates and Percentage Agreement Values for the Behaviours in the Body Category During Each Phase*

| Behaviour | *k* | *z* | *p* | Percentage Agreement | |
|---|---|---|---|---|---|
| | | | | *Tol* =0 | *Tol* = 1 |
| Phase 2A: Subject 1 | | | | | |
| Lean Forward | 0.327 | 37.90 | < .001 | 95.7 | 99.50 |
| Lean Back | 0.228 | 24.90 | < .001 | 96.2 | 99.50 |
| Lean Side | 0.061 | 6.54 | < .001 | 97.6 | 99.50 |
| Rotate Away | 0.093 | 9.96 | < .001 | 98.1 | 99.50 |
| Protective Object | -0.009 | -0.88 | 0.381 | 91 | 99.50 |
| Body Fidget | 0.087 | 9.66 | < .001 | 71.9 | 96.70 |
| Fidgeting | | | | | |
| Phase 2B: Subject 2 | | | | | |

| Behaviour | *k* | *z* | *p* | Percentage Agreement *Tol* =0 | *Tol* = 1 |
|---|---|---|---|---|---|
| Lean Forward | 0.109 | 9.45 | < .001 | 98.8 | 100 |
| Lean Back | 0.123 | 10.70 | < .001 | 97.6 | 100 |
| Lean Side | -0.007 | -0.58 | 0.564 | 93.4 | 100 |
| Rotate Away | - | - | - | 100 | 100 |
| Protective Object | - | - | - | 100 | 100 |
| Body Fidget | 0.059 | 5.81 | < .001 | 84.3 | 97 |
| Fidgeting | | | | | |
| Phase 2C: Subject 3 | | | | | |
| Lean Forward | 0.113 | 8.91 | < .001 | 93 | 100 |
| Lean Back | 0.091 | 7.18 | < .001 | 89.70 | 100 |
| Lean Side | 0.160 | 12.70 | < .001 | 88.8 | 100 |
| Rotate Away | -0.003 | -0.27 | 0.790 | 97.30 | 100 |
| Protective Object | 0.253 | 20.00 | < .001 | 98.2 | 100 |
| Body Fidget | | | | | |
| Fidgeting | 0.485 | 41.70 | < .001 | 54.30 | 93.30 |

*Note:* There were 210 five-second intervals during Phase 2A, 166 during Phase 2B, and 223 during Phase 2C.

**Limbs**

The individual behaviours that comprise the *Limbs* category are: *Arms Crossed, One Arm Crossed, Two Arms Crossed, Arms Akimbo Head, Arms Akimbo Hips, Legs Crossed Tight, Legs Open Stance, Leg Resting on Other, Tucked on Seat, Leg Kicking, Knees Up,* and *Legs Other.* Within the *Limbs* category there was no single behaviour that was coded as being the highest or lowest frequency across all three subjects. The frequency at which each *Limbs* behaviour was coded during each Phase, along with mean frequency, SD, median, range, skewness, and kurtosis for each can be found within Appendix D.

The distribution of Fleiss' Kappa coefficient estimates for the *Limbs* category, collapsed across subjects and behaviours, was found to center around 0.36, with a grand mean of 0.35, a median of 0.38, and a standard deviation of 0.27. Reliability estimates ranged from poor (-0.002) for the *Leg Resting on Other* behaviour to substantial (0.71) for the *Legs Open Stance* behaviour, with a skewness of -1.11 and a kurtosis of -1.44. The distribution of percentage agreement values with a tolerance of zero for the *Limbs* category, collapsed across subjects and behaviours, was found to center around 98, with a grand mean of 98.25, a median of 98.98, and a standard deviation of 2.09. Values ranged from 93.27 for the *Leg Kicking* behaviour to 100.00 for the *Arms Akimbo Hips, Tucked on Seat, Knees up,* and *Legs Other* behaviour, with a skewness of -

1.41 and a kurtosis of 1.75. See Table 5 for mean Fleiss' Kappa coefficient estimates and mean percentage agreement values for each individual behaviour in this category, collapsed across each Phase, and Table 6 for values specific to each individual behaviour within each Phase.

Table 5

*Estimated Reliability and Agreement Means for Each Behaviour in the Limbs Category, Collapsed Across Three Subjects*

| | | | *PA* (Tol =0) | | *PA* (Tol =1) | |
|---|---|---|---|---|---|---|
| Behaviour | *k M* | *k* SD | *M* | SD | *M* | SD |
| Arms Crossed | 0.525 | | 97.30 | | 100.00 | |
| One Arm Crossed | 0.167 | | 95.95 | 5.73 | 99.75 | 0.35 |
| Two Arms Crossed | 0.37 | | 98.35 | 2.33 | 99.75 | 0.35 |
| Arms Akimbo Head | 0.394 | 0.07 | 99.27 | 0.64 | 100.00 | 0.00 |
| Arms Akimbo Hips | | | 100.00 | 0.00 | 100.00 | 0.00 |
| Legs Crossed Tight | 0.614 | | 96.90 | | 100.00 | |
| Legs Open Stance | 0.713 | | 99.60 | | 100.00 | |
| Leg Resting on | -0.002 | | 98.70 | | 100.00 | |
| Tucked on Seat | | | 100.00 | | 100.00 | |
| Leg Kicking | 0.049 | | 93.27 | 11.66 | 99.57 | 0.75 |
| Knees Up | | | 100.00 | | 100.00 | |
| Legs Other | | | 100.00 | | 100.00 | |

The ICC inter-rater reliability coefficients for the overall *Limbs* category were estimated to range from ICC(2,8) = .117 to ICC(2,8) = .494 across all three subjects, with an estimated grand mean coefficient of ICC(2,8) = .305. The inter-rater reliability coefficient estimate for the *Limbs* category was found to be fair during Phase 2A, (ICC(2,8) = .305), with a 95% confidence interval ranging from 0.09 to 0.76, ($F(5,54) = 5.82$, $p$ <.001). Similarly, during Phase 2B, the inter-rater reliability coefficient estimate for the *Limbs* category was found to be moderate, (ICC(2,8) = .049), with a 95% confidence interval ranging from 0.22 to 0.87, ($F(5,54) = 10.8$, $p$ < .001. However, estimates declined for the *Limbs* category during Phase 2C and were found to be slight, (ICC(2,8) = .117), with a 95% confidence interval ranging from -0.01 to 0.43, ($F(9,68) = 2.06$, $p$ < .046). A summary of these overall intraclass correlation estimates can be found in Appendix C and a complete reporting of the individual behaviour estimates and values within this category can be found in Table 6.

Table 6

*Fleiss' Kappa Estimates and Percentage Agreement Values for the Behaviours in the Hands Category During Each Phase*

| | | | | Percentage Agreement | |
|---|---|---|---|---|---|
| Behaviour | *k* | *z* | *p* | *Tol* = 0 | *Tol* = 1 |
| Phase 2A: Subject 1 | | | | | |
| Arms Crossed | | | | | |
| One Arm Crossed | 0.167 | 18.30 | < .001 | 91.9 | 99.50 |
| Two Arms Crossed | 0.370 | 42.70 | < .001 | 96.7 | 99.50 |
| Arms Akimbo Head | 0.442 | 42.90 | < .001 | 99 | 100.00 |
| Arms Akimbo Hips | - | - | - | 100 | 100.00 |
| Legs Crossed Tight | | | | | |
| Legs Open Stance | | | | | |
| Leg Resting on Other | | | | | |
| Tucked on Seat | | | | | |
| Leg Kicking | - | - | - | 100 | 100.00 |
| Knees Up | | | | | |
| Legs Other | | | | | |
| Phase 2B: Subject 2 | | | | | |
| Arms Crossed | | | | | |
| One Arm Crossed | - | - | - | 100 | 100 |
| Two Arms Crossed | - | - | - | 100 | 100 |
| Arms Akimbo Head | 0.346 | 29.9 | < .001 | 98.8 | 100 |
| Arms Akimbo Hips | - | - | - | 100 | 100 |
| Legs Crossed Tight | | | | | |
| Legs Open Stance | | | | | |
| Leg Resting on Other | | | | | |
| Tucked on Seat | | | | | |
| Leg Kicking | - | - | - | 100 | 100 |
| Knees Up | | | | | |
| Legs Other | | | | | |
| Phase 2C: Subject 3 | | | | | |
| Arms Crossed | 0.525 | 41.50 | 0 | 97.3 | 100 |
| One Arm Crossed | | | | | |
| Two Arms Crossed | | | | | |
| Arms Akimbo Head | - | - | - | 100.00 | 100 |
| Arms Akimbo Hips | - | - | - | 100 | 100 |
| Legs Crossed Tight | 0.614 | 48.6 | < .001 | 96.9 | 100 |
| Legs Open Stance | 0.713 | 56.4 | < .001 | 99.6 | 100 |
| Leg Resting on Other | -0.002 | -0.133 | 0.894 | 98.7 | 100 |
| Tucked on Seat | - | - | - | 100 | 100 |
| Leg Kicking | 0.049 | 4.02 | < .001 | 79.8 | 98.7 |

| Behaviour | *k* | *z* | *p* | Percentage Agreement *Tol* = 0 | *Tol* = 1 |
|---|---|---|---|---|---|
| Knees Up | - | - | - | 100 | 100 |
| Legs Other | - | - | - | 100 | 100 |

*Note:* There were 210 five-second intervals during Phase 2A, 166 during Phase 2B, and 223 during Phase 2C.

**Head & Shoulders**

The individual behaviours that comprise the *Head and Shoulders* category are: *Full Shrug*, *Shrug Without Arms*, *Asymmetric Shrug*, *Head Silt Tilt*, *Head Tilt Up*, *Head Tilt Down*, *Head Down and Away*, *Head Nod*, *Head Shake*, and *Head Averted.* Within the *Head and Shoulders* category, the most frequently coded behaviour across the first two subjects was *Head Nod*, while *Head Side Tilt* was the most frequently coded behaviour for Subject 3. The behaviours *Full Shrug* and *Head Down and Away* were found to be the lowest frequency behaviours across all three subjects. The frequency at which each *Head and Shoulders* behaviour was coded during each Phase, along with mean frequency, SD, median, range, skewness, and kurtosis for each can be found within Appendix D.

The distribution of Fleiss' Kappa coefficient estimates for the *Head and Shoulders* category, collapsed across subjects and behaviours, was found to center around 0.2, with a grand mean of 0.21, a median of 0.18, and a standard deviation of 0.12. Reliability estimates ranged from slight (0.05) for the *Head Tilt Down* behaviour to moderate (0.44) for the *Head Shake* behaviour, with a skewness of 0.74 and a kurtosis of -0.15. The distribution of percentage agreement values with a tolerance of zero for the *Head and Shoulders* category, collapsed across subjects and behaviours, was found to center around 84, with a grand mean of 82.65, a median of 86.08, and a standard deviation of 16.21. Values ranged from 49.43 for the *Head Nod* behaviour to 98.17 for the *Full Shrug* behaviour, with a skewness of -1.21 and a kurtosis of 0.49. At a tolerance of 1, the percentage agreement values for the *Head and Shoulders* category, collapsed across subjects and behaviours, was found to center around 99, with a grand mean of 98.11, a median of 99.30, and a standard deviation of 2.86. Values ranged from 90.53 for the *Head Nod* behaviour to 99.83 for the *Shrug Without Arms,* and *Asymmetric Shrug* behaviours, with a skewness of -2.47 and a kurtosis of 6.56. See Table 7 for mean Fleiss' Kappa coefficient estimates and mean percentage agreement values for each individual behaviour in this category, collapsed across each Phase, and Table 8 for values specific to each individual behaviour within each Phase.

Table 7

*Estimated Reliability and Agreement Means for Each Behaviour in the Head and Shoulders Category, Collapsed Across Three Subjects*

| | | | *PA (Tol =0)* | | *PA (Tol =1)* | |
|---|---|---|---|---|---|---|
| Behaviour | *k M* | *k SD* | *M* | *SD* | *M* | *SD* |
| Full Shrug | 0.13 | 0.16 | 98.17 | 1.80 | 99.70 | 0.27 |
| Shrug Without Arms | 0.124 | 0.11 | 95.57 | 3.20 | 99.83 | 0.29 |
| Asymmetric Shrug | 0.18 | 0.10 | 96.77 | 2.60 | 99.83 | 0.29 |
| Head Side Tilt | 0.321 | 0.07 | 62.93 | 6.70 | 97.13 | 1.65 |
| Head Tilt Up | 0.187 | 0.07 | 86.77 | 9.96 | 99.07 | 0.90 |
| Head Tilt Down | 0.052 | 0.07 | 85.40 | 16.68 | 98.63 | 1.95 |
| Head Down Away | 0.102 | 0.07 | 95.17 | 6.41 | 99.70 | 0.27 |
| Head Nod | 0.329 | 0.09 | 49.43 | 3.76 | 90.53 | 3.03 |
| Head Shake | 0.444 | 0.04 | 73.53 | 7.25 | 97.13 | 0.15 |
| Head Averted | 0.228 | 0.14 | 82.80 | 8.36 | 99.53 | 0.45 |

The ICC inter-rater reliability coefficients for the overall *Head and Shoulders* category were estimated to range from ICC(2,8) = .503 to ICC(2,8) = .875 across all three subjects, with an estimated grand mean coefficient of ICC(2,8) = .700. The inter-rater reliability coefficient estimate for the *Head and Shoulders* category was found to be substantial during Phase 2A, (ICC(2,8) = .721), with a 95% confidence interval ranging from -0.52 to 0.90, ($F$(9,69) = 33, $p$ < .001). During Phase 2B, the inter-rater reliability coefficient estimate for the *Head and Shoulders* category was found to be almost perfect, (ICC(2,8) = .875), with a 95% confidence interval ranging from 0.75 to 0.96, ($F$(9.75) = 83.3, $p$ < .001). However, estimates for the *Head and Shoulders* category during Phase 2C and were found to be moderate, (ICC(2,8) = .503), with a 95% confidence interval ranging from 0.27 to 0.79, ($F$(9,59) = 10.7, $p$ < .001). A summary of these overall intraclass correlation estimates can be found in Appendix C and a complete reporting of the individual behaviour estimates and values within this category can be found in Table 8.

Table 8

*Fleiss' Kappa Estimates and Percentage Agreement Values for the Behaviours in the Hands Category During Each Phase*

| | | | | Percentage Agreement | |
|---|---|---|---|---|---|
| Behaviour | *k* | *z* | *p* | *Tol* = 0 | *Tol* = 1 |
| | | | Phase 2A: Subject 1 | | |
| Full Shrug | 0.0203 | 2.26 | 0.0238 | 98.1 | 99.50 |
| Shrug Without Arms | 0.185 | 20.60 | < .001 | 92.4 | 99.50 |
| Asymmetric Shrug | 0.138 | 15.70 | < .001 | 96.7 | 99.50 |
| Head Side Tilt | 0.272 | 28.40 | < .001 | 66.7 | 97.10 |
| Head Tilt Up | 0.258 | 28.20 | < .001 | 86.2 | 99.00 |
| Head Tilt Down | 0.132 | 14.00 | < .001 | 97.6 | 99.50 |
| Head Down Away | 0.152 | 17.00 | < .001 | 97.6 | 99.50 |
| Head Nod | 0.360 | 37.40 | < .001 | 52.4 | 93.30 |
| Head Shake | 0.438 | 46.40 | < .001 | 80 | 97.10 |
| Head Averted | 0.247 | 25.80 | < .001 | 80 | 99.50 |
| | | | Phase 2B: Subject 2 | | |
| Full Shrug | - | - | - | 100 | 100 |
| Shrug Without Arms | -0.001 | -0.10 | 0.917 | 98.8 | 100 |
| Asymmetric Shrug | 0.110 | 9.51 | < .001 | 99.4 | 100 |
| Head Side Tilt | 0.29 | 25.4 | < .001 | 66.9 | 98.8 |
| Head Tilt Up | 0.173 | 14.9 | < .001 | 97 | 100 |
| Head Tilt Down | -0.008 | -0.682 | 0.495 | 92.2 | 100 |
| Head Down Away | - | - | - | 100 | 100 |
| Head Nod | 0.403 | 40.5 | < .001 | 45.2 | 87.3 |
| Head Shake | 0.482 | 42.7 | < .001 | 65.7 | 97 |
| Head Averted | 0.0821 | 7.1 | < .001 | 92.2 | 100 |
| | | | Phase 2C: Subject 3 | | |
| Full Shrug | 0.239 | 19.90 | < .001 | 96.4 | 99.6 |
| Shrug Without Arms | 0.187 | 14.80 | < .001 | 95.50 | 100 |
| Asymmetric Shrug | 0.292 | 23.10 | < .001 | 94.2 | 100 |
| Head Side Tilt | 0.401 | 33.20 | < .001 | 55.20 | 95.5 |
| Head Tilt Up | 0.130 | 10.70 | < .001 | 77.1 | 98.2 |
| Head Tilt Down | 0.033 | 2.75 | 0.006 | 66.4 | 96.4 |
| Head Down Away | 0.053 | 4.26 | < .001 | 87.9 | 99.6 |
| Head Nod | 0.225 | 19.2 | < .001 | 50.7 | 91 |
| Head Shake | 0.411 | 34.1 | < .001 | 74.9 | 97.3 |
| Head Averted | 0.355 | 28.4 | < .001 | 76.2 | 99.1 |

*Note:* There were 210 five-second intervals during Phase 2A, 166 during Phase 2B, and 223 during Phase 2C.

**Face & Eyes**

The individual behaviours that comprise the *Face and Eyes* category are: *Lip Bite, Lip Lick, Pursed Lips, Biting Self, Biting Object, Oral Fixation Self, Oral Fixation Object Gaze Aversion, Eyes Looking Away, Eyes Searching, Blink, Full Eye Closure,* and *Eye Rolling.* Within the *Face and eyes* category, the most frequently coded behaviour across all three subjects was *Blinking*, while the lowest frequency behaviours were *Biting Self*, *Biting Object,* and *Oral Fixation Object*. The frequency at which each *Face and Eyes* behaviour was coded during each Phase, along with mean frequency, SD, median, range, skewness, and kurtosis for each can be found within Appendix D.

The distribution of Fleiss' Kappa coefficient estimates for the *Face and Eyes* category, collapsed across subjects and behaviours, was found to center around 0.18, with a grand mean of 0.19, a median of 0.17, and a standard deviation of 0.14. Reliability estimates ranged from poor (0.00) for the *Biting Self* behaviour to moderate (0.45) for the *Lip Lick* behaviour, with a skewness of 0.61 and a kurtosis of -0.32. The distribution of percentage agreement values with a tolerance of zero for the *Face and Eyes* category, collapsed across subjects and behaviours, was found to center around 75, with a grand mean of 69.91, a median of 82.40, and a standard deviation of 34.96. Values ranged from 4.30 for the *Blinking* behaviour to 100.00 for the *Oral Fixation Object* behaviour, with a skewness of -1.04 and a kurtosis of -0.24. At a tolerance of 1, the percentage agreement values for the *Face and Eyes* category, collapsed across subjects and behaviours, was found to center around 98, with a grand mean of 97.04, a median of 99.07, and a standard deviation of 3.51. Values ranged from 91.27 for the *Blinking* behaviour to 100.00 for the *Lip Bite, Biting Self, Biting Other,* and *Oral Fixation Object* behaviours, with a skewness of -0.63 and a kurtosis of -1.61. See Table 9 for mean Fleiss' Kappa coefficient estimates and mean percentage agreement values for each individual behaviour in this category, collapsed across each Phase, and Table 10 for values specific to each individual behaviour within each Phase.

Table 9

*Estimated Reliability and Agreement Means for Each Behaviour in the Face and Eyes Category, Collapsed Across Three Subjects*

| | | | *PA (Tol =0)* | | *PA (Tol =1)* | |
|---|---|---|---|---|---|---|
| Behaviour | *k M* | *k SD* | *M* | SD | *M* | SD |
| Lip Bite | 0.357 | 0.56 | 98.87 | 1.21 | 100.00 | 0.00 |
| Lip Lick | 0.449 | 0.05 | 78.03 | 3.91 | 98.53 | 0.38 |

| Behaviour | *k M* | *k SD* | *PA (Tol =0)* *M* | SD | *PA (Tol =1)* *M* | SD |
|---|---|---|---|---|---|---|
| Pursed Lips | 0.165 | 0.11 | 82.40 | 5.91 | 99.53 | 0.50 |
| Biting Self | -0.001 | - | 99.87 | 0.23 | 100.00 | 0.00 |
| Biting Other | - | - | 100.00 | 0.00 | 100.00 | 0.00 |
| Oral Fixation Self | 0.065 | 0.11 | 96.20 | 5.10 | 99.87 | 0.23 |
| Oral Fixation Object | - | - | 100.00 | 0.00 | 100.00 | 0.00 |
| Gaze Aversion | 0.147 | | 6.73 | | 92.80 | |
| Eyes Looking Away | 0.184 | 0.02 | 33.85 | 4.03 | 92.35 | 8.84 |
| Eyes Searching | 0.058 | 0.03 | 54.60 | 7.92 | 94.50 | 0.99 |
| Blink | 0.237 | 0.04 | 4.30 | 1.49 | 91.27 | 3.96 |
| Full Eye Closure | 0.3 | 0.09 | 65.27 | 6.68 | 93.57 | 4.07 |
| Eye Rolling | 0.102 | 0.04 | 88.77 | 9.02 | 99.07 | 0.95 |

The ICC inter-rater reliability coefficients for the overall *Face and Eyes* category were estimated to range from ICC(2,8) = .850 to ICC(2,8) = .944 across all three subjects, with an estimated grand mean coefficient of ICC(2,8) = .903. The inter-rater reliability coefficient estimate for the *Face and Eyes* category was found to be almost perfect during Phase 2A, (ICC(2,8) = .915), with a 95% confidence interval ranging from 0.84 to 0.97, ($F(11,102) = 118$, $p < .001$). Similarly, during Phase 2B, the inter-rater reliability coefficient estimate for the *Face and Eyes* category was also found to be almost perfect, (ICC(2,8) = .944), with a 95% confidence interval ranging from 0.89 to 0.98, ($F(11,106) = 177$, $p < .001$). Estimates for the *Face and Eyes* category during Phase 2C and were also found to be almost perfect, (ICC(2,8) = .850), with a 95% confidence interval ranging from 0.71 to 0.95, ($F(10,70) = 46.5$, $p < .001$). A summary of these overall intraclass correlation estimates can be found in Appendix C and a complete reporting of the individual behaviour estimates and values within this category can be found in Table 10.

Table 10

*Fleiss' Kappa Estimates and Percentage Agreement Values for the Behaviours in the Face and Eyes Category During Each Phase*

| Behaviour | *k* | *z* | *P* | Percentage Agreement *Tol* = 0 | *Tol* = 1 |
|---|---|---|---|---|---|
| | | Phase 2A: Subject 1 | | | |
| Lip Bite | -0.001 | -0.09 | 0.926 | 99 | 100.00 |
| Lip Lick | 0.463 | 47.70 | < .001 | 74.3 | 98.10 |
| Pursed Lips | 0.287 | 29.90 | < .001 | 79.5 | 99.00 |

| Behaviour | *k* | *z* | *P* | Percentage Agreement *Tol* = 0 | Percentage Agreement *Tol* = 1 |
|---|---|---|---|---|---|
| Biting Self | - | - | - | 100 | 100.00 |
| Biting Other | - | - | - | 100 | 100.00 |
| Oral Fixation Self | - | - | - | 100 | 100.00 |
| Oral Fixation Object | - | - | - | 100 | 100.00 |
| Gaze Aversion | | | | | |
| Eyes Looking Away | 0.198 | 23.20 | < .001 | 31 | 98.60 |
| Eyes Searching | 0.038 | 3.99 | < .001 | 49 | 93.80 |
| Blink | 0.227 | 42.20 | < .001 | 2.86 | 86.7 |
| Full Eye Closure | 0.399 | 43.00 | < .001 | 62.4 | 91.9 |
| Eye Rolling | 0.090 | 9.64 | < .001 | 78.6 | 98.1 |
| | | Phase 2B: Subject 2 | | | |
| Lip Bite | 0.0707 | 6.11 | < .001 | 97.6 | 100 |
| Lip Lick | 0.391 | 34.30 | < .001 | 77.7 | 98.8 |
| Pursed Lips | 0.071 | 6.13 | < .001 | 89.2 | 100 |
| Biting Self | - | - | - | 100 | 100 |
| Biting Other | - | - | - | 100 | 100 |
| Oral Fixation Self | -0.01 | -0.841 | 0.4 | 90.4 | 100 |
| Oral Fixation Object | - | - | - | 100 | 100 |
| Gaze Aversion | | | | | |
| Eyes Looking Away | 0.17 | 16.3 | < .001 | 36.7 | 86.1 |
| Eyes Searching | 0.078 | 7.13 | < .001 | 60.2 | 95.2 |
| Blink | 0.28 | 45.4 | < .001 | 4.22 | 93.4 |
| Full Eye Closure | 0.263 | 23.5 | < .001 | 72.9 | 98.2 |
| Eye Rolling | 0.069 | 5.97 | < .001 | 95.8 | 100 |
| | | Phase 2C: Subject 3 | | | |
| Lip Bite | 1.000 | 79.00 | < .001 | 100 | 100 |
| Lip Lick | 0.494 | 40.90 | < .001 | 82.10 | 98.7 |
| Pursed Lips | 0.138 | 11.00 | < .001 | 78.5 | 99.6 |
| Biting Self | -0.001 | -0.04 | 0.965 | 99.6 | 100 |
| Biting Other | - | - | - | 100 | 100 |
| Oral Fixation Self | 0.140 | 12.50 | < .001 | 98.2 | 99.6 |
| Oral Fixation Object | - | - | - | 100 | 100.000 |
| Gaze Aversion | 6.730 | 17.80 | < .001 | 6.730 | 92.800 |
| Eyes Looking Away | | | | | |
| Eyes Searching | | | | | |
| Blink | 0.203 | 27.7 | < .001 | 5.83 | 93.7 |
| Full Eye Closure | 0.239 | 21.2 | < .001 | 60.5 | 90.6 |
| Eye Rolling | 0.148 | 12.3 | < .001 | 91.9 | 99.1 |

*Note:* There were 210 five-second intervals during Phase 2A, 166 during Phase 2B, and 223 during Phase 2C.

**Reflexive**

The individual behaviours that comprise the *Reflexive* category are: *Deep Inhale*, *Sigh, Yawn, Sniff, Clear Throat, Cough,* and *Hard Swallow*. Within the *Reflexive* category, the most frequently coded behaviour across all three subjects was *Deep Inhale,* while the lowest frequency behaviour was *Cough*. The frequency at which each *Reflexive* behaviour was coded during each Phase, along with mean frequency, SD, median, range, skewness, and kurtosis for each can be found within Appendix D.

The distribution of Fleiss' Kappa coefficient estimates for the *Reflexive* category, collapsed across subjects and behaviours, was found to center around 0.17, with a grand mean of 0.15, a median of 0.20, and a standard deviation of 0.11. Reliability estimates ranged from poor (0.00) for *Yawn* behaviour to fair (0.25) for the *Deep Inhale* behaviour, with a skewness of -0.84 and a kurtosis of -1.42. The distribution of percentage agreement values with a tolerance of zero for the *Reflexive* category, collapsed across subjects and behaviours, was found to center around 95, with a grand mean of 94.65, a median of 96.27, and a standard deviation of 7.35. Values ranged from 78.60 for the *Deep Inhale* behaviour to 100.00 for the *Cough* behaviour, with a skewness of -2.24 and a kurtosis of 5.44. At a tolerance of 1, the percentage agreement values for the *Reflexive* category, collapsed across subjects and behaviours, was found to center around 99, with a grand mean of 99.76, a median of 99.83, and a standard deviation of 0.23. Values ranged from 99.40 for the *Deep Inhale* behaviour to 100.00 for the *Yawn* and *Cough* behaviours, with a skewness of -0.54 and a kurtosis of -0.90. See Table 11 for mean Fleiss' Kappa coefficient estimates and mean percentage agreement values for each individual behaviour in this category, collapsed across each Phase, and Table 12 for values specific to each individual behaviour within each Phase.

Table 11

*Estimated Reliability and Agreement Means for Each Behaviour in the Reflexive Category, Collapsed Across Three Subjects*

| | | | *PA (Tol =0)* | | *PA (Tol =0)* | |
|---|---|---|---|---|---|---|
| Behaviour | *k M* | *k* SD | *M* | SD | *M* | SD |
| Deep Inhale | 0.253 | 0.03 | 78.60 | 16.93 | 99.40 | 0.66 |
| Sigh | 0.178 | 0.12 | 95.50 | 1.93 | 99.53 | 0.50 |

| Behaviour | *k M* | *k* SD | *PA (Tol =0)* | | *PA (Tol =0)* | |
|---|---|---|---|---|---|---|
| | | | *M* | SD | *M* | SD |
| Yawn | -0.002 | - | 99.57 | 0.75 | 100.00 | 0.00 |
| Sniff | 0.046 | 0.08 | 96.27 | 3.35 | 99.83 | 0.29 |
| Clear Throat | 0.216 | 0.29 | 97.93 | 2.72 | 99.83 | 0.29 |
| Cough | - | - | 100.00 | 0.00 | 100.00 | 0.00 |
| Hard Swallow | 0.232 | 0.19 | 94.70 | 6.19 | 99.70 | 0.27 |

The ICC inter-rater reliability coefficients for the overall *Reflexive* category were estimated to range from ICC(2,8) = .342 to ICC(2,8) = .523 across all three subjects, with an estimated grand mean coefficient of ICC(2,8) = .426. The inter-rater reliability coefficient estimate for the *Reflexive* category was found to be fair during Phase 2A, (ICC(2,8) = .342), with a 95% confidence interval ranging from 0.13 to 0.74, ($F(6,55) = 7.51$, $p < .001$. Similarly, during Phase 2B, the inter-rater reliability coefficient estimate for the *Reflexive* category was found to be moderate, (ICC(2,8) = .413), with a 95% confidence interval ranging from 0.18 to 0.79, ($F(6,62) = 8.56$, $p < .001$). Estimates for the *Reflexive* category during Phase 2C and were also found to be moderate, (ICC(2,8) = .523), with a 95% confidence interval ranging from 0.25 to 0.86, ($F(6,49) = 10.2$, $p < .001$). A summary of these overall intraclass correlation estimates can be found in Appendix C and a complete reporting of the individual behaviour estimates and values within this category can be found in Table 12.

Table 12

*Fleiss' Kappa Estimates and Percentage Agreement Values for the Behaviours in the Reflexive Category During Each Phase*

| Behaviour | *k* | *z* | *P* | Percentage Agreement | |
|---|---|---|---|---|---|
| | | | | *Tol* = 0 | *Tol* = 1 |
| | | Phase 2A: Subject 1 | | | |
| Deep Inhale | 0.258 | 27.90 | < .001 | 86.2 | 99.50 |
| Sigh | 0.267 | 30.50 | < .001 | 93.8 | 99.00 |
| Yawn | - | - | - | 100 | 100.00 |
| Sniff | 0.141 | 14.70 | < .001 | 92.4 | 99.50 |
| Clear Throat | 0.540 | 60.40 | < .001 | 94.8 | 99.50 |
| Cough | - | - | - | 100 | 100.00 |
| Hard Swallow | 0.098 | 10.60 | < .001 | 96.2 | 99.50 |
| | | Phase 2B: Subject 2 | | | |
| Deep Inhale | 0.224 | 19.4 | < .001 | 90.4 | 100 |
| Sigh | 0.042 | 3.59 | < .001 | 97.6 | 100 |

| | | | | Percentage Agreement | |
|---|---|---|---|---|---|
| Behaviour | *k* | *z* | *P* | *Tol* = 0 | *Tol* = 1 |
| Yawn | - | - | - | 100 | 100 |
| Sniff | -0.002 | -0.16 | 0.876 | 98.2 | 100 |
| Clear Throat | 0.110 | 9.51 | < .001 | 99.4 | 100 |
| Cough | - | - | - | 100 | 100 |
| Hard Swallow | - | - | - | 100 | 100 |
| Phase 2C: Subject 3 | | | | | |
| Deep Inhale | 0.277 | 22.10 | < .001 | 59.2 | 98.7 |
| Sigh | 0.225 | 18.50 | < .001 | 95.10 | 99.6 |
| Yawn | -0.002 | -0.13 | 0.894 | 98.7 | 100 |
| Sniff | -0.002 | -0.18 | 0.859 | 98.2 | 100 |
| Clear Throat | -0.001 | -0.04 | 0.965 | 99.6 | 100 |
| Cough | - | - | - | 100 | 100 |
| Hard Swallow | 0.366 | 29.20 | < .001 | 87.9 | 99.600 |

*Note:* There were 210 five-second intervals during Phase 2A, 166 during Phase 2B, and 223 during Phase 2C.

**Paralinguistics I**

The individual behaviours that comprise the *Paralinguistics I* category are: *Pause, Latency, Dissociation, Fillers, Speech Errors, Laughter, Vocal Raises, Interruptions,* and *Scoff.* Within the *Paralinguistics I* category, the most frequently coded behaviour across all three subjects was *Fillers*, while the lowest frequency behaviour was *Dissociation*. The frequency at which each *Paralinguistics I* behaviour was coded during each Phase, along with mean frequency, SD, median, range, skewness, and kurtosis for each can be found within Appendix D.

The distribution of Fleiss' Kappa coefficient estimates for the *Paralinguistics I* category, collapsed across subjects and behaviours, was found to center around 0.29, with a grand mean of 0.29, a median of 0.28, and a standard deviation of 0.13. Reliability estimates ranged from slight (0.12) for the *Vocal Raises* behaviour to moderate (0.45) for the *Laughter* behaviour, with a skewness of 0.09 and a kurtosis of -1.75. The distribution of percentage agreement values with a tolerance of zero for the *Paralinguistics I* category, collapsed across subjects and behaviours, was found to center around 86, with a grand mean of 80.51, a median of 92.30, and a standard deviation of 20.73. Values ranged from 37.83 for the *Fillers* behaviour to 100.00 for the *Scoff* and *Dissociation* behaviours, with a skewness of -1.17 and a kurtosis of 0.91. At a tolerance of 1, the percentage agreement values for the *Paralinguistics I* category, collapsed across subjects and behaviours, was found to center around 97, with a grand mean of 96.38, a median of 99.63, and a standard deviation of 6.51. Values ranged from 79.63 for the *Fillers* behaviour to 100.00 for *Dissociation* and *Scoff* behaviours, with a skewness of -2.63 and a kurtosis of 7.24. See Table 13 for mean Fleiss' Kappa coefficient estimates and mean percentage agreement values for each individual behaviour in this category, collapsed across each Phase, and Table 14 for values specific to each individual behaviour within each Phase.

Table 13

*Estimated Reliability and Agreement Means for Each Behaviour in the Paralinguistics I Category, Collapsed Across Three Subjects*

| | | | *PA (Tol =0)* | | *PA (Tol =1)* | |
|---|---|---|---|---|---|---|
| Behaviour | *k M* | *k* SD | *M* | SD | *M* | SD |
| Pause | 0.352 | 0.04 | 71.10 | 3.03 | 97.33 | 0.78 |
| Latency | 0.284 | 0.19 | 92.37 | 2.83 | 99.70 | 0.27 |
| Dissociation | - | - | 100.00 | 0.00 | 100.00 | 0.00 |
| Fillers | 0.43 | 0.06 | 37.83 | 2.74 | 79.63 | 3.64 |
| Speech Errors | 0.19 | 0.06 | 62.50 | 19.01 | 94.73 | 5.64 |
| Laughter | 0.451 | 0.39 | 92.30 | 8.22 | 99.70 | 0.27 |
| Vocal Raises | 0.121 | 0.09 | 76.00 | 2.35 | 97.67 | 1.67 |
| Interruptions | 0.189 | 0.17 | 92.53 | 3.80 | 98.63 | 1.58 |
| Scoff | - | - | 100.00 | 0.00 | 100.00 | 0.00 |

The ICC inter-rater reliability coefficients for the overall *Paralinguistics I* category were estimated to range from ICC(2,8) = .676 to ICC(2,8) = .830 across all three subjects, with an estimated grand mean coefficient of ICC(2,8) = .732. The inter-rater reliability coefficient estimate for the *Paralinguistics I* category was found to be substantial during Phase 2A, (ICC(2,8) = .676), with a 95% confidence interval ranging from 0.45 to 0.89, ($F(8,69) = 25.7$, $p < .001$). During Phase 2B, the inter-rater reliability coefficient estimate for the *Paralinguistics I* category was found to be almost perfect, (ICC(2,8) = .830), with a 95% confidence interval ranging from 0.67 to 0.95, ($F(8,87) = 59.6$, $p < .001$). Estimates for the *Paralinguistics I* category during Phase 2C and were found to be substantial, (ICC(2,8) = .690), with a 95% confidence interval ranging from 0.47 to 0.89, ($F(9,65) = 21$, $p < .001$). A summary of these overall intraclass correlation estimates can be found in Appendix C and a complete reporting of the individual behaviour estimates and values within this category can be found in Table 14.

Table 14

*Fleiss' Kappa Estimates and Percentage Agreement Values for the Behaviours in the Paralinguistics I Category During Each Phase*

| | | | | Percentage Agreement | |
|---|---|---|---|---|---|
| Behaviour | *k* | *z* | *P* | *Tol* = 0 | *Tol* = 1 |
| | | | Phase 2A: Subject 1 | | |
| Pause | 0.357 | 36.80 | < .001 | 67.6 | 97.10 |
| Latency | 0.324 | 35.70 | < .001 | 90 | 99.50 |
| Dissociation | - | - | - | 100 | 100.00 |
| Fillers | 0.397 | 44.60 | < .001 | 36.2 | 76.70 |
| Speech Errors | 0.132 | 14.10 | < .001 | 76.2 | 97.10 |
| Laughter | 0.645 | 65.10 | < .001 | 83.3 | 99.50 |
| Vocal Raises | 0.197 | 20.10 | < .001 | 73.3 | 99.00 |
| Interruptions | 0.258 | 29.60 | < .001 | 92.4 | 99.00 |
| Scoff | - | - | - | 100 | 100.00 |
| | | | Phase 2B: Subject 2 | | |
| Pause | 0.308 | 27.3 | < .001 | 72.9 | 98.2 |
| Latency | 0.447 | 38.60 | < .001 | 91.6 | 100 |
| Dissociation | - | - | - | 100 | 100 |
| Fillers | 0.503 | 51.00 | < .001 | 41 | 83.7 |
| Speech Errors | 0.182 | 15.90 | < .001 | 70.5 | 98.8 |
| Laughter | -0.001 | -0.05 | 0.958 | 99.4 | 100 |
| Vocal Raises | 0.149 | 13.70 | < .001 | 77.1 | 95.8 |
| Interruptions | -0.004 | -0.31 | 0.754 | 96.4 | 100 |
| Scoff | - | - | - | 100 | 100 |
| | | | Phase 2C: Subject 3 | | |
| Pause | 0.390 | 26.30 | < .001 | 72.8 | 96.7 |
| Latency | 0.082 | 6.83 | < .001 | 95.50 | 99.6 |
| Dissociation | - | - | - | 100 | 100 |
| Fillers | 0.390 | 38.00 | < .001 | 36.3 | 78.5 |
| Speech Errors | 0.257 | 22.60 | < .001 | 40.8 | 88.3 |
| Laughter | 0.710 | 56.60 | < .001 | 94.2 | 99.6 |
| Vocal Raises | 0.017 | 1.41 | 0.16 | 77.6 | 98.2 |
| Interruptions | 0.314 | 29.20 | < .001 | 88.8 | 96.9 |
| Scoff | | | | | |

*Note:* There were 210 five-second intervals during Phase 2A, 166 during Phase 2B, and 223 during Phase 2C.

## Paralinguistics II

The individual behaviours that comprise the *Paralinguistics II* category are: *Corrections, Verbal Diversions, Generalizations, Lack of Memory, Lack of Memory Recovered, Lack of*

*Memory Reasons, Dismissiveness, Humor, Extreme Description, Pardon Self, Pardon Other,* and *Sarcasm.* Within the *Paralinguistics II* category, the most frequently coded behaviour across all three subjects was *Generalizations*, while the lowest frequency behaviour was *Pardon Self*. The frequency at which each *Paralinguistics II* behaviour was coded during each Phase, along with mean frequency, SD, median, range, skewness, and kurtosis for each can be found within Appendix D.

The distribution of Fleiss' Kappa coefficient estimates for the *Paralinguistics II* category, collapsed across subjects and behaviours, was found to center around 0.13, with a grand mean of 0.12, a median of 0.15, and a standard deviation of 0.09. Reliability estimates ranged from poor (0.00) for the *Pardon Self* behaviour to fair (0.26) for the *Lack of Memory* behaviour, with a skewness of 0.10 and a kurtosis of -1.46. The distribution of percentage agreement values with a tolerance of zero for the *Paralinguistics II* category, collapsed across subjects and behaviours, was found to center around 94, with a grand mean of 93.81, a median of 94.63, and a standard deviation of 4.35. Values ranged from 86.93 for the *Generalizations* behaviour to 100.00 for the *Sarcasm* behaviour, with a skewness of -0.50 and a kurtosis of -0.62. At a tolerance of 1, the percentage agreement values for the *Paralinguistics II* category, collapsed across subjects and behaviours, was found to center around 99, with a grand mean of 99.71, a median of 99.83, and a standard deviation of 0.29. Values ranged from 98.93 for the *Extreme Descriptions* behaviour to 100.00 for the *Pardon Other* and *Sarcasm* behaviours, with a skewness of -1.84 and a kurtosis of 4.25. See Table 15 for mean Fleiss' Kappa coefficient estimates and mean percentage agreement values for each individual behaviour in this category, collapsed across each Phase, and Table 16 for values specific to each individual behaviour within each Phase.

Table 15

*Estimated Reliability and Agreement Means for Each Behaviour in the Paralinguistics II Category, Collapsed Across Three Subjects*

| | | | *PA (Tol =0)* | | *PA (Tol =1)* | |
|---|---|---|---|---|---|---|
| Behaviour | *k M* | *k* SD | *M* | SD | *M* | SD |
| Corrections | 0.171 | 0.11 | 95.67 | 1.65 | 99.70 | 0.27 |
| Verbal Diversions | 0.025 | 0.03 | 87.53 | 7.24 | 99.63 | 0.32 |
| Generalizations | 0.174 | 0.07 | 86.93 | 6.82 | 99.50 | 0.10 |
| Lack of Memory | 0.261 | 0.12 | 87.60 | 5.01 | 99.83 | 0.29 |
| Recovered | 0.014 | 0.02 | 96.20 | 1.60 | 99.83 | 0.29 |
| Reasons | 0.146 | 0.06 | 94.83 | 2.57 | 99.83 | 0.29 |

| Behaviour | *k M* | *k* SD | *PA (Tol =0)* | | *PA (Tol =1)* | |
|---|---|---|---|---|---|---|
| | | | *M* | SD | *M* | SD |
| Dismissiveness | 0.059 | 0.07 | 94.23 | 1.19 | 99.53 | 0.45 |
| Humor | 0.239 | 0.10 | 95.50 | 3.25 | 99.83 | 0.29 |
| Extreme Description | 0.06 | 0.05 | 93.43 | 7.37 | 98.93 | 1.10 |
| Pardon Self | -0.002 | - | 99.37 | 1.10 | 99.83 | 0.29 |
| Pardon Other | 0.171 | 0.18 | 94.43 | 5.42 | 100.00 | 0.00 |
| Sarcasm | - | - | 100.00 | 0.00 | 100.00 | 0.00 |

The ICC inter-rater reliability coefficients for the overall *Paralinguistics II* category were estimated to range from ICC(2,8) = .356 to ICC(2,8) = .386 across all three subjects, with an estimated grand mean coefficient of ICC(2,8) = .371. The inter-rater reliability coefficient estimate for the *Paralinguistics II* category was found to be fair during Phase 2A, (ICC(2,8) = .386), with a 95% confidence interval ranging from 0.20 to 0.67, ($F(11,66) = 9.76$, $p < .001$). Similarly, during Phase 2B, the inter-rater reliability coefficient estimate for the *Paralinguistics II* category was also found to be fair, (ICC(2,8) = .370), with a 95% confidence interval ranging from 0.19 to 0.66, ($F(11,101) = 7.58$, $p < .001$). Estimates for the *Paralinguistics II* category during Phase 2C and were found to be fair once again, (ICC(2,8) = .356), with a 95% confidence interval ranging from 0.16 to 0.67, ($F(10,74) = 5.86$, $p < .001$). A summary of these overall intraclass correlation estimates can be found in Appendix C and a complete reporting of the individual behaviour estimates and values within this category can be found in Table 16.

Table 16

*Fleiss' Kappa Estimates and Percentage Agreement Values for the Behaviours in the Paralinguistics II Category During Each Phase*

| Behaviour | *k* | *z* | *P* | Percentage Agreement | |
|---|---|---|---|---|---|
| | | | | *Tol = 0* | *Tol = 1* |
| | | Phase 2A: Subject 1 | | | |
| Corrections | 0.297 | 34.20 | < .001 | 95.7 | 99.50 |
| Verbal Diversions | 0.0477 | 5.18 | < .001 | 92.9 | 99.50 |
| Generalizations | 0.246 | 26.50 | < .001 | 87.6 | 99.50 |
| Lack of Memory | 0.148 | 15.90 | < .001 | 82.4 | 99.50 |
| Recovered | 0.030 | 3.32 | 0.001 | 96.2 | 99.50 |
| Reasons | 0.210 | 23.00 | < .001 | 93.3 | 99.50 |
| Dismissiveness | 0.046 | 5.19 | < .001 | 92.9 | 99.50 |
| Humor | 0.249 | 28.10 | < .001 | 91.9 | 99.50 |
| Extreme Description | 0.075 | 8.91 | < .001 | 95.7 | 99.00 |

| Behaviour | *k* | *z* | *P* | Percentage Agreement | |
|---|---|---|---|---|---|
| | | | | *Tol = 0* | *Tol = 1* |
| Pardon Self | -0.002 | -0.18 | 0.859 | 98.1 | 99.50 |
| Pardon Other | 0.054 | 5.23 | < .001 | 98.6 | 100.00 |
| Sarcasm | - | - | - | 100 | 100.00 |
| Phase 2B: Subject 2 | | | | | |
| Corrections | 0.11 | 9.55 | < .001 | 94 | 100 |
| Verbal Diversions | -0.009 | -0.83 | 0.407 | 90.4 | 99.4 |
| Generalizations | 0.109 | 10.30 | < .001 | 93.4 | 99.4 |
| Lack of Memory | 0.241 | 20.90 | < .001 | 88 | 100 |
| Recovered | 0.016 | 1.41 | 0.159 | 94.6 | 100 |
| Reasons | 0.109 | 9.38 | < .001 | 93.4 | 100 |
| Dismissiveness | -0.005 | -0.42 | 0.676 | 95.2 | 100 |
| Humor | 0.131 | 11.30 | < .001 | 98.2 | 100 |
| Extreme Description | -0.001 | -0.05 | 0.958 | 99.4 | 100 |
| Pardon Self | - | - | - | 100 | 100 |
| Pardon Other | 0.079 | 6.82 | < .001 | 96.4 | 100 |
| Sarcasm | - | - | - | 100 | 100 |
| Phase 2C: Subject 3 | | | | | |
| Corrections | 0.107 | 9.17 | < .001 | 97.3 | 99.6 |
| Verbal Diversions | 0.036 | 2.82 | 0.005 | 79.30 | 100 |
| Generalizations | 0.167 | 13.30 | < .001 | 79.8 | 99.6 |
| Lack of Memory | 0.393 | 31.00 | < .001 | 92.4 | 100 |
| Recovered | -0.003 | -0.22 | 0.824 | 97.8 | 100 |
| Reasons | 0.119 | 9.40 | < .001 | 97.8 | 100 |
| Dismissiveness | 0.135 | 11.50 | < .001 | 94.6 | 99.1 |
| Humor | 0.337 | 26.70 | < .001 | 96.4 | 100 |
| Extreme Description | 0.105 | 9.25 | < .001 | 85.2 | 97.8 |
| Pardon Self | - | - | - | 100 | 100 |
| Pardon Other | 0.379 | 29.90 | < .001 | 88.3 | 100 |
| Sarcasm | | | | | |

*Note:* There were 210 five-second intervals during Phase 2A, 166 during Phase 2B, and 223 during Phase 2C.

## Incongruence

The individual behaviours with incongruent modifiers that comprise the *Incongruence* category are: *Full Shrug, Shrug Without Arms, Asymmetric Shrug, Head Nod, Head Shake, Humor*, *Hand Shrug, Hand Stop, Laughter, Head Down and Away, Hand Dismiss,* and *Obscene Gesture.* Within the *Incongruence* category, the most frequently coded incongruent behaviour across all three subjects was *Head Shake*. There was no single incongruent behaviour that was coded with the least frequency across all three subjects. The frequency at which each *Incongruence* behaviour was coded during each Phase, along with mean frequency, SD, median, range, skewness, and kurtosis for each can be found within Appendix D.

The distribution of Fleiss' Kappa coefficient estimates for the *Incongruence* category, collapsed across subjects and behaviours, was found to center around 0.13, with a grand mean of 0.15, a median of 0.12, and a standard deviation of 0.11. Reliability estimates ranged from slight (0.05) for the *Head Down and Away* incongruent behaviour to moderate (0.43) for the *Laughter* incongruent behaviour, with a skewness of 1.84 and a kurtosis of 2.94. The distribution of percentage agreement values with a tolerance of zero for the *Incongruence* category, collapsed across subjects and behaviours, was found to center around 96, with a grand mean of 96.40, a median of 97.80, and a standard deviation of 4.12. Values ranged from 84.87 for the *Head Shake* incongruent behaviour to 100.00 for the *Obscene Gesture* incongruent behaviour, with a skewness of -2.28 and a kurtosis of 5.87. At a tolerance of 1, the percentage agreement values for the *Incongruence* category, collapsed across subjects and behaviours, was found to center around 99, with a grand mean of 99.88, a median of 99.87, and a standard deviation of 0.21. Values ranged from 99.70 for the *Hand Shrug* incongruent behaviour to 100.00 for *Shrug Without Arms*, *Asymmetric Shrug*, *Hand Stop*, *Head Down and Away* and *Obscene Gesture* incongruent behaviours, with a skewness of -0.32 and a kurtosis of -1.44. See Table 17 for mean Fleiss' Kappa coefficient estimates and mean percentage agreement values for each individual behaviour in this category, collapsed across each Phase, and Table 18 for values specific to each individual behaviour within each Phase.

Table 17

*Estimated Reliability and Agreement Means for Each Behaviour in the Incongruence Category, Collapsed Across Three Subjects*

| | | | *PA (Tol =0) M* | | *PA (Tol =1) M* | |
|---|---|---|---|---|---|---|
| Behaviour | *k M* | *k* SD | *M* | SD | *M* | SD |
| Full Shrug | 0.083 | 0.12 | 98.63 | 1.58 | 99.87 | 0.23 |
| Shrug Without | 0.093 | 0.08 | 97.50 | 1.95 | 100.00 | 0.00 |
| Asymmetric Shrug | 0.126 | 0.18 | 97.73 | 3.10 | 100.00 | 0.00 |
| Head Nod | 0.064 | 0.09 | 96.00 | 4.42 | 99.87 | 0.23 |
| Head Shake | 0.3 | 0.05 | 84.87 | 3.59 | 99.70 | 0.52 |
| Humor | 0.13 | 0.03 | 97.87 | 1.85 | 99.83 | 0.29 |

| Behaviour | *k M* | *k* SD | *PA (Tol =0) M* | | *PA (Tol =1) M* | |
|---|---|---|---|---|---|---|
| | | | *M* | SD | *M* | SD |
| Hand Shrug | 0.166 | 0.23 | 94.50 | 4.10 | 99.70 | 0.27 |
| Hand Stop | 0.081 | 0.12 | 98.50 | 1.27 | 100.00 | 0.00 |
| Laughter | 0.425 | 0.23 | 93.33 | 6.65 | 99.83 | 0.29 |
| Head Down and | 0.054 | | 99.30 | 0.99 | 100.00 | 0.00 |
| Hand Dismiss | 0.118 | | 98.55 | 2.05 | 99.75 | 0.35 |
| Obscene Gesture | | | 100.00 | 0.00 | 100.00 | 0.00 |

The ICC inter-rater reliability coefficients for the overall *Incongruence* category were estimated to range from ICC(2,8) = .342 to ICC(2,8) = .484 across all three subjects, with an estimated grand mean coefficient of ICC(2,8) = .431. The inter-rater reliability coefficient estimate for the *Incongruence* category was found to be fair during Phase 2A, (ICC(2,8) = .342), with a 95% confidence interval ranging from 0.25 to 0.46, ($F(11,58) = 6.39$, $p < .001$). During Phase 2B, the inter-rater reliability coefficient estimate for the *Incongruence* category was found to be moderate, (ICC(2,8) = .467), with a 95% confidence interval ranging from 0.37 to 0.58, ($F(11,47) = 9.94$, $p < .001$). Estimates for the *Incongruence* category during Phase 2C and were also found to be moderate, (ICC(2,8) = .484), with a 95% confidence interval ranging from 0.23 to 0.81, ($F(7,45) = 10.7$, $p < .001$). A summary of these overall intraclass correlation estimates can be found in Appendix C and a complete reporting of the individual behaviour estimates and values within this category can be found in Table 18.

Table 18

*Fleiss' Kappa Estimates and Percentage Agreement Values for the Behaviours in the Incongruence Category During Each Phase*

| Behaviour | *k* | *z* | *p* | Percentage Agreement | |
|---|---|---|---|---|---|
| | | | | *Tol* = 0 | *Tol* = 1 |
| | | Phase 2A: Subject 1 | | | |
| Full Shrug | -0.001 | -0.09 | 0.926 | 99 | 100.00 |
| Shrug Without Arms | 0.163 | 15.90 | < .001 | 97.6 | 100.00 |
| Asymmetric Shrug | -0.001 | -0.09 | 0.926 | 99 | 100.00 |
| Head Nod | 0.163 | 15.90 | < .001 | 97.6 | 100.00 |
| Head Shake | 0.358 | 38.80 | < .001 | 89 | 100.00 |
| Humor | 0.150 | 17.00 | < .001 | 96.7 | 99.50 |
| Hand Shrug | 0.069 | 7.76 | < .001 | 95.2 | 99.50 |
| Hand Stop | 0.163 | 15.90 | < .001 | 97.6 | 100.00 |
| Laughter | 0.260 | 27.10 | < .001 | 86.7 | 99.50 |

| Behaviour | *k* | *z* | *p* | Percentage Agreement *Tol* = 0 | Percentage Agreement *Tol* = 1 |
|---|---|---|---|---|---|
| Head Down and Away | 0.054 | 5.23 | < .001 | 98.6 | 100.00 |
| Hand Dismiss | 0.118 | 12.40 | < .001 | 97.1 | 99.5 |
| Obscene Gesture | - | - | - | 100 | 100.00 |
| Phase 2B: Subject 2 | | | | | |
| Full Shrug | - | - | - | 100 | 100 |
| Shrug Without Arms | -0.001 | -0.05 | 0.958 | 99.4 | 100 |
| Asymmetric Shrug | - | - | - | 100 | 100 |
| Head Nod | -0.001 | -0.052 | 0.958 | 99.4 | 100 |
| Head Shake | 0.259 | 22.4 | < .001 | 83.1 | 100 |
| Humor | - | - | - | 100 | 100 |
| Hand Shrug | -0.002 | -0.156 | 0.876 | 98.2 | 100 |
| Hand Stop | -0.001 | -0.052 | 0.958 | 99.4 | 100 |
| Laughter | - | - | - | 100 | 100 |
| Head Down and Away | - | - | - | 100 | 100 |
| Hand Dismiss | - | - | - | 100 | 100 |
| Obscene Gesture | - | - | - | 100 | 100 |
| Phase 2C: Subject 3 | | | | | |
| Full Shrug | 0.167 | 14.20 | < .001 | 96.9 | 99.6 |
| Shrug Without Arms | 0.117 | 9.25 | < .001 | 95.50 | 100 |
| Asymmetric Shrug | 0.253 | 20.00 | < .001 | 94.2 | 100 |
| Head Nod | 0.029 | 2.37 | 0.018 | 91 | 99.6 |
| Head Shake | 0.282 | 22.90 | < .001 | 82.5 | 99.1 |
| Humor | 0.109 | 8.64 | < .001 | 96.90 | 100 |
| Hand Shrug | 0.431 | 34.50 | < .001 | 90.1 | 99.6 |
| Hand Stop | | | | | |
| Laughter | 0.59 | 46.6 | < .001 | 93.3 | 100 |
| Head Down and Away | | | | | |
| Hand Dismiss | | | | | |
| Obscene Gesture | | | | | |

*Note:* There were 210 five-second intervals during Phase 2A, 166 during Phase 2B, and 223 during Phase 2C.

## Duration

The individual behavioural modifiers that comprise the *Duration* category are: *Self Soothe, Protective Object, Arms Crossed, One Arm Crossed, Two Arms Crossed, Legs Crossed Tight, Legs Open Stance, Leg Resting on Other, Tucked on Seat, Knees Up, Leg Other, Pause, Latency,* and *Dissociation.* There was no single behaviour with highest or lowest frequency across all three subjects within this category. The frequency at which each *Duration* behaviour was coded during each Phase, along with mean frequency, SD, median, range, skewness, and kurtosis for each can be found within Appendix D.

The distribution of ICC coefficient estimates for the *Duration* category, collapsed across subjects and behaviours, was found to center around 0.22, with a grand mean of 0.24, a median of 0.19, and a standard deviation of 0.22. Reliability estimates ranged from poor (0.00) for the *Leg Resting on Other* behaviour to substantial (0.61) for the *Arms Crossed* behaviour, with a skewness of 0.72 and a kurtosis of -0.68. The distribution of percentage agreement values with a tolerance of zero for the *Duration* category, collapsed across subjects and behaviours, was found to center around 96, with a grand mean of 95.05, a median of 97.83, and a standard deviation of 8.15. Values ranged from 70.90 for the *Pause* behaviour to 100.00 for the *Tucked on Seat, Knees Up, Leg Other, and Dissociation* behaviours, with a skewness of -2.44 and a kurtosis of 6.08. At a tolerance of 1, the percentage agreement values for the *Duration* category, collapsed across subjects and behaviours, was found to center around 96, with a grand mean of 95.94, a median of 97.83, and a standard deviation of 5.78. Values ranged from 82.93 for the *Pause* behaviour 100.00 for the *Tucked on Seat, Knees Up, Leg Other, and Dissociation* behaviours, with a skewness of -1.83 and a kurtosis of 2.36. See Table 19 for mean ICC coefficient estimates and mean percentage agreement values for each individual behaviour in this category, collapsed across each Phase, and Table 20 for values specific to each individual behaviour within each Phase.

Table 19

*Estimated Reliability and Agreement Means for Each Behaviour in the Hands Category, Collapsed Across Three Subjects*

| | | | *PA (Tol =0)* | | *PA (Tol =1)* | |
|---|---|---|---|---|---|---|
| Behaviour | *k M* | *k* SD | *M* | SD | *M* | SD |
| Self Soothe | 0.252 | 0.30 | 84.17 | 25.09 | 83.20 | 22.78 |
| Protective Object | 0.129 | 0.18 | 96.40 | 4.76 | 96.70 | 4.25 |
| Arms Crossed | 0.612 | | 97.30 | | 97.30 | |
| One Arm Crossed | 0.081 | | 95.95 | 5.73 | 95.95 | 5.73 |
| Two Arms Crossed | 0.123 | | 98.35 | 2.33 | 98.35 | 2.33 |
| Legs Crossed Tight | 0.381 | | 96.90 | | 96.90 | |
| Leg Open Stance | 0.009 | | 99.60 | | 99.60 | |
| Leg Resting on | 0 | | 98.70 | | 98.70 | |

| Behaviour | *k M* | *k* SD | *PA (Tol =0)* *M* | SD | *PA (Tol =1)* *M* | SD |
|---|---|---|---|---|---|---|
| Tucked on Seat | - | | 100.00 | | 100.00 | |
| Knees Up | - | | 100.00 | | 100.00 | |
| Leg Other | - | | 100.00 | | 100.00 | |
| Pause | 0.548 | 0.18 | 70.90 | 2.88 | 82.93 | 4.01 |
| Latency | 0.259 | 0.05 | 92.37 | 2.83 | 93.50 | 3.35 |
| Dissociation | - | | 100.00 | | 100.00 | |

The ICC inter-rater reliability coefficients for the overall *Duration* category were estimated to range from ICC(2,8) = .353 to ICC(2,8) = .724 across all three subjects, with an estimated grand mean coefficient of ICC(2,8) = .588. The inter-rater reliability coefficient estimate for the *Duration* category was found to be fair during Phase 2A, (ICC(2,8) = .353), with a 95% confidence interval ranging from 0.14 to 0.73, ($F(7,71) = 6.36$, $p < .001$). During Phase 2B, the inter-rater reliability coefficient estimate for the *Duration* category was found to be substantial, (ICC(2,8) = .686), with a 95% confidence interval ranging from 0.45 to 0.91, ($F(7,71) = 22.1$, $p < .001$). Estimates for the *Body* category during Phase 2C were also found to be substantial, (ICC(2,8) = .724), with a 95% confidence interval ranging from 0.55 to 0.88, ($F(13,91) = 20.5$, $p < .001$). A summary of these overall intraclass correlation estimates can be found in Appendix C and a complete reporting of the individual behaviour estimates and values within this category can be found in Table 20.

Table 20

*ICC Estimates and Percentage Agreement Values for the Behaviours in the Duration Category During Each Phase*

| Behaviour | *ICC* | 95% CI *LL* | *UL* | Percentage Agreement *Tol* = 0 | *Tol* = 1 |
|---|---|---|---|---|---|
| | | Phase 2A: Subject 1 | | | |
| Self Soothe | 0.587 | 0.54 | 0.64 | 55.2 | 57.10 |
| Protective Object | 0 | -0.02 | 0.023 | 91 | 91.90 |
| Arms Crossed | | | | | |
| One Arm Crossed | 0.081 | 0.05 | 0.117 | 91.9 | 91.90 |
| Two Arms Crossed | 0.123 | 0.09 | 0.164 | 96.7 | 96.70 |
| Legs Crossed Tight | | | | | |
| Leg Open Stance | | | | | |
| Leg Resting on Other | | | | | |
| Tucked on Seat | | | | | |

| Behaviour | ICC | 95% CI LL | 95% CI UL | Percentage Agreement Tol = 0 | Percentage Agreement Tol = 1 |
|---|---|---|---|---|---|
| Knees Up | | | | | |
| Leg Other | | | | | |
| Pause | 0.754 | 0.72 | 0.792 | 67.6 | 78.6 |
| Latency | 0.224 | 0.18 | 0.275 | 90 | 91 |
| Dissociation | - | - | - | 100 | 100.00 |
| Phase 2B: Subject 2 | | | | | |
| Self Soothe | 0.169 | 0.126 | 0.221 | 98.2 | 93.4 |
| Protective Object | - | - | - | 100 | 100 |
| Arms Crossed | | | | | |
| One Arm Crossed | - | - | - | 100 | 100 |
| Two Arms Crossed | - | - | - | 100 | 100 |
| Legs Crossed Tight | | | | | |
| Leg Open Stance | | | | | |
| Leg Resting on Other | | | | | |
| Tucked on Seat | | | | | |
| Knees Up | | | | | |
| Leg Other | | | | | |
| Pause | 0.474 | 0.413 | 0.539 | 72.9 | 83.7 |
| Latency | 0.318 | 0.262 | 0.382 | 91.6 | 92.2 |
| Dissociation | - | - | - | 100 | 100 |
| Phase 2C: Subject 3 | | | | | |
| Self Soothe | 0.000 | -0.02 | 0.028 | 99.1 | 99.1 |
| Protective Object | 0.257 | 0.21 | 0.312 | 98.20 | 98.2 |
| Arms Crossed | 0.612 | 0.56 | 0.663 | 97.3 | 97.3 |
| One Arm Crossed | | | | | |
| Two Arms Crossed | | | | | |
| Legs Crossed Tight | 0.381 | 0.33 | 0.44 | 96.9 | 96.9 |
| Leg Open Stance | 0.009 | -0.01 | 0.039 | 99.6 | 99.600 |
| Leg Resting on Other | 0.000 | -0.02 | 0.028 | 98.700 | 98.700 |
| Tucked on Seat | - | - | - | 100 | 100 |
| Knees Up | - | - | - | 100 | 100 |
| Leg Other | - | - | - | 100 | 100 |
| Pause | 0.416 | 0.361 | 0.475 | 72.2 | 86.5 |
| Latency | 0.234 | 0.187 | 0.288 | 95.5 | 97.3 |
| Dissociation | - | - | - | 100 | 100 |

*Note:* There were 210 five-second intervals during Phase 2A, 166 during Phase 2B, and 223 during Phase 2C.

**Affect I**

The individual behaviours that comprise the *Affect I* category differed across Subjects. In Phase 2A & B, the behaviours that comprised the *Affect I* category were *Micro Expressions* (coded nominally), *Masking*, and *Neutralizing.* However, in Phase 2C, the behaviours that comprise the *Affect I* category are: *Micro Expression* (*ME) Happiness*, *ME Sadness*, *ME Anger*, *ME Fear*, *ME Surprise*, *ME Disgust*, *ME Contempt*, *Micro Expression Masking (MEM) Happiness*, *MEM Sadness*, *MEM Anger*, *MEM Fear*, *MEM Surprise*, *MEM Disgust*, *MEM Contempt* and *MEM Neutral.* Across all three subjects, *Neutralizing* was found to be the most common Micro Expression Mask. The frequency at which each *Affect I* behaviour was coded during each Phase, along with mean frequency, SD, median, range, skewness, and kurtosis for each can be found within Appendix D.

The distribution of Fleiss' Kappa coefficient estimates for the *Affect I* category, collapsed across subjects and behaviours, was found to center around 0.02, with a grand mean of 0.03, a median of 0.01, and a standard deviation of 0.04. Reliability estimates ranged from poor (-0.01) to slight (0.12), with a skewness of 1.11 and a kurtosis of -0.09. The distribution of percentage agreement values with a tolerance of zero for the *Affect I* category, collapsed across subjects and behaviours, was found to center around 90, with a grand mean of 87.49, a median of 93.05, and a standard deviation of 17.39. Values ranged from 42.65 for the *Micro Expression* behaviour to 100.00 for the *MEM Surprise* behaviour, with a skewness of -1.85 and a kurtosis of 2.42. At a tolerance of 1, the percentage agreement values for the *Affect I* category, collapsed across subjects and behaviours, was found to center around 98, with a grand mean of 97.23, a median of 100, and a standard deviation of 9.18. Values ranged from 61.70 to 100.00, with a skewness of -3.85 and a kurtosis of 15.28. See Table 21 for mean Fleiss' Kappa coefficient estimates and mean percentage agreement values for each individual behaviour in this category, collapsed across each Phase, and Table 22 for values specific to each individual behaviour within each Phase.

Table 21

*Estimated Reliability and Agreement Means for Each Behaviour in the Affect I Category, Collapsed Across Three Subjects*

| | | | *PA (Tol =0)* | | *PA (Tol =1)* | |
|---|---|---|---|---|---|---|
| Behaviour | *k M* | *k* SD | *M* | SD | *M* | SD |
| Micro Expression | 0.046 | 0.01 | 42.65 | 1.06 | 61.70 | 2.97 |

| Behaviour | *k M* | *k* SD | *PA (Tol =0)* | | *PA (Tol =1)* | |
|---|---|---|---|---|---|---|
| | | | *M* | SD | *M* | SD |
| Micro Expression, Masking | 0.006 | 0.01 | 80.00 | 2.69 | 89.90 | 0.14 |
| Micro Expression, | 0.035 | 0.04 | 49.30 | 8.34 | 99.75 | 0.35 |
| Micro Expression | | | | | | |
| Happiness | 0 | - | 90.10 | - | 100.00 | - |
| Sadness | 0.024 | - | 93.30 | - | 100.00 | - |
| Anger | 0.093 | - | 92.80 | - | 99.60 | - |
| Fear | - | - | 95.10 | - | 100.00 | - |
| Surprise | -0.01 | - | 91.90 | - | 100.00 | - |
| Contempt | 0.119 | - | 92.40 | - | 100.00 | - |
| Disgust | 0.083 | - | 92.40 | - | 100.00 | - |
| Micro Expression Mask | | | | | | |
| Happiness | 0.011 | - | 93.70 | - | 99.60 | - |
| Sadness | - | - | 98.70 | - | 100.00 | - |
| Anger | - | - | 99.60 | - | 100.00 | - |
| Fear | - | - | 99.10 | - | 100.00 | - |
| Surprise | - | - | 100.00 | - | 100.00 | - |
| Contempt | - | - | 99.60 | - | 100.00 | - |
| Disgust | - | - | 99.60 | - | 100.00 | - |
| Neutral | 0.079 | - | 64.60 | - | 99.60 | - |

The ICC inter-rater reliability coefficient for *Affect I* was only calculated for Phase 2C, due to the limited number of *Affect* categories in Phase 2A&B. Estimates for the *Affect I* category during Phase 2C were found to be moderate, (ICC(2,8) = .440), with a 95% confidence interval ranging from 0.25 to 0.69, ($F(14,77) = 8.79$, $p < .001$). A summary of these overall intraclass correlation estimates can be found in Appendix C and a complete reporting of the individual behaviour estimates and values within this category can be found in Table 22.

Table 22

*Fleiss' Kappa Estimates and Percentage Agreement Values for the Behaviours in the Hands Category During Each Phase*

| Behaviour | *k* | *z* | *P* | Percentage Agreement | |
|---|---|---|---|---|---|
| | | | | *Tol = 0* | *Tol = 1* |
| Phase 2A: Subject 1 | | | | | |
| Micro Expression | 0.054 | 8.07 | < .001 | 41.9 | 63.80 |
| Micro Expression, Masking | 0.0105 | 1.44 | 0.151 | 78.1 | 90.00 |
| Micro Expression, Neutralizing | 0.064 | 6.54 | < .001 | 55.2 | 99.50 |
| Phase 2B: Subject 2 | | | | | |

| Behaviour | *k* | *z* | *P* | Percentage Agreement *Tol = 0* | Percentage Agreement *Tol = 1* |
|---|---|---|---|---|---|
| Micro Expression | 0.038 | 5.05 | < .001 | 43.4 | 59.6 |
| Micro Expression, Masking | 0.001 | 0.13 | 0.896 | 81.9 | 89.8 |
| Micro Expression, Neutralizing | 0.005 | 0.44 | 0.662 | 43.4 | 100 |
| Phase 2C: Subject 3 | | | | | |
| Micro Expression | | | | | |
| Happiness | 0.000 | -0.04 | 0.97 | 90.1 | 100 |
| Sadness | 0.024 | 1.92 | 0.055 | 93.3 | 100 |
| Anger | 0.093 | 7.62 | < .001 | 92.8 | 99.6 |
| Fear | -0.006 | -0.49 | 0.624 | 95.1 | 100 |
| Surprise | -0.010 | -0.81 | 0.421 | 91.9 | 100 |
| Contempt | 0.119 | 9.41 | < .001 | 92.4 | 100 |
| Disgust | 0.083 | -0.04 | 0.965 | 92.4 | 100 |
| Micro Expression Mask | | | | | |
| Happiness | 0.011 | 0.93 | 0.35 | 93.7 | 99.6 |
| Sadness | -0.002 | -0.13 | 0.894 | 98.7 | 100 |
| Anger | -0.001 | -0.04 | 0.965 | 99.6 | 100 |
| Fear | -0.001 | -0.09 | 0.929 | 99.1 | 100 |
| Surprise | - | - | - | 100 | 100 |
| Contempt | -0.001 | -0.04 | 0.965 | 99.6 | 100 |
| Disgust | -0.001 | -0.04 | 0.965 | 99.6 | 100 |
| Neutral | 0.079 | 6.31 | < .001 | 64.6 | 99.6 |

*Note:* There were 210 five-second intervals during Phase 2A, 166 during Phase 2B, and 223 during Phase 2C.

**Affect II**

The individual behaviours that comprise the *Affect II* category also differed across subjects. In Phase 2A & B, the behaviours that comprised the *Affect II* category were *Incongruent Expression and Incongruent Expression Expected* (coded nominally). However, in Phase 2C, the behaviours that comprise the *Affect II* category are: *Incongruent Expression* (*IE) Happiness, IE Sadness, IE Anger, IE Fear, IE Surprise, IE Disgust, IE Contempt, IE Neutral, Incongruent Expression Expected (IEE) Happiness, IEE Sadness, IEE Anger, IEE Fear, IEE Surprise, IEE Disgust, IEE Contempt* and *IEE Neutral.* In Phase 2C, *Happiness* was found to be the most frequent Incongruent Expression. The frequency at which each *Affect II* behaviour was coded during each Phase, along with mean frequency, SD, median, range, skewness, and kurtosis for each can be found within Appendix D.

The distribution of Fleiss' Kappa coefficient estimates for the *Affect II* category, collapsed across subjects and behaviours, was found to center around 0.07, with a grand mean of 0.0, a median of 0.07, and a standard deviation of 0.07. Reliability estimates ranged from poor (0.00) to fair (0.23), with a skewness of 0.72 and a kurtosis of -0.44. The distribution of percentage agreement values with a tolerance of zero for the *Affect II* category, collapsed across subjects and behaviours, was found to center around 97, with a grand mean of 96.79, a median of 98.90 and a standard deviation of 5.05. Values ranged from 83.65for the *Incongruent Expression* and *Incongruent Expression Expected* behaviours to 100.00 for the *IE Neutral*, *IEE Surprise,* and *IEE Disgust* behaviours, with a skewness of -2.25 and a kurtosis of 4.17. At a tolerance of 1, the percentage agreement values for the *Affect II* category, collapsed across subjects and behaviours, was found to center around 99, with a grand mean of 98.80, a median of 100, and a standard deviation of 3.50. Values ranged from 88.15 to 100.00, with a skewness of -2.76 and a kurtosis of 6.47. See Table 23 for mean Fleiss' Kappa coefficient estimates and mean percentage agreement values for each individual behaviour in this category, collapsed across each Phase, and Table 24 for values specific to each individual behaviour within each Phase.

Table 23

*Estimated Reliability and Agreement Means for Each Behaviour in the Hands Category, Collapsed Across Three Subjects*

| | | | *PA (Tol =0)* | | *PA (Tol =1)* | |
|---|---|---|---|---|---|---|
| Behaviour | *k M* | *k* SD | *M* | SD | *M* | SD |
| Incongruent Expression | 0.077 | 0.03 | 83.65 | 7.85 | 88.15 | 10.82 |
| Expected Expression | 0.072 | 0.01 | 83.65 | 7.85 | 90.30 | 1.84 |
| Incongruent Expression | | | | | | |
| Happiness | 0.233 | - | 93.70 | - | 100.00 | - |
| Sadness | -0.001 | - | 99.10 | - | 100.00 | - |
| Anger | -0.001 | - | 99.10 | - | 100.00 | - |
| Fear | -0.001 | - | 99.60 | - | 100.00 | - |
| Surprise | 0.094 | - | 99.10 | - | 100.00 | - |
| Contempt | -0.003 | - | 97.80 | - | 100.00 | - |
| Disgust | 0.094 | - | 99.10 | - | 100.00 | - |
| Neutral | - | - | 100.00 | - | 100.00 | - |
| Expected Expression | | | | | | |
| Happiness | 0.188 | - | 98.70 | - | 100.00 | - |
| Sadness | 0.123 | - | 97.30 | - | 100.00 | - |
| Anger | 0.055 | - | 98.20 | - | 100.00 | - |

| Behaviour | *k M* | *k* SD | *PA (Tol =0)* | | *PA (Tol =1)* | |
|---|---|---|---|---|---|---|
| | | | *M* | SD | *M* | SD |
| Fear | -0.001 | - | 99.60 | - | 100.00 | - |
| Surprise | - | - | 100.00 | - | 100.00 | - |
| Contempt | -0.002 | - | 98.20 | - | 100.00 | - |
| Disgust | - | - | 100.00 | - | 100.00 | - |
| Neutral | 0.164 | - | 95.50 | - | 100.00 | - |

The ICC inter-rater reliability coefficient for *Affect II* was only calculated for Phase 2C, due to the limited number of *Affect* categories in Phase 2A&B. Estimates for the *Affect II* category during Phase 2C were found to be fair, (ICC(2,8) = .301), with a 95% confidence interval ranging from 0.14 to 0.55, ($F$(15,104) = 4.86, $p$ < .001). A summary of these overall intraclass correlation estimates can be found in Appendix C and a complete reporting of the individual behaviour estimates and values within this category can be found in Table 24.

Table 24

*Fleiss' Kappa Estimates and Percentage Agreement Values for the Behaviours in the Affect II Category During Each Phase*

| Behaviour | *k* | *z* | *P* | Percentage Agreement | |
|---|---|---|---|---|---|
| | | | | *Tol = 0* | *Tol = 1* |
| Phase 2A: Subject 1 | | | | | |
| Incongruent Expression | 0.0597 | 9.56 | < .001 | 78.1 | 80.50 |
| Expected Expression | 0.0621 | 8.38 | < .001 | 78.1 | 89.00 |
| Phase 2B: Subject 2 | | | | | |
| Incongruent Expression | 0.095 | 9.9 | < .001 | 89.2 | 95.8 |
| Expected Expression | 0.082 | 11.00 | < .001 | 89.2 | 91.6 |
| Phase 2C: Subject 3 | | | | | |
| Incongruent Expression | | | | | |
| Happiness | 0.233 | 18.40 | < .001 | 93.7 | 100 |
| Sadness | -0.001 | -0.09 | 0.929 | 99.1 | 100 |
| Anger | -0.001 | -0.089 | 0.929 | 99.1 | 100 |
| Fear | -0.001 | -0.04 | 0.965 | 99.6 | 100 |
| Surprise | 0.094 | 7.41 | < .001 | 99.1 | 100 |
| Contempt | -0.003 | -0.22 | 0.824 | 97.8 | 100 |
| Disgust | 0.094 | 7.41 | < .001 | 99.1 | 100 |
| Neutral | - | - | - | 100 | 100 |
| Expected Expression | | | | | |
| Happiness | 0.188 | 14.80 | < .001 | 98.7 | 100 |
| Sadness | 0.123 | 9.68 | < .001 | 97.3 | 100 |

| Behaviour | *k* | *z* | *P* | Percentage Agreement | |
|---|---|---|---|---|---|
| | | | | *Tol = 0* | *Tol = 1* |
| Anger | 0.055 | 4.31 | < .001 | 98.2 | 100 |
| Fear | -0.001 | -0.04 | 0.965 | 99.6 | 100 |
| Surprise | - | - | - | 100 | 100 |
| Contempt | -0.002 | -0.18 | 0.859 | 98.2 | 100 |
| Disgust | - | - | - | 100 | 100 |
| Neutral | 0.164 | 13.00 | < .001 | 95.5 | 100 |

*Note:* There were 210 five-second intervals during Phase 2A, 166 during Phase 2B, and 223 during Phase 2C.

## Discussion

The overall objective of the current study was to develop a coding manual, train a group of coders, and work toward achieving inter-rater reliability for the factors outlined in the manual. Given the breadth of literature with regard to nonverbal behaviour, the opportunity for consultation within the academic community, and the intuitive capabilities of the coding software, the initial manual development process was successful. At this current stage of the research program, acceptable levels of inter-rater reliability have been achieved for various categories, while others require further remediation. Based on the results presented above, further manual development, coder training, and feedback will continue until inter-rater reliability has been achieved for all categories outlined above. This process is necessary for the continuation of this research program.

The following section presents an overview of, not only the levels of reliability and agreement, but of additional factors which effect the reliability of the current coding tool. Factors such as subject and coder individual differences, common behavioural confounds, statistical considerations, and behaviour descriptions are important aspects to consider when interpreting the results of the current study.

### Hands

While strong levels of agreement were found across the majority of behaviours in this category, exceptions were found for *Illustrators, Hands Together,* and *Self Soothe* (for Subject 1 only). However, acceptable levels of reliability were consistently observed for *Illustrators, Covering Mouth, Wiping Mouth, Hand to Eye, Hand to Nose, Hand Below Knee, Hand Clench,* and *Hands Together,* while sequential improvements were noted for the *Hands Symmetrical, Hands Asymmetrical, Chin Resting in Hand, Hand to Ear, Hand Shrug,* and *Hand Stop* behaviours. All coders agreed that the *Obscene Gesture* behaviour did not occur at all across all three subjects. Individual behaviours requiring further inspection and remediation include *Hand to Forehead, Hand Touching Chin, Hand Above Knee, Hand Scratch/Pick, Self Soothe, Hand Hiding,* and *Hand Dismiss.*

Many of the *Hand* behaviours were found to have consistent patterns across coders and subjects; most notably *Hands Symmetrical, Covering Mouth, Wiping Mouth, Hand to Eye, Hand to Ear,* and *Hand to Nose. Hand Shrug* was also found to be consistent and demonstrated improvements during the third video. While the *Hand Clench* and *Hand Below Knee* behaviours were coded with slightly less consistency, they proved to be robust against across-subject variability.

Behaviours that revealed inconsistent coding patterns consisted of *Hands Asymmetrical, Hand to Forehead, Hand Resting on Chin, Hand Above Knee, Hand Hiding, Self Soothe,* and *Illustrators.* Additionally, only two coders were able to effectively code *Hand Dismiss* behaviours without having it confounded with *Illustrators*. The *Hand Touching Chin and Shielding Eyes* behaviours were not found to have any true instances of the behaviour across the three videos, yet two coders consistently coded false positives in these categories. Finally, behaviours that varied substantially between subjects consisted of *Hand to Forehead, Hand to Ear, Hand Below Knee, Hands Symmetrical, Hand Clench,* and *Illustrators.*

Commonly confounded behaviours within the category of *Hands* were found to be *Hands Asymmetrical* and *Hands Symmetrical*, *Fidget* and *Self Soothe*, and *Shielding Eyes* and *Hand to Eye.* Many of the hand behaviours, such as *Hand Clench, Hand Shrug,* and *Hand Stop* were commonly mistaken for *Illustrators*. Additionally, the *Hand Scratch/Pick* behaviour was commonly confounded with many of the behaviours in which the hand touches a specific part of the face.

Overall, reliability appeared to improve over time for the behavioural category of *Hands*. While substantial reliability was found for the hand movements of Subjects 1 and 2, reliability was found to be almost perfect for the third subject. This overall category appeared to be coded consistently with some degree of subject variability. Specifically, Subject 2’s hand movements were less frequent and more difficult to code when they did appear.

**Body**

Strong levels of agreement were observed across each behaviour in the body category with no exceptions. However, the individual point-by-point reliability analyses did not return any individual behaviours with acceptable levels of reliability across all three subjects. Importantly, the newly developed Phase 2C *Fidgeting* behaviour returned a moderate level of reliability in addition to being the strongest reliability estimate across all behaviours and subjects within this category. While sequential improvements were observed for the *Protective Object* category, the *Lean Forward* and *Lean Back* behaviours were observed to decline over time. Further, the behaviours *Leaning Side, Rotated Away*, and *Body Fidget* require further investigation and remediation.

One *Body* behaviour, *Fidgeting* was found to have a consistent pattern of coding in the third video across all coders. However, there were additional behaviours that revealed inconsistent coding patterns across all three videos. There were occasional false positives found for the *Leaning Side* behaviour across all three videos, and similarly very few true instances of *Rotated Away* behaviours alongside numerous false positives for each. Behaviours that varied consistently between subjects included *Leaning Forward, Leaning Back,* and *Protective Object.*

Commonly confounded behaviours within this category were found to be confusing any *Leaning* behaviour with *Fidgeting* and *Illustrators,* and *Head Side Tilt* was confounded with *Leaning Side.* Additionally, *Leaning Back* was frequently confounded with *Illustrators, Fidgeting,* and *Head Tilt Up,* and *Leaning Forward* was frequently confounded with *Illustrators* for two coders in particular.

Overall, reliability appeared to improve substantially over the three subjects. While only slight reliability was observed for Subjects 1 and 2, reliability was found to be almost perfect for the third subject. There was some degree of overall subject variability, with Subject 3 being significantly easier to code for behaviours in this category.

**Limbs**

Strong levels of agreement were found across all behaviours in this category and the *Arms Crossed, Legs Crossed Tight,* and *Legs Open Stance* behaviours new to Phase 2C returned acceptable reliability estimates for Subject 3. Additionally, the *Arms Akimbo, Head* returned acceptable reliability over the first two subjects. However, the *Leg Resting on Other* and *Leg Kicking* behaviour require further remediation and inspection. All coders agreed that the *Arms*

*Akimbo, Hips* behaviour did not occur across all three videos, and that *Legs Other, Knees Up,* and *Tucked on Seat* did not occur during the last video.

Within the category of *Limbs*, it was difficult to determine consistent levels of coding due to the fact that many of the behaviours occurred so infrequently, and because many of the behaviours within this category were added in Phase 2C. Overall, the *Arms Akimbo, Head* behaviour appeared to have a consistent pattern of coding across all coders. The *Arms Crossed* behaviour was consistently coded within the third video, with one instance of early coding and one instance of late coding and the *Legs Crossed Tight* and *Legs Open Stance* behaviour were coded with high levels of consistency within the third video. Patterns of in consistent coding were found for the *Leg Resting on Other* and *Leg Kicking* behaviours, however, there is not enough information to determine systematic variability between subjects. Commonly confounded behaviours included *Leg Kicking* and *Fidgeting,* and *Leg Resting on Other* and *Legs Crossed Tight.*

Overall, reliability appeared to improve across the first two subjects but declined for the third subject. However, after further investigation of this anomaly, it was discovered that this decline during the third video was a result of a specific coder's "over coding" of the *Leg Kicking* behaviour. This coding was highly impactful on the overall reliability estimate for the *Limbs* category due to the small number of behaviours coded within this category and the overall infrequency of the behaviours observed. When removing this particular coder from the analyses, the reliability estimate for the overall *Limbs* category was found to be substantial, thus continuing the trend of improvement over time. Overall, this category was consistently coded across all thee subjects.

**Head and Shoulders**

While strong levels of agreement were found across the majority of behaviours in this category, exceptions were found for the *Head Side Tilt, Head Nod,* and *Head Shake* behaviours. However, acceptable levels of reliability were consistently observed for *Head Side Tilt, Head Nod,* and *Head Shake* and sequential improvements were noted for the *Full Shrug, Asymmetric Shrug,* and *Head Averted* behaviours. A sequential decline was noted for the *Head Tilt Up* behaviour and further inspection and remediation is required for the *Shrug Without Arms, Head Tilt Down,* and *Head Down and Away* behaviours.

Consistent patterns of coding were observed for the behaviours of *Head Side Tilt, Head Nod*, and *Head Shake.* Notably, *Head Side Tilt* and *Head Nod* were found to be robust against between subject variations. Lower levels of coder consistency were for behaviours such as *Full Shrug* wherein there were very few true instances of the behaviour but frequent false positive coding, *Head Tilt Up, Head Tilt Down, Shrug Without Arms,* and *Head Averted.* Systematic differences between subjects that were found to decrease reliability were found for the *Head Tilt Up, Head Side Tilt, Asymmetric Shrug, Head Nod, Head Shake, and Head Averted.* Commonly confounded behaviours included *Head Averted* and *Head Tilt Down and Away, Head Tilt Down* and *Head Nod, Scratching,* or *Illustrators,* and *Head Tilt Down and Away* and functional movements (e.g., looking at the time). Additionally, all shrugging behaviours were commonly confounded with each other.

Overall reliability appeared to improve from the first subject to the second subject, but declined slightly for Subject 3, though still within acceptable range. While substantial and almost perfect reliability was found for the *Head and Shoulders* category for Subjects 1 and 2, respectively, moderate reliability was found for this category for Subject 3. There was some degree of between-subject variability that decreased the reliability estimate for the third video.

**Face and Eyes**

Levels of agreement were variable across the majority of behaviours in this category, through strong levels of agreement were found for the *Lip Bite, Pursed Lips, Biting Self, Biting Other, Oral Fixation Self, Oral Fixation Other, Gaze Aversion, Eyes Looking Away, Eyes Searching, Blinking, Full Eye Closure,* and *Eye Rolling.* Acceptable levels of reliability were consistently observed for the *Lip Lick, Blink,* and *Full Eye Closure* behaviours, while a sequential improvement was observed for the *Lip Bite* behaviour, and a sequential decline was observed for the *Pursed Lips* behaviour. All coders agreed that the *Biting Other* and *Oral Fixation Object* behaviours did not occur across all three subjects. Individual behaviours requiring further inspection and remediation include *Biting Self, Oral Fixation Self, Gaze Aversion,* and *Eye Rolling.*

The behaviours that demonstrated the highest consistency with regard to coding patterns were *Lip Lick, Biting Self,* and *Blinking.* Patterns with a lower degree of consistent coding patterns included *Lip Bite, Pursed Lips,* and *Eye Rolling.* Though the *Oral Fixation Self* behaviour was coded quite frequently by only one coder on the second video, this coder was found to be correct in her interpretation of this behaviour. *Eyes Looking Away*, and *Eyes Searching* also demonstrated inconsistent coding patterns and were subsumed into the *Gaze Aversion* behaviour. However, *Gaze Aversion* was also found to be coded relatively inconsistently. Finally, though the *Full Eye Closure* behaviour demonstrated lower levels of coder consistency across all three videos, the high frequency of observation was robust against this inconsistency.

Systematic differences between subjects were found for the *Full Eye Closure*, *Oral Fixation Self*, *Pursed Lips*, and *Lip Bite* behaviours. Common confounds were discovered between *Lip Bite* and *Lip Lick*, *Pursed Lips* and a sad affective display, *Biting Self* and *Lip Bite*, *Oral Fixation Self* and *Covering Mouth*, *Eyes Looking Away* and *Eyes Searching*, *Eye Rolling* and *Eyes Looking away*, and *Blink* and *Full Eye Closure*.

Overall reliability remained consistently strong throughout all three subjects for the *Face and Eyes* behavioural category and was found to be within the almost perfect range. Overall, this category appeared to demonstrate consistent patterns of coding and very little between-subject variability.

**Reflexive**

Strong levels of agreement were found across all behaviours in the *Reflexive* category. Levels of acceptable reliability were consistently observed for the *Deep Inhale* behaviour while sequential improvements were observed for the *Hard Swallow* behaviours. A sequential decline was found within the *Clearing Throat* behaviour. Individual behaviours requiring further remediation and investigation include the *Yawn* and *Sniff* behaviours. All coders agreed that the *Cough* behaviour did not occur across all three subjects.

Although consistent patterns of coding were found for the *Deep Inhale* and *Hard Swallow* behaviours, inconsistent patterns were found for the *Sign*, *Yawn* and *Sniff* behaviours. Additionally, systematic between subject variability was observed for the *Sniff*, *Sigh*, *Clearing Throat*, and *Hard Swallow* behaviours. Common confounds were observed to be *Deep Inhale* and *Sigh*, and *Yawn* and *Deep Inhale*.

Overall reliability appeared to improve and plateau over time for the behavioural *Reflexive* category. While fair reliability was observed for the first subject, reliability improved and was found to be within the moderate range for Subjects 2 and 3. Although there was some degree of coding variability overall, it was quite minimal, and coding remained relatively stable throughout this category.

**Paralinguistic I**

While strong levels of agreement were found across the majority of the behaviours in this category, exceptions were found for the *Pause* and *Vocal Raise* behaviours. Acceptable levels of reliability were consistently observed for the *Pause, Fillers,* and *Laughter* behaviours while sequential improvements were observed for *Speech Errors* and *Interruptions*. A sequential decline was observed for the *Latency* behaviour and all coders agreed that the *Scoff* and *Dissociation* behaviour did not occur across all three subjects. The *Vocal Raise* behaviour requires further remediation and inspection.

Consistent coding patterns were observed for the *Pauses*, *Fullers*, and *Laughter* behaviours, while inconsistent coding patterns were found for the *Latency, Speech Errors, Vocal Raises,* and *Interruption* behaviours. Systematic between subject variations were found for *Latency, Fillers, Speech Errors,* and *Interruptions* and common confounds included *Speech Errors* and *Corrections*, and *Vocal Raises* and listing behaviours. *Laughter* was also observed to be confounded with *Humor* and smiling behaviours.

Overall reliability appeared to remain relatively consistent over time for the behavioural category of *Paralinguistics I*. While substantial reliability was found for the paralinguistic behaviour for Subjects 1 and 3, reliability was found to be almost perfect for the second subject. No overall systematic differences were found between subjects and consistent coding patterns were observed throughout.

**Paralinguistic II**

Strong levels of agreement were consistently found across all individual behaviours in this category. While there were no behaviours that demonstrated consistent levels of acceptable levels of reliability, sequential improvements in reliability were observed for the *Lack of Memory, Humor,* and *Pardon Other* behaviours, ultimately reaching acceptable levels of reliability. Sequential declines in reliability were observed for the *Corrections* and *Generalization* behaviours and all coders agreed that the *Scoff* behaviour did not occur across all three subjects. Individual behaviours requiring further inspection and remediation include *Verbal Diversions, Dismissiveness,* and *Extreme Descriptions.*

Consistent patterns of coding were observed for *Lack of Memory* and *Pardon Self,* while inconsistent coding patterns were observed for the *Verbal Diversions, Corrections, Generalization, Dismissiveness, Extreme Descriptions, Humor,* and *Pardon Other* behaviours. These inconsistent coding patterns were predominately owing to a difficulty with identifying these behaviours within the same 5-second interval across coders. That is to say that rather than a variability across the identification of these behaviours, it would be more accurate to say that there was a variability in the timely identification. Behaviours that were observed to be susceptible to subject variability included *Lack of Memory, Dismissiveness, Humor*, and *Extreme Descriptions.* Common confounds included *Corrections* and *Speech Errors, Generalizations* and *Extreme Descriptions, Lack of Memory* and *Verbal Diversions,* and *Laughter* and *Humor.*

Overall reliability estimates remained stable over time for the behavioural category of *Paralinguistics II.* Reliability estimates were found to be fair across all three subjects. Although some individual differences in coding patterns were observed, there was no systemic coder variability overall.

**Incongruence**

Strong levels of agreement were found across all behaviours in this category. However, acceptable levels of reliability were only consistently observed for the *Head Shake* and *Laughter* incongruent behaviours and sequential improvements were observed for the *Asymmetric Shrug* and *Hand Shrug* incongruent behaviours, ultimately reaching acceptable levels of reliability for the third subject. Individual incongruent behaviours requiring further remediation and inspection include *Full Shrug, Shrug Without Arms, Head Nod, Humor, Hand Stop, Head Down and Away,* and *Hand Dismiss.* Coding patterns and subject variability depended heavily on the parent behaviour of the incongruent modifier. However, it was observed that the majority of the incongruent behaviours were heavily dependent on both subject and coder individual differences. It was evident that two coders demonstrated some difficulty in identifying incongruent behaviour. It was also observed that incongruence also depends on the subject. The determination of incongruence is dependent on the clarity of the physical movements in addition to the clarity of the narrative being delivered. While some subjects may demonstrate clarity in both of these areas, others may demonstrate more ambiguity in one or both areas.

Overall reliability estimates appeared to improve and plateau over time for the behavioural category of *Incongruence*. While fair reliability was found for Subject 1, Subjects 2 and 3 were returned moderate reliability estimates.

**Duration**

While strong levels of agreement were found across the majority of behaviours in this category, the expectation was found for the *Self Soothe* behaviour. Acceptable levels of reliability were consistently observed for *Arms Crossed*, *Legs Crossed Tight, Pauses*, and *Latency*, while sequential improvements were observed for *Protective Objects*. Sequential decline was observed for the *Self Soothe, Duration* behaviour, and further inspection and remediation is required on the *Leg Open Stance* and *Leg Resting on Other* duration behaviours. As with *Incongruence*, the duration category comprises all the duration modifiers within the coding system. Therefore, the performance of these data depends heavily on the parent behaviours and interpretations surrounding coder consistency and subject variability are difficult to ascertain.

Overall reliability appeared to improve and plateau over time for the behavioural category of *Duration.* While fair reliability was observed for the first subject, substantial levels of reliability were observed for the second and third subjects.

**Affect I & II**

Overall reliability was found to be Moderate for the Affect category housing Micro Expressions and Fair for the Affect category housing Incongruent Expressions. The overall individual affective behaviours were found to be poor in both categories, which is unsurprising, given the complexity and level of detail required to master these domains. The exception was found to be the incongruent expression of *Happiness*, which demonstrated acceptable reliability for the third subject. Individual coder differences appeared to play a substantial role in this variability, as it was noted that certain coders displayed a tendency to “pick up on” certain affective displays to the neglect of others. These patterns were consistent across all three subjects.

**Current Code Recommendations**

Based on the results described above, the following recommendations are made with regard to which individual behaviours and overall categories appear to be the most reliable within the current coding system. At this early stage in the research program, the current system shows extreme promise, especially with respect to the overall behavioural categories.

**Individual behaviours.** Floyd (1998) cautions against the interpretation of individual behaviours. He states that these results are most useful when providing feedback to coders and are not typically a valid representation of the reliability of the overall coding tool. However, for the purposes of tool process and development, it is imperative to present these findings in order to provide context to the development process. In addition to the behaviours that were found to be consistently reliable, Table 25 also presents individual behaviours that were found to improve in reliability over time with additional training and feedback.

Table 25

*Recommended Individual Behaviours*

| Acceptable | Showed Improvement |
|---|---|
| Illustrators | Hands Symmetrical |
| Covering Mouth | Hands Asymmetrical |
| Wiping Mouth | Chin Resting in Hand |
| Hand to Eye | Hand to Ear |
| Hand to Nose | Hand Shrug |
| Hand Below Knee | Hand Stop |
| Hand Clench | Protective Object |
| Hands Together | Full Shrug |
| Arms Crossed | Asymmetric Shrug |
| Arms Akimbo, Head | Head Averted |
| Legs Crossed Tight | Lip Bite |
| Legs Open Stance | Sigh |
| Head Side Tilt | Hard Swallow |
| Head Nod | Speech Errors |
| Head Shake | Interruptions |
| Lip Lick | Lack of Memory |
| Blink | Humor |
| Full Eye Closure | Pardon Other |
| Deep Inhale | |
| Pause | |
| Fillers | |
| Laughter | |

Of particular note in Table 25 are the individual behaviours which, despite demonstrating varying degrees of inconsistency in the coding patterns and/or subject variability, appeared to be robust in the face of these inconsistencies. These noteworthy behaviours were: *Illustrators, Hand Below Knee, Hand Clench, Head Side Tilt, Head Nod, Head Shake, Full Eye Closure,* and *Fillers.* Additional behaviours yielded demonstrated improvements despite varying degrees of inconsistency in the coding patterns and/or subject variability. These behaviours were: *Hands Symmetrical, Hands Asymmetrical, Chin Resting in Hand, Hand to Ear, Protective Object, Full Shrug, Asymmetric Shrug, Head Averted, Lip Bite, Sigh, Speech Errors, Interruption, Humor,* and *Pardon Other.*

Further, more information is required for the following individual behaviours: *Shielding Eyes, Fidgeting, Arms Crossed, Legs Crossed Tight, Legs Open Stance, Leg Resting on Other, Leg Kicking, Gaze Aversion, Biting Self,* and *Yawn.* Although the within-subject coding patterns were found to be stable for these individual behaviours, additional points of between-subject comparison are needed in order to be confident in these results. Finally, the following individual behaviours were never coded: *Obscene Gesture, Biting Object, Oral Fixation Object, Cough,* and *Dissociation.* However, it is not uncommon for coding systems to retain rare, but theoretically important behaviours (Floyd et al., 1998). These behaviours will be re-evaluated during subsequent rounds of training. The individual behaviours which did not reach acceptable levels of reliability will continue to be monitored throughout subsequent rounds of training. Reliability estimates for these individual behaviours provide valuable information for coder training and feedback. However, due to the overly conservative nature of the individual analyses, reliability estimates for the behavioural groups provide a far more accurate representation of the reliability for the present coding system.

**Behavioural categories.** Estimated reliability of behavioural categories is thought to be the most representative measure of a newly developed coding system (Floyd et al., 1998). Table 26 presents a summary of the behavioural categories that maintained acceptable levels of reliability across three subjects, as well as the behavioural categories that demonstrated sequential improvement across three subjects.

Table 26

*Recommended Behavioral Categories*

| Acceptable | Showed Improvement | Requires Remediation |
|---|---|---|
| Hands | Body | Paralinguistics II |
| Head and Shoulders | Limbs* | Affect I |
| Face and Eyes | Reflexive | Affect II |
| Paralinguistic I | Incongruence | |
| Duration | | |

Of note in Table 26 are the behavioural groups of *Hands* and *Head and Shoulders.* These groups demonstrated consistent levels of acceptable reliability despite fluctuating degrees of subject variability. However, the behavioural groups of *Paralinguistics II* and both *Affect* categories did not reach acceptable levels of reliability by the third round of coding. The results for the *Affect* categories are unsurprising. As is described further in the limitation section, mastery of the Facial Action Coding System (Ekman, 1997) requires an immense amount of specialized training which was not provided to the present coding team, as time and resources were prohibitive. This contributed to the difficulty in achieving reliability for these categories. The behaviours which comprise the *Paralinguistics II* category will undergo further remediation, as there is a great degree of variability in the coding patterns. While some coders were able to identify the target behaviours with consistent accuracy, others were more variable in their coding. Overall, nine of the twelve behavioural categories show a great degree of promise.

## Additional Observations

In addition to the interpretation of estimated reliability, further consideration should be given to statistical challenges, variations in individual differences (both coder and subject), commonly confounded behaviour, and various other factors in order to better contextualize the coding process and results.

**Statistical challenges.** The primary statistical challenge encountered in the current study was that of the reliability-agreement paradox, as mentioned above. The Fleiss' Kappa statistic, in addition to being a conservative estimate of reliability, behaves irregularly when applied to low frequency observations (Cicchetti & Feinstein, 1990; Falotico & Quatto, 2015; Feinstein & Cicchetti, 1990; Gwet, 2008; Karstad et al., 2018). This restriction of range can lower IRR estimates because the true score variance is reduced even when the error variance remains the same (Hallgren, 2012). Thus, the Fleiss' Kappa statistic can severely underrepresent agreement in cases where the subject displays a limited number of behaviours. This produces very large values for expected agreement and comparatively lower values for reliability, even when agreement is quite high (Floyd et al., 1998; Viera & Garret, 2005). Indeed, this current study found that the reliability estimate generated from Fleiss' Kappa was profoundly affected by the prevalence of the target behaviour, wherein agreement was found to be quite high with significantly lower reliability estimates.

Further, reliability at the specific code (individual behaviour) level has generally been found to be an overly conservative requirement that can prolong a study such as this (Floyd et al., 1998). Instead, the aggregate behavioural categories were investigated using the ICC analysis, which provided a more accurate indicator of the reliability of the coding tool as a whole. However, due to the relatively small number of individual behaviours in each category, and, in part, due to individual differences in coding patterns, the ICC analyses occasionally produced large confidence intervals.

**Individual differences.** Substantial variability with regard to within-subject coding patterns was observed to profoundly impact reliability estimates at both the individual behaviour level, and the aggregate behavioural category level. While the aggregate categories appeared to be robust against these variations, a number of individual behaviours were affected. For example, within the aggregate category of *Incongruence*, one coder in particular demonstrated consistently lower levels of reliability as compared to the rest of the group. Indeed, when examining the raw data, this coder demonstrated substantially lower rates of identifying the true instances of incongruence as compared to her peers. At the individual behaviour level, patterns of strength and weakness emerged when examining the raw data; It appeared that certain coders demonstrated proficiency for coding certain behaviours but not others. Thus "Coder Dependent" behaviours were identified in cases where reliability estimates were consistently low across three subjects and substantial variability was found within-subject, with some coders accurately identifying the target behaviours and others displaying consistent difficulty in target identification. Future remediation will continue to target individual coder feedback, providing explicit examples of target behaviours in addition to their own differential false positives. Clarifying these behavioural nuances may help to circumvent this difficulty.

Similarly, behaviours that were deemed "Subject Dependent" exhibited inconsistent patterns of between-subjects reliability. For the most part, individual coding patterns remained constant and fluctuated as a unit, resulting from individual *subject* differences, rather than individual coder differences. Indeed, the decrease in reliability estimates as a result of ambiguity in the target's behaviour has been well documented (Floyd et al., 1998; Karstad et al., 2018), and cannot be attributed to coder error alone. However, this is not simply a question of a particular subject being "easier" or "harder" to code. Rather, as individuals exhibit their own unique idiosyncratic patterns of nonverbal behaviour, it is often the case that one subject may exhibit great clarity in hand movements while being more ambiguous with head and facial movements. Remediation for subject-dependent behaviours will include expanding the repertoire of examples in the AAI-NVB Manual to include these ambiguous behavioural examples.

While guided by the results of the analyses, the identification of potential coder- and subject-dependent behaviours resulted from a close inspection of the raw data. This process represented an invaluable step in truly understanding the underlying processes and challenges inherent in the current study. However, as this current study includes only three subjects, additional datapoints are necessary in order to increase confidence in these results. For example, it is difficult to determine, at this stage, whether increases in reliability estimates during Phase 2C result from the additional feedback and training that was provided to coders, or whether this increase was due to individual subject differences alone. Additionally, it is difficult to determine which behaviours are truly coder- or subject-dependent, as both levels of individual difference are likely to have some degree of effect on each behaviour. Finally, it is likely that there exists an interaction effect between coder and subject individual differences wherein certain coders may be able to effectively tolerate the potential ambiguity of nonverbal behaviour, while others may not.

**Common confounds and global feedback.** An additional source of decreased reliability can be attributed to commonly confounded behaviours. While some confounds were common across all videos, others were specific to individual subjects, supporting the concept of individual subject differences. The most commonly confounded behavioural category was found to be the *Hands* category. As the codes in the AAI-NVB Manual adhere to the rule of mutual exclusivity, behavioural exemptions required clarification. For example, there was noticeable trend in coding the *Hands Apart Asymmetrical* and *Symmetrical* while the subject's arms were crossed. It was clarified that, if arms are crossed, this should be coded as *Arm Crossing,* and that both hands and arms need to be apart and in a resting position in order for the behaviour to qualify as *Hands Apart Symmetrical* or *Asymmetrical.* Another common confound was between the *Fidgeting* and *Self Soothing* behaviours. The nuances between the two are sometimes difficult to ascertain, and coders were instructed to keep in mind that a self soothe behaviour displays more of a repetitive and smooth movement quality as compared to fidgeting. Further, the *Body* and *Object* fidgeting behaviours were merged together in Phase 2C in order reduce any additional confounds. The *Hand Above* and *Hand Below Knee* behaviours required the clarification that the hand or hands needed to be in a resting position, rather than just touching the knee briefly.

The third subject provided excellent opportunities to confront increased ambiguity in hand behaviours. The *Illustrators* behaviour is easily confounded with many other hand movements when the individual's general idiosyncratic movement patterns are ambiguous relative to other subjects. For example, *Illustrators* were commonly confounded with behaviours such as *Hand Shrug* and *Hand Stop*. Coders were given concrete examples to illustrate the differences, as well as additional literature on the function and presentation of *Illustrators.* As a general rule of thumb, coders were advised to ask themselves whether the behaviour is *Illustrating* the narrative or *reacting* to it. For example, if a subject's narrative was describing a situation in which they explicitly told another person to stop what they were doing, while simultaneously exhibiting the Hand Stop behaviour, this would be coded as an *Illustrator*. The AAI-NVB Manual also instructs that coders refrain from coding functional/willful movements, such as blowing their noses or reaching for their phones to check the time. Finally, the *Hand Hiding* behaviour was often coded incorrectly in cases where the camera angle and leg positions precluded a full view of the subject's lap. Coders were subsequently instructed not to code a behaviour as *Hand Hiding* in these cases, as they were unable to ascertain if the individual was truly hiding their hands.

The second most commonly confounded category was found to be within the *Paralanguage* codes. Firstly, *Corrections* and *Speech Errors* were commonly confounded with one another and additional training was provided in order to identify the difference between the two. *Corrections* tend to be more overt, with the subject explicitly indicating that they made a mistake: "Three months ago…wait…actually, it was two months ago". *Speech Errors* tend to involve a more rapid self-correction without the subject directly drawing attention to the error. Initially, *Vocal Raises* were often confounded with the subject posing a question. While this confound was corrected in Phase 2C, an additional confound arose and numerous false positives were coded in cases where the subject was listing items, qualities, or descriptions. The explicit definition of *Vocal Raises* includes an indication that the vocal raise occurs at the end of a sentence only, and that it be observed in the absence of a question. Finally, *Extreme Descriptions* and *Generalizations* were often confounded with one another as they are quite similar in nature. Whereas *Extreme Descriptions* refer to a superlative such as “Worst”, “Best”, “Most”, “Least”, *Generalizations* refer to the frequency at which something occurred, such as “Always” or “Never”.

A number of false positives were also seen in behaviours involving the positioning of the head or torso. As a general rule, when coding head tilting and torso leaning positions, a positive observation should include a hinging at the neck or hips. Additionally, many of these behaviours were commonly confounded with regulators. As mentioned above, regulators serve to control the pacing of talk-turns in a conversation. While not included in the AAI-NVB Manual as a behaviour to code, a definition and description of regulators was included to clarify any potential confounds. Overall, the identification of these confounds allowed for unique opportunities to provide additional training, feedback, and manual refinement. Importantly, confounds for behaviors were typically found to remain in the same category. That is, hand behaviours were most often confounded with other hand behaviors, head movements confounded with other head movements, and so on. This phenomenon provides support for the decision to categorize individual behaviors within these larger groups.

**Other Considerations.** Additional process considerations to the development of a novel coding tool were noted. The identification of behaviours such as *Lack of Memory, Dismissiveness, Verbal Diversions, Generalizations, Pardon Self, Pardon Other, Humor* and *Corrections* require that the subject send a verbal message to the interviewer that can typically span seconds or even minutes. Therefore, the reliability of these behaviours was substantially diminished due to the 5-second interval data entry protocol, as the specific time-point selected could easily vary from coder to coders. These behaviours would likely be better represented by a frequency-only count rather than a point-by-point analysis. These variations may have also contributed to the wide confidence interval for the *Paralanguage II* behavioural category.

### Limitations

Although the current study yielded promising results, there also exists a variety of limitations. Firstly, although each member of the coding team received preliminary training and education in facial musculature movements, facial action units, and micro and macro facial expressions, they are not certified through the Ekman Group on the entire Facial Action Coding System (Ekman, 1997) due to prohibitive time and financial factors. Secondly, the present study did not consider eye tracking, vocal behavior (tone, pace, tempo, volume), or other peripheral aspects of nonverbal communication (Burgoon et al., 2016). By no means should the AAI-NVB manual be considered a comprehensive list of nonverbal behaviors. The behaviours selected for investigation were those considered to be relevant to the current study. Further, due to the small sample size of subjects (n = 3), it was difficult to determine whether sequential improvements could be associated with increases in training and feedback, or simply random between-subject variance. Additional rounds of investigation are needed in this regard.

Although non-specific to the current study, all research involving observational coding is subject to the limitations of coder bias. These biases may be influenced by a number of factors such as individual coder differences, coding complexity, and expectation biases (Harris & Lahey, 1982). In addition to the statistical challenges outline in previous sections, the number of individual behaviors in each behavioral category did not reach the recommended sample size for ICC analyses. In this regard, this was constrained by the number of behaviors in the present taxonomy and further amalgamation of categories would have potentially rendered them less meaningful. However, this likely contributed to occasional instances of wide confidence intervals associated with the ICC analyses. Finally, it is possible that there exists a paucity of heterogeneity, as all AAI video recordings were collected at a mental health clinic located in a relatively affluent area.

**Future Directions**

As previously stated, the current study represents the first phase in a larger research program designed to address the following questions: (1) What is the relationship between the identified nonverbal behaviours and adult attachment style? (2) What is the relationship between the identified nonverbal behaviours and in participants with early relational trauma? (3) Do certain distinct nonverbal behaviours have a tendency to cluster together across individuals? (4) Is it possible to apply this coding system outside of the AAI context? If so, (5) can this coding system be applied to transdiagnostic identification of mental health difficulties? However, in order to address these subsequent questions, it is imperative that nonverbal behaviours be carefully identified and categorized and that these behaviours can be reliably identified across multiple coders. As such, we aim to immediately continue the process of manual refinement and coder training until such time as each behavioural category reaches acceptable levels of reliability. Targeted individual coder training will be implemented to address the coder-dependent behaviours and a larger breadth of examples within the AAI-NVB Manual will be created to contend with behaviours that appear to vary across subjects due to individual subject differences. Additionally, statistical simulations will be run to test other analytical approaches in order to find an approach that more accurately fits the data. Once acceptable levels of reliability have been achieved across the entirety of the coding tool, coders will be randomly assigned the remainder of the videos in the AAI catalogue and the resultant data will be employed to address the future goals above.

**Conclusion**

The present study represents the first stage of a larger research program that ultimately aims to validate a new nonverbal coding tool against the original coding of the AAI in order to examine similarities and differences between content that is delivered verbally, versus content that "leaks" nonverbally. Through the coding of video-recorded AAIs, specific nonverbal behaviours and behavioural categories provided evidence to support the assumption that these identified behaviours can be catalogued and reliably coded.

Importantly, it was noted that careful consideration should be given to the individual differences between coders and coding style, and between the subjects themselves. Not every individual subject will display behaviours in the same way, just as not every individual coder will be primed to identify those behaviours in the same way. Though some behaviours appear to be impervious to these differences, other behaviours are more susceptible. Therefore, it is important to be sensitive to these differences in order to provide thoughtful feedback and training to the coding team. Additionally, providing coder support and constructive feedback is imperative to the success of developing a novel coding tool. Finally, it is vital to appreciate the intricacies of human behaviour (both verbal and nonverbal) and to ensure that each behaviour is clearly outlined and fully understood by the coding team.

The implications of this study are threefold. Firstly, this study will help to more effectively align the AAI with the theoretical underpinnings of attachment. The SSP, which the AAI is modelled after, was behaviourally based on the biological function of infant-mother attachment and aligned with Bowlby's proposition of attachment as a behavioural system. Thus, returning to the ethological-evolutionary foundation of attachment theory is imperative, and will better align the AAI with its attachment and evolutionary roots.

Secondly, the present study fills the methodological gaps that previously existed within the three studies that attempted to investigate nonverbal behaviour within the context of the AAI (Karlsson, 2005; Lambert 2012; Roisman et al., 2004). A nonverbal baseline was established to assess nonverbal behaviour in periods of neutrality versus periods of attachment activation, inter-rater reliability was utilized to bolster the utility of this instrument, and a larger cross-section of nonverbal codes was considered.

Lastly, the present study shows promise for the potential clinical utility of improved identification of attachment difficulties. Based on the coding completed to date, there is evidence to support that the majority of these behavioural categories can be coded reliably, thus the potential for the further development of a reliable tool for nonverbal behaviour appears promising. As this research program progresses and thematic and temporal behavioural clusters are identified, we hope to discover groups of behaviours that are indicative of specific attachment difficulties, adding valuable information for identification and early intervention of attachment related difficulties.

**References**

Afifi, T. O., MacMillan, H. L., Boyle, M., Taillieu, T., Cheung, K., & Sareen, J. (2014). Child abuse and mental disorders in Canada. *CMAJ : Canadian Medical Association*

Ainsworth, M. D. S., & Bell, S. M. (1970). Attachment, exploration, and separation: Illustrated by the behavior of one-year-olds in a strange situation. *Child development*, *41*(1), 49-67.

Ainsworth, M.D.S. (1985a). Attachments across the life span. *Bulletin of the New York Academy of Medicine, 61*, 792-812.

Ainsworth, M.D.S. (1985b). Patterns of infant-mother attachments: Antecedents and effects on development. *Bulletin of the New York Academy of Medicine*, *61*, 771-791.

Ainsworth, M.D.S., & Bowlby, J. (1991). An ethological approach to personality development. *American Psychologist*, *46,* 331-341.

American Psychological Association (1985). Standards for educational and psychological testing. Washington, DC: Author.

Baranowsky, A. B., & Gentry, J. E. (2015). *Trauma practice: Tools for stabilization and recovery (3rd ed.)* Hogrefe Publishing, Boston, MA.

Becker-Stoll, F., Delius, A., & Scheitenberger, S. (2001). Adolescents' nonverbal emotional expressions during negotiation of a disagreement with their mothers: An attachment approach. *International Journal of Behavioral Development*, *25*(4), 344-353.

Behrens, K. Y., Hesse, E., & Main, M. (2007). Mothers' attachment status as determined by the Adult Attachment Interview predicts their 6-yearolds' responses to separation and reunion: A study conducted in Japan. *Developmental Psychology, 43,* 1553–1567.

Blanchard, M., & Farber, B. A. (2016). Lying in psychotherapy: Why and what clients don't tell their therapist about therapy and their relationship. *Counselling Psychology Quarterly, 29*(1), 90-112.

Bomyea, J., & Lang, A. J. (2015). The role of executive functioning in PTSD and its treatment. *Current Psychiatry Reviews*, *11*(3), 160-171. doi:http://dx.doi.org/10.2174/1573400511666150528212418

Breslau, N., & Kessler, R. C. (2001). The stressor criterion in DSM-IV posttraumatic stress disorder: An empirical investigation. *Biological Psychiatry, 50*(9), 699-704. doi:http://dx.doi.org/10.1016/S0006-3223(01)01167-2

Bretherton, I. (1992). The origins of attachment theory: John Bowlby and Mary Ainsworth. *Developmental Psychology*, *28,* 759-775

Brouwer, S., Reneman, M. F., Dijkstra, P. U., Groothoff, J. W., Schellekens, J. M. H., & Göeken, L. N. H. (2003). Test–retest reliability of the Isernhagen work systems functional capacity evaluation in patients with chronic low back pain. *Journal of Occupational Rehabilitation*, *13*(4), 207-218.

Bryington, A. A., Palmer, D. J., & Watkins, M. W. (2002). The estimation of interobserver agreement in behavioral assessment. *The Behavior Analyst Today*, *3*(3), 323

Buller, D. B., & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory, 3,* 203–242.

Burgoon, J., Guerrero, L., & Floyd, K. (2016). *Nonverbal Communication.* New York, NY: Routledge.

Carlson, J., & Kottler, J. A. (2011). *Duped: Lies and deception in psychotherapy*. Routledge.

Casey, R. J., & Fuller, L. L. (1994). Maternal regulation of children's emotions. *Journal of Nonverbal Behavior, 18,* 57–89.

Cassidy, J., & Shaver, P. R. (1999). *Handbook of attachment: Theory, research, and clinical applications* Guilford Publications, New York, NY

Cervone, D. & Pervin, L. (2013). *Personality : theory and research*. Hoboken, NJ: Wiley.

Chu, A., & DePrince, A. P. (2006). Development of dissociation: Examining the relationship between parenting, maternal trauma and child dissociation. *Journal of Trauma & Dissociation, 7*(4), 75-89. doi:http://dx.doi.org/10.1300/J229v07n04_05

Ciarrochi, J., Robb, H., & Godsell, C. (2005). Letting a little nonverbal air into the room: Insights from acceptance and commitment therapy: Part 1: Philosophical and theoretical underpinnings. *Journal of Rational-Emotive & Cognitive-Behavior Therapy, 23*(2), 79-106. doi:http://dx.doi.org/10.1007/s10942-005-0005-y

Cicchetti, D. V., & Feinstein, A. R. (1990). High agreement but low kappa: II. Resolving the paradoxes. *Journal of clinical epidemiology*, *43*(6), 551-558.

Cicchetti, D., & Cohen, D. J. (1995). Perspectives on developmental psychopathology. In D. Cicchetti, & D. J. Cohen (Eds.), *Developmental psychopathology, vol. 1: Theory and methods; developmental psychopathology,* (pp. 3-20, Chapter xx, 787 Pages) John Wiley & Sons, Oxford.

Crittenden, P. M. (2000). A dynamic-maturational approach to continuity and change in pattern of attachment. In P. M. Crittenden, & A. H. Claussen (Eds.), *The organization of attachment relationships: Maturation, culture, and context* (pp. 343-357, Chapter xii, 432 Pages) Cambridge University Press, New York, NY.

D'Agostino, T. A., & Bylund, C. L. (2014). Nonverbal accommodation in health care communication. *Health Communication, 29*(6), 563-573. doi:http://dx.doi.org/10.1080/10410236.2013.783773

Darwin, C. (1872/1998). *The expression of emotion in man and animals.* New York: Oxford University Press

DePaulo, B. M. (1992). Nonverbal behavior and self-presentation. *Psychological Bulletin, 111,* 203–243.

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.

Dozier, M., & Kobak, R. R. (1992). Psychophysiology in attachment interviews: Converging evidence for deactivating strategies. *Child development*, *63*(6), 1473-1480.

Easterbrooks, M.A., & Goldberg, W. (1990). Security of toddler-parent attachment: Relation to children's sociopersonality functioning during kindergarten. In M.T. Greenburg, D. Cicchetti, & E.M. Cummings (Eds.), *Attachment in the preschool years (*pp. 221-244). Chicago: University of Chicago 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. (1971). Universal and cultural differences in facial expressions of emotion. In J. K. Cole (Ed.), *Nebraska symposium on motivation* (pp. 207–283). Lincoln: University of Nebraska Press.

Ekman, P. (1997). Emotion families . In Rauch, I. & Carr, G. F. (Eds.), Semiotics Around the World: Synthesis in Diversity (pp. 191-193). Berlin: Mouton de Gruyter.

Ekman, P. (1997). *What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS)*. Oxford University Press, USA.

Ekman, P. (1999). Basic Emotions. In Dalgleish, T. & Power, M. J. (Eds.), Handbook of Cognition and Emotion (pp. 45-60). New York, NY: John Wiley & Sons Ltd.

Ekman, P. (2003). Sixteen Enjoyable Emotions. Emotion Researcher, 18, 6-7.

Ekman, P. (2009). *Telling lies: Clues to deceit in the marketplace, politics, and marriage* W W Norton & Co, New York: NY.

Ekman, P. (2016). What Scientists Who Study Emotion Agree About. Perspectives on Psychological Science, 11(1), 31-34.

Ekman, P., & Friesen, W. V. (1969a). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *semiotica*, *1*(1), 49-98.

Ekman, P., & Friesen, W. V. (1969b). Nonverbal leakage and clues to deception. *Psychiatry*, *32*(1), 88-106.

Ekman, P., Davidson, R. J., Ricard, M., & Wallace, B. A. (2005). Buddhist and Psychological Perspectives on Emotions and Well-Being.Current Directions in Psychological Science, 14(2), 59-63.

Falotico, R., & Quatto, P. (2015). Fleiss' kappa statistic without paradoxes. *Quality & Quantity*, *49*(2), 463-470.

Feinstein, A. R., & Cicchetti, D. V. (1990). High agreement but low kappa: I. The problems of two paradoxes. *Journal of clinical epidemiology*, *43*(6), 543-549.

Feng, G. C. (2015). Mistakes and how to avoid mistakes in using intercoder reliability indices. *Methodology*.

Floyd, F. J., Baucom, D. H., Godfrey, J., & Palmer, C. (1998). Observational methods. In N. Schooler (Ed.), *Comprehensive clinical psychology.* Oxford: Elsevier Science Limited. doi: 10.101.6/B0080-4270(73)00223-6

Foley, G. N., & Gentile, J. P. (2010). Nonverbal communication in psychotherapy. *Psychiatry (Edgmont)*, *7*(6), 38.

Fonagy P, Steele M, Steele H, Target M. (1998) *Reflective-function manual: version 5.0. For application to the Adult Attachment Interview*. Unpublished manuscript.

Ford, J. D., Racusin, R., Ellis, C. G., Daviss, W. B., Reiser, J., Fleischer, A., & Thomas, J. (2000). Child maltreatment, other trauma exposure and posttraumatic symptomatology among children with oppositional defiant and attention deficit hyperactivity disorders. *Child Maltreatment, 5*(3), 205-217.

Foroughe, M. F., & Muller, R. T. (2012). Dismissing (avoidant) attachment and trauma in dyadic parent-child psychotherapy.*Psychological Trauma: Theory, Research, Practice, and Policy, 4*(2), 229-236. doi:http://dx.doi.org/10.1037/a0023061

Gabbard, G. (2007). Unconscious enactments in psychotherapy. *Psychiatric Annals, 37(4),* 269-275.

Gamer, M., Lemon, J., & Fellows Puspendra Singh, I. (2012). irr: Various Coefficients of Interrater Reliability and Agreement. R package version 0.84

Gander, M., & Buchheim, A. (2015). Attachment classification, psychophysiology and frontal EEG asymmetry across the lifespan: a review. *Frontiers in Human Neuroscience, 9* (79). doi: 10.3389/fnhum.2015.00079

Gediman, H. K., & Lieberman, J. S. (1996). *The many faces of deceit: Omissions, lies, and disguise in psychotherapy*. Jason Aronson, Incorporated.

George, C., & Main, M. (1979). Social interactions of young abused children: Approach, avoidance, and aggression. *Child development*, 306-318.

George, C., Kaplan, N., & Main, M. (1996). Adult Attachment Interview. Unpublished Manuscript, Department of Psychology, University of California, Berkeley (third edition).

George, C., West, M., & Pettem, O. (1999). The adult attachment projective: Disorganization of adult attachment at the level of representation. In J. Solomon, & C. George (Eds.), *Attachment disorganization; attachment disorganization* (pp. 318-346, Chapter xxiii, 420 Pages) Guilford Press, New York, NY.

Green, J., & Goldwyn, R. (2002). Annotation: Attachment disorganization and psychopathology: new findings in attachment research and their potential implications for developmental psychopathology in childhood. *Journal of Child Psychology & Psychiatry & Allied Disciplines, 43*, 835-846.

Greenberg, L. S., & Angus, L. (2004). The contributions of emotion processes to narrative change in psychotherapy. *Handbook of narrative and psychotherapy: Practice, theory and research*, 331-351.

Grossmann, K.E., Grossmann, K., & Waters, E. (2005). *Attachment from infancy to adulthood: The major longitudinal studies.* New York: Guilford Press.

Gwet, K. L. (2008). Computing inter-rater reliability and its variance in the presence of high agreement. *British Journal of Mathematical and Statistical Psychology*, *61*(1), 29-48.

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: an overview and tutorial. *Tutorials in quantitative methods for psychology*, *8*(1), 23.

Harrigan, J. A. (2013). Methodology: Coding and studying nonverbal behavior. In J. A. Hall, & M. L. Knapp (Eds.), *Nonverbal communication; nonverbal communication* (pp. 35-68, Chapter xi, 882 Pages) De Gruyter Mouton, Boston, MA. doi:http://dx.doi.org.ezproxy.library.yorku.ca/10.1515/9783110238150.35

Harrigan, J., Rosenthal, R., Scherer, K. R., & Scherer, K. (Eds.). (2008). *New handbook of methods in nonverbal behavior research*. Oxford University Press.

Harris, F. C., & Lahey, B. B. (1982). Subject reactivity in direct observational assessment: A review and critical analysis. *Clinical Psychology Review*, *2*(4), 523-538.

Heppner, P., Wampold, B., Owen, J., Wang, K. & Thompson, M. (2016). *Research design in counseling*. Boston, MA: Cengage Learning.

Herman, J. L. (1997). *Trauma and recovery: The aftermath of violence – from domestic abuse to political terror* (rev ed.). New York, NY: Basic Books

Hesse, E. (1999). The adult attachment interview: Historical and current perspectives. In J. Cassidy, & P. R. Shaver (Eds.), *Handbook of attachment: Theory, research, and clinical applications*(pp. 395-433, Chapter xvii, 925 Pages) The Guilford Press, New York, NY.

Izard, C. E. (1977). *Human emotions.* New York: Plenum.

Karlsson, R. (2005). *Nonverbal behaviors and mental representations of secure and insecure attachment during the Adult Attachment Interview (AAI).* (Unpublished doctoral dissertation). University of California, Berkley, California.

Karstad, K., Rugulies, R., Skotte, J., Munch, P. K., Greiner, B. A., Burdorf, A., ... & Holtermann, A. (2018). Inter-rater reliability of direct observations of the physical and psychosocial working conditions in eldercare: An evaluation in the DOSES project. *Applied ergonomics*, *69*, 93-103.

Kasl, S. V., & Mahl, G. F. (1965). Relationship of disturbances and hesitations in spontaneous speech to anxiety. *Journal of personality and social psychology, 1*(5), 425.

Keltner, D. & Ekman, P. (2000). Facial Expression of Emotion. In Lewis, M. & Haviland-Jones, J. (Eds.), Handbook of emotions, 2nd edition(pp. 236-249). New York: Guilford Publications, Inc.

Keltner, D. & Ekman, P. (2003). Introduction: Expression of Emotion. In Davidson, R. J., Scherer, K. R., & Goldsmith H. H. (Eds.), Handbook of Affective Sciences (pp. 411-414). New York: Oxford University Press.

Kessler, R.C., Sonnega, A., Bromet, E., Hughes, M., & Nelson, C.B. (1995). Posttraumatic stress disorder in the National Comorbidity Survey. *Archives of General Psychiatry, 52,* 1048-1060.

Kobak, R. R., Cole, H. E., Ferenz-Gillies, R., Fleming, W. S., & Gamble, W. (1993). Attachment and emotion regulation during mother-teen problem solving: A control theory analysis. *Child development*, *64*(1), 231-245.

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of chiropractic medicine*, *15*(2), 155-163.

Kottner, J., Audige, L., Brorson, S., Donner, A., Gajewski, B. J., Hróbjartsson, A., . . . Streiner, D. L. (2011). Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. *International Jourfloydnal of Nursing Studies, 48*(6), 661-671. doi:http://dx.doi.org.ezproxy.

Laible, D., & Thompson, R. A. (2007). *Early socialization: A relationship perspective.* Guilford Press, New York, NY.

Lambert, H. (2012). *Observation of nonverbal expressions as a sign of congruence during the Adult Attachment Interview.* (Unpublished Thesis). Texas Tech University, Texas.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *biometrics*, 159-174.

Lazarus, R. S. (1991). *Emotion and adaptation.* New York: Oxford University Press.

Le Poire, B. A., Shepard, C., & Duggan, A. (1999). Nonverbal involvement, expressiveness, and pleasantness as predicted by parental and partner attachment style. *Communications Monographs*, *66*(4), 293-311.

Liao, S. C., Hunt, E. A., & Chen, W. (2010). Comparison between inter-rater reliability and inter-rater agreement in performance assessment. *Annals Academy of Medicine Singapore*, *39*(8), 613.

Lyons-Ruth, K., Dutra, L., Schuder, M. R., & Bianchi, I. (2006). From infant attachment disorganization to adult dissociation: Relational adaptations or traumatic experiences? *Psychiatric Clinics of North America, 29*(1), 63-86. doi:http://dx.doi.org/10.1016/j.psc.2005.10.011

Main, M. (1983). Exploration, play, and cognitive functioning related to infant-mother attachment. *Infant Behavior & Development, 6*, 167-174

Main, M., & Goldwyn, R. (1984a). Predicting rejection of her infant from mother's representation of her own experience: implications for the abused-abusing intergenerational cycle. *Child Abuse and Neglect*, 8, 203-217.

Main, M., & Goldwyn, R. (1984b). *Adult attachment scoring and classification system*. Unpublished manuscript, University of California, Berkeley.

Main, M., & Solomon, J. (1990). Procedures for identifying infants as disorganized/disoriented during the ainsworth strange situation. *Attachment in the preschool years: Theory, research, and intervention.* (pp. 121-160) University of Chicago Press, Chicago, IL.

Mann, J., ten-Have, T., Plunkett, J. W., & Meisels, S. J. (1991). Time sampling: methodological critique. *Child Development*, 62, 227-241.

McFarlane, A. C. (2010). The long-term costs of traumatic stress: intertwined physical and psychological consequences. *World Psychiatry*, *9*(1), 3–10.

Menz, H. B., Fotoohabadi, M. R., Wee, E., & Spink, M. J. (2012). Visual categorisation of the arch index: a simplified measure of foot posture in older people. *Journal of foot and ankle research*, *5*(1), 10.

Mikulincer, M., & Shaver, P. R. (2003). The attachment behavioral system in adulthood: Activation, psychodynamics, and interpersonal processes. *Advances in experimental social psychology*, *35*, 56-152.

Morrisset, C.E., Barnard, K.E., Greenberg, M.T., Booth, C.L., & Spiecker, S.J. (1990). Environmental influences on early language development: The context of social risk. *Developmental Psychopathology, 2,* 127-149

*of Psychiatry*, *148*(1), 10-20.

Olivier F. & Marco G. (2016) BORIS: A free, versatile open-source event-logging software for video/audio coding and live observations. Methods in Ecology and Evolution, 7(11), 1324-1330, doi: 10.1111/2041-210X.12584

Paniagua, F. A. (2004). Utility of verbal-nonverbal correspondence-training techniques in outpatient pediatric settings. *Psychological Reports*, *94*(1), 317-326. doi:http://dx.doi.org/10.2466/PR0.94.1.317-326

Patterson, M. L. (2003). Commentary: Evolution and nonverbal behavior: Functions and mediating processes. *Journal of Nonverbal Behavior*, *27*(3), 201-207. doi:http://dx.doi.org/10.1023/A:1025346132037

Perry, B. D., Pollard, R. A., Blakley, T. L., Baker, W. L., & Vigilante, D. (1995). Childhood trauma, the neurobiology of adaptation, and "use-dependent" development of the brain: How "states" become "traits." *Infant Mental Health Journal, 16*, 271-291

Philippot, P., Feldman, R., & Coats, E. (2003). *Nonverbal behavior in clinical settings*. Oxford University Press, New York, NY. doi:http://dx.doi.org/10.1093/med:psych/9780195141092.001.0001

Purnell, C. (2010). Childhood trauma and adult attachment. *Healthcare Counselling and Psychotherapy Journal*, 10, 1–7.

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org.

Ramseyer, F., & Tschacher, W. (2011). Nonverbal synchrony in psychotherapy: Coordinated body movement reflects relationship quality and outcome. *Journal of Consulting and Clinical Psychology, 79*(3), 284-295. doi:http://dx.doi.org/10.1037/a0023419

Raskin, D. C., & Esplin, P. W. (1991). Statement validity assessment: Interview procedures and content analysis of children's statements of sexual abuse. *Behavioral Assessment*.

Reneman, M. F., Brouwer, S., Meinema, A., Dijkstra, P. U., Geertzen, J. H. B., & Groothoff, J. W. (2004). Test–retest reliability of the Isernhagen work systems functional capacity evaluation in healthy adults. *Journal of Occupational Rehabilitation*, *14*(4), 295-305

Roisman, G. I., Tsai, J. L., & Chiang, K. S. (2004). The emotional integration of childhood experience: Physiological, facial expressive, and self-reported emotional response during the adult attachment interview. *Developmental Psychology, 40*(5), 776-789. doi:http://dx.doi.org/10.1037/0012-1649.40.5.776

Sameroff, A. (2010). A unified theory of development: A dialectic integration of nature and nurture. *Child Development, 81*(1), 6-22. doi:http://dx.doi.org/10.1111/j.1467-8624.2009.01378.

Shea, S.C. (1998). *Psychiatric Interviewing: The art of understanding.* Philadelphia, PA: Sanders.

Simoneau, T.L., Miklowitz, D.J., & Saleem, R. (1998). Expressed emotion and interactional patterns in the families of bipolar patients. *Journal of Abnormal Psychology*, 107, 497–597.

Steele, H., & Steele. M. (2008). *Clinical applications of the adult attachment interview*. Guilford Press, New York, NY.

Steller, M., & Koehnken, G. (1989). Criteria-based statement analysis. In D. C. Raskin (Ed.), *Psychological methods in criminal investigation and evidence; psychological methods in criminal investigation and evidence* (pp. 217-245, Chapter xvi, 399 Pages) Springer Publishing Co, New York, NY.

Storr, C. L., lalongo, N. S., Anthony, J. C., & Breslau, N. (2007). Childhood antecedents of exposure to traumatic events and posttraumatic stress disorder. *The American Journal of Psychiatry*, *164*(1), 119-125. doi:http://dx.doi.org/10.1176/appi.ajp.164.1.119

Tarabulsy, G. M., Bernier, A., Provost, M. A., Maranda, J., Larose, S., Moss, E., Tessier, R. (2005). Another look inside the gap: Ecological contributions to the transmission of attachment in a sample of adolescent mother-infant dyads. *Developmental Psychology, 41*, 212–224.

Terr, L. C. (1991). Childhood traumas: An outline and overview. *The American Journal*

Tomkins, S. S. (1962). *Affect, imagery, consciousness: Vol. 1. The positive affects.* New York: Springer-Verlag.

Troisi, A. (1999). Ethological research in clinical psychiatry: the study of nonverbal behavior during interviews. *Neuroscience and Biobehavioral Reviews*, 23, 905–13.

van IJzendoorn, M. (1995). Adult attachment representations, parental responsiveness, and infant attachment: A meta-analysis on the predictive validity of the Adult Attachment Interview. *Psychological Bulletin, 117*, 387-403.

van IJzendoorn, M.H., Schuengel, C., & Bakermans-Kranenburg, M.J. (1999). Disorganized attachment in early childhood: meta-analysis of precursors, concomitants, and sequelae. *Developmental Psychopathology, 11*(2), 255-249.

Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: the kappa statistic. *Fam med*, *37*(5), 360-363.

Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities*. John Wiley & Sons.

Zander, E., Willfors, C., Berggren, S., Choque-Olsson, N., Coco, C., Elmund, A., ... & Linder, J. (2016). The objectivity of the Autism Diagnostic Observation Schedule (ADOS) in naturalistic clinical settings. *European child & adolescent psychiatry*, *25*(7), 769-780.

Zaparniuk, J., Yuille, J. C., & Taylor, S. (1995). Assessing the credibility of true and false statements. *International Journal of Law and Psychiatry*.

Zlotnick, C., Zakriski, A. L., Shea, M. T., & Costello, E. (1996). The long-term sequelae of sexual abuse: Support for a complex posttraumatic stress disorder. *Journal of Traumatic Stress, 9*(2), 195-205.

Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 14, pp. 1–59). New York: Academic Press.

# Appendix A

Complete Accounting of Phase 2C (Current) AAI-NVB Manual Behaviors

Table A.1

*Current AAI-NVB Manual Behaviours and Ethogram*

| Behavior | Modifier 1 | Modifier 2 |
|---|---|---|
| Illustrator | | |
| Hands Apart, Resting | Symmetrical<br>Asymmetrical | |
| Hand to Mouth | Cover Mouth<br>Wiping/Rubbing | |
| Hand to Eye | | |
| Hand to Forehead or Chin | Forehead<br>Chin | Touching/Rubbing Chin<br>Resting Chin |
| Hand to Ear | | |
| Hand to Nose | | |
| Hand to Knee | Above knee – Upper thigh<br>Holding below or on kneecap | |
| Hand Clench | | |
| Hand Shrug | Incongruent – Yes<br>Incongruent – No | |
| Hand Stop | | |
| Hand Scratch/Pinch/Pick | | |
| Hands Together | | |
| Protective Object[a] | | |
| Self Soothe[a] | Neck Covered – Yes<br>Repetitive<br>Holding/Hugging | |
| Hand Hiding | | |
| Hand Dismiss | | |
| Shielding Eyes | | |
| Obscene Gesture | | |
| Trunk | Leaning Forward<br>Leaning Back<br>Leading Side<br>Rotated Away | |
| Arms Crossed[a] | | |
| Arms Akimbo | On Hip<br>On Head | |
| Legs | Crossed Tight<br>Open Stance<br>One leg resting on other<br>Tucked on Seat<br>Knees Up<br>Other | |
| Leg Kicking | | |
| Shrug | Full Shrug<br>Without Arms<br>Asymmetric | Incongruent – Yes<br>Incongruent – No |
| Fidgeting | | |
| Tilt | Side Tilt<br>Tilt Up<br>Tilt Down | |

| | | |
|---|---|---|
| | Tilt Down and Away | |
| Nod | Incongruent – Yes | |
| | Incongruent – No | |
| Shake | Incongruent – Yes | |
| | Incongruent – No | |
| Averted | | |
| Lip Bite | | |
| Lip Lick | | |
| Pursed Lips | | |
| Biting | Self Bite | |
| | Object Bite | |
| Oral Fixation | Self | |
| | Object | |
| Micro Expression | Happiness | Masking (With Expression) |
| | Sadness | Neutralized |
| | Anger | |
| | Fear | |
| | Surprise | |
| | Contempt | |
| | Disgust | |
| Incongruent Expression | Happiness | Happiness – Expected |
| | Sadness | Sadness – Expected |
| | Anger | Anger – Expected |
| | Fear | Fear – Expected |
| | Surprise | Surprise – Expected |
| | Contempt | Contempt – Expected |
| | Disgust | Disgust – Expected |
| | Neutral | Neutral - Expected |
| Eye Shift | | |
| Eye Closure/Blinking | Blink | |
| | Full Closure | |
| Eye Rolling | | |
| Deep Inhale/Yawn/Sigh/Sniff | Yawn | |
| | Sigh | |
| | Deep Inhale | |
| | Sniff | |
| Clear Throat/Cough/Gulp | Clearing | |
| | Coughing | |
| | Hard Swallow | |
| Pauses[a] | Silent | |
| | Filled | |
| | Dissociation | |
| Latency[a] | Silent | |
| | Filled | |
| | Dissociation | |
| Fillers | | |
| Speech Errors | | |
| Laughter | Incongruent – Yes | |
| | Incongruent – No | |
| Vocal Raises | | |
| Interruptions | | |
| Corrections | | |
| Diversions | | |
| Generalization | | |
| Lack of Memory | Recovered – Yes | Reasons – Yes |
| | Recovered – No | Reasons – No |

| | |
|---|---|
| Dismissive | |
| Humor | Incongruent – Yes |
| | Incongruent – No |
| Extreme Descriptions | |
| Pardons | Self Pardon |
| | Other Pardon |

[a] Indicates that the behaviour is a state event and should be coded with both start- and endpoints.

# Appendix B

## Sequential Changes to the AAI-NVB Manual: Phases 2A-2C

Table B.1

*Sequential Changes to the AAI-NVB Manual: Phases 2A-C*

| CODE | FACTOR | Phase 1 | Phase 2A | Phase 2B | Phase 2C |
|---|---|---|---|---|---|
| CATEGORY: HANDS | | | ✓ | ✓ | ✓ |
| CM | Covering Mouth | | ✓ | ✓ | ✓ |
| H2E | Hand to Eye | | ✓ | ✓ | ✓ |
| H2F | Hand to Forehead | | ✓ | ✓ | ✓ |
| H2N | Hand to Nose | | ✓ | ✓ | ✓ |
| H2R | Hand to Ear | | ✓ | ✓ | ✓ |
| HAA | Hands Apart Asymmetrical | | ✓ | ✓ | ✓ |
| HAS | Hands Apart Symmetrical | | ✓ | ✓ | ✓ |
| HAK | Hand Above Knee | | ✓ | ✓ | ✓ |
| HBK | Hand Below Knee | | ✓ | ✓ | ✓ |
| HC | Hands Clench | | ✓ | ✓ | ✓ |
| HD | Hand Dismiss | | ✓ | ✓ | ✓ |
| HHi | Hand Hiding | | ✓ | ✓ | ✓ |
| HRC | Chin Resting on Hand | | ✓ | ✓ | ✓ |
| HSh | Hand Shrug | | ✓ | ✓ | ✓ |
| HSPP | Hand Scratch, Pick, Pinch | | ✓ | ✓ | ✓ |
| HSto | Hand Stop | | ✓ | ✓ | ✓ |
| HT | Hands Together | | ✓ | ✓ | ✓ |
| HTC | Hand Touching Chin | | ✓ | ✓ | ✓ |
| Ill | Illustrator | | ✓ | ✓ | ✓ |
| OG | Obscene Gesture | | ✓ | ✓ | ✓ |
| SE | Shielding Eyes | | ✓ | ✓ | ✓ |
| SS | Self Soothe, Repetitive Motion | ✓ | ✓ | ✓ | ✓ |
| SS_BH | Self Soothe, Body Holding | | ✓ | ✓ | ✓ |
| SS_NC | Self Soothe, Neck Covered | | ✓ | ✓ | ✓ |
| WM | Wiping Mouth | | ✓ | ✓ | ✓ |
| FO | Fidget Object | | ✓ | ✓ | |
| CATEGORY: BODY | | | ✓ | ✓ | ✓ |
| FDG | Fidgeting | ✓ | | | ✓ |
| BF | Body Fidget | | ✓ | ✓ | |
| LBA | Leaning Back | | ✓ | ✓ | ✓ |

| CODE | FACTOR | Phase 1 | Phase 2A | Phase 2B | Phase 2C |
|---|---|---|---|---|---|
| LF | Leaning Forward | | ✓ | ✓ | ✓ |
| LS | Leaning Side | | ✓ | ✓ | ✓ |
| PO | Protective Object | ✓ | ✓ | ✓ | ✓ |
| RA | Rotated Away | | ✓ | ✓ | ✓ |
| CATEGORY: LIMBS | | | ✓ | ✓ | ✓ |
| AAHi | Arms Akimbo, Hips | | ✓ | ✓ | ✓ |
| AAHe | Arms Akimbo, Head | | ✓ | ✓ | ✓ |
| KU | Knees Up | | | | ✓ |
| LCT | Legs Crossed Tight | | | | ✓ |
| LK | Leg Kicking | | ✓ | ✓ | ✓ |
| LO | Legs Other | | | | ✓ |
| LOS | Legs Open Stance | | | | ✓ |
| LRO | Legs Resting on Other | | | | ✓ |
| TOS | Tucked on Seat | | | | ✓ |
| AC | Arms Crossed | | | | ✓ |
| OAC | One Arm Crossed | | ✓ | ✓ | |
| TAC | Two Arms Crossed | | ✓ | ✓ | |
| CATEGORY: HEAD AND SHOULDER | | | ✓ | ✓ | ✓ |
| AS | Asymmetric Shrug | | ✓ | ✓ | ✓ |
| FS | Full Shrug | | ✓ | ✓ | ✓ |
| SWA | Shrug without arms | | ✓ | ✓ | ✓ |
| HA | Head Averted | | ✓ | ✓ | ✓ |
| HDA | Tilt Down and Away | | ✓ | ✓ | ✓ |
| HN | Head Nod | | ✓ | ✓ | ✓ |
| HS | Head Shake | | ✓ | ✓ | ✓ |
| HTD | Tilt Down | | ✓ | ✓ | ✓ |
| HTU | Tilt Up | | ✓ | ✓ | ✓ |
| HST | Side Tilt | | ✓ | ✓ | ✓ |
| CATEGORY: INCONGRUENCE | | ✓ | ✓ | ✓ | ✓ |
| AS_In | Asymmetrical Shrug, Incongruent | | ✓ | ✓ | ✓ |
| FS_In | Full Shrug, Incongruent | | ✓ | ✓ | ✓ |
| SWA_In | Shrug without Arms, Incongruent | | ✓ | ✓ | ✓ |
| HSh_In | Hand Shrug Incongruent | | ✓ | ✓ | ✓ |
| H_In | Humour, Incongruent | | ✓ | ✓ | ✓ |
| HN_In | Head Nod, Incongruent | | ✓ | ✓ | ✓ |
| L_In | Laughter, Incongruent | | ✓ | ✓ | ✓ |

| CODE | FACTOR | Phase 1 | Phase 2A | Phase 2B | Phase 2C |
|---|---|---|---|---|---|
| HS_In | Head Shake, Incongruent | | ✓ | ✓ | ✓ |
| HSto_In | Hand Stop, Incongruent | | ✓ | ✓ | |
| HD_In | Hand Dismiss, Incongruent | | ✓ | ✓ | |
| HDA_In | Tilt Down and Away, Incongruent | | ✓ | ✓ | |
| OG_In | Obscene Gesture, Incongruent | | ✓ | ✓ | |
| CATEGORY: DURATION | | | ✓ | ✓ | ✓ |
| SS_D | Self Soothe, Duration | ✓ | ✓ | ✓ | ✓ |
| PO_D | Protective Object, Duration | ✓ | ✓ | ✓ | ✓ |
| LCT_D | Legs Crossed Tight, Duration | | | | ✓ |
| LOS_D | Legs Open Stance, Duration | | | | ✓ |
| LRO_D | Legs Resting on Other, Duration | | | | ✓ |
| TOS_D | Tucked on Seat, Duration | | | | ✓ |
| KU_D | Knees Up, Duration | | | | ✓ |
| LO_D | Legs Other, Duration | | | | ✓ |
| Dist_D | Dissociation, Duration | ✓ | ✓ | ✓ | ✓ |
| AC_D | Arms Crossed, Duration | | | | ✓ |
| OAC_D | One Arm Crossed, Duration | | ✓ | ✓ | |
| TAC_D | Two Arms Crossed, Duration | | ✓ | ✓ | |
| P_D | Pause, Duration | ✓ | ✓ | ✓ | ✓ |
| LAT_D | Latency, Duration | ✓ | ✓ | ✓ | ✓ |
| CATEGORY: FACE & EYES | | | ✓ | ✓ | ✓ |
| LB | Lip Bite | | ✓ | ✓ | ✓ |
| LL | Lip Lick, Smack, Mouth Movement | | ✓ | ✓ | ✓ |
| PL | Pursed Lips | | ✓ | ✓ | ✓ |
| BS | Biting Self | | ✓ | ✓ | ✓ |
| OFS | Oral Fixation Self | | ✓ | ✓ | ✓ |
| OFO | Oral Fixation Other | | ✓ | ✓ | ✓ |
| BO | Biting Object | | ✓ | ✓ | ✓ |
| GA | Gaze Aversion | ✓ | | | ✓ |
| ELA | Eyes Looking Away | | ✓ | ✓ | |
| B | Blink | | ✓ | ✓ | ✓ |
| FEC | Full Eye Closure | | ✓ | ✓ | ✓ |
| ER | Eye Rolling | | ✓ | ✓ | ✓ |
| ES | Eyes Searching | | ✓ | ✓ | |
| CATEGORY: OTHER | | | ✓ | ✓ | ✓ |
| DI | Deep Inhale | | ✓ | ✓ | ✓ |

| CODE | FACTOR | Phase 1 | Phase 2A | Phase 2B | Phase 2C |
|---|---|---|---|---|---|
| S | Sigh | | √ | √ | √ |
| Y | Yawn | | √ | √ | √ |
| Sn | Sniff | | √ | √ | √ |
| C | Coughing | | √ | √ | √ |
| CT | Clearing Throat | | √ | √ | √ |
| HSW | Hard Swallow | | √ | √ | √ |
| CATEGORY: PARALINGUISTICS 1 | | | √ | √ | √ |
| Dis | Dissociation | √ | √ | √ | √ |
| F | Fillers | √ | √ | √ | √ |
| SErr | Speech Errors | | √ | √ | √ |
| L | Laughter | √ | √ | √ | √ |
| VR | Vocal Raises | | √ | √ | √ |
| I | Interruptions | | √ | √ | √ |
| P | Pause | √ | √ | √ | √ |
| LAT | Latency | √ | √ | √ | √ |
| SC | Scoff | | √ | √ | |
| CATEGORY: PARALINGUISTICS 2 | | | √ | √ | √ |
| Cor | Corrections | | √ | √ | √ |
| VD | Verbal Diversions | √ | √ | √ | √ |
| G | Generalizations | √ | √ | √ | √ |
| LM | Lack of Memory | √ | √ | √ | √ |
| LM_Rc | Lack of Memory, Recovered | | √ | √ | √ |
| LM_Re | Lack of Memory, Reasons | | √ | √ | √ |
| DM | Dismissive | √ | √ | √ | √ |
| H | Humour | √ | √ | √ | √ |
| ED | Extreme Descriptions | | √ | √ | √ |
| PS | Pardon, Self | | √ | √ | √ |
| POt | Pardon, Other | | √ | √ | √ |
| SAR | Sarcasm | | √ | √ | |
| CATEGORY: AFFECT 1 | | | √ | √ | √ |
| ME | Micro Expression | √ | √ | √ | |
| MEM | Micro Expression Masking | | √ | √ | |
| MEN | Micro Expression Neutralized | | √ | √ | |
| ME_Hap | Micro Expression, Happiness | | | | √ |
| ME_Sad | Micro Expression, Sadness | | | | √ |
| ME_Ang | Micro Expression, Anger | | | | √ |

| CODE | FACTOR | Phase 1 | Phase 2A | Phase 2B | Phase 2C |
|---|---|---|---|---|---|
| ME_Fe | Micro Expression, Fear | | | | ✓ |
| ME_Sur | Micro Expression, Surprise | | | | ✓ |
| ME_Con | Micro Expression, Contempt | | | | ✓ |
| ME_Dis | Micro Expression, Disgust | | | | ✓ |
| MEM_Hap | Micro Expression Mask, Happiness | | | | ✓ |
| MEM_Sad | Micro Expression Mask, Sadness | | | | ✓ |
| MEM_Ang | Micro Expression Mask, Anger | | | | ✓ |
| MEM_Fe | Micro Expression Mask, Fear | | | | ✓ |
| MEM_Sur | Micro Expression Mask, Surprise | | | | ✓ |
| MEM_Con | Micro Expression Mask, Contempt | | | | ✓ |
| MEM_Dis | Micro Expression Mask, Disgust | | | | ✓ |
| MEM_Neu | Micro Expression Mask, Neutral | | | | ✓ |
| CATEGORY: AFFECT 2 | | | ✓ | ✓ | ✓ |
| IE | Incongruent Expression | | ✓ | ✓ | |
| IEE | Incongruent Expression Expected | | ✓ | ✓ | |
| IE_Hap | Incongruent Expression, Happiness | | | | ✓ |
| IE_Sad | Incongruent Expression, Sadness | | | | ✓ |
| IE_Ang | Incongruent Expression, Anger | | | | ✓ |
| IE_Fe | Incongruent Expression, Fear | | | | ✓ |
| IE_Sur | Incongruent Expression, Surprise | | | | ✓ |
| IE_Con | Incongruent Expression, Contempt | | | | ✓ |
| IE_Dis | Incongruent Expression, Disgust | | | | ✓ |
| IE_Neu | Incongruent Expression, Neutral | | | | ✓ |
| IEE_Hap | Incongruent Expression Expected, Happiness | | | | ✓ |
| IEE_Sad | Incongruent Expression Expected, Sadness | | | | ✓ |
| IEE_Ang | Incongruent Expression Expected, Anger | | | | ✓ |
| IEE_Fe | Incongruent Expression Expected, Fear | | | | ✓ |
| IEE_Sur | Incongruent Expression Expected, Surprise | | | | ✓ |
| IEE_Con | Incongruent Expression Expected, Contempt | | | | ✓ |
| IEE_Dis | Incongruent Expression Expected, Disgust | | | | ✓ |
| IEE_Neu | Incongruent Expression Expected, Neutral | | | | ✓ |

# Appendix C

Intraclass Correlation Results for each Category Across Three Subjects

Table C.1

*Intraclass Correlation Results for each Behavioural Category Across for Three Subjects*

| Group | Subject 1 | | Subject 2 | | Subject 3 | |
|---|---|---|---|---|---|---|
| | ICC | CI (95%) | ICC | CI (95%) | ICC | CI (95%) |
| Hands | 0.8 | 0.702, 0.886 | 0.705 | 0.581, 0.824 | 0.881 | 0.81, 0.937 |
| Body | 0.146 | -0.007, 0.612 | 0.0482 | -0.04, 0.433 | 0.855 | 0.662, 0.974 |
| Limbs | 0.305 | 0.088, 0.759 | 0.494 | 0.22, 0.867 | 0.117[a] | -0.014, 0.431 |
| H & S | 0.721 | 0.519, 0.901 | 0.875 | 0.75, 0.96 | 0.503 | 0.271, 0.789 |
| Incongruence | 0.342 | 0.25, 0.456 | 0.467 | 0.368, 0.579 | 0.484 | 0.233, 0.813 |
| Duration | 0.353 | 0.138, 0.725 | 0.686 | 0.452, 0.906 | 0.724 | 0.548, 0.879 |
| Face & Eyes | 0.915 | 0.835, 0.97 | 0.944 | 0.888, 0.98 | 0.843 | 0.701, 0.945 |
| Other | 0.342 | 0.128, 0.743 | 0.413 | 0.175, 0.794 | 0.523 | 0.25, 0.856 |
| Para 1 | 0.676 | 0.452, 0.89 | 0.830 | 0.665, 0.949 | 0.690 | 0.473, 0.888 |
| Para 2 | 0.386 | 0.198, 0.67 | 0.370 | 0.187, 0.657 | 0.356 | 0.156, 0.667 |
| Affect 1 | | | | | 0.440 | 0.25, 0.687 |
| Affect 2 | | | | | 0.301 | 0.142, 0.551 |

[a] 0.759 without leg kicking

# Appendix D

Descriptive Statistics for Behaviour Frequency by Category During Each Phase

Table E.1

*Descriptive Statistics for the Hands Behaviours in Each Phase*

| Behavior | M | Mdn | SD | Min | Max | Skew | Kurt |
|---|---|---|---|---|---|---|---|
| Phase 2A: Subject 1 | | | | | | | |
| Illustrators | 60.7 | 62 | 22.221 | 23 | 98 | -0.137 | -0.143 |
| Hands | 0.3 | 0 | 0.483 | 0 | 1 | 1.035 | -1.224 |
| Hands | 1.6 | 1 | 2.221 | 0 | 7 | 1.85 | 3.593 |
| Covering Mouth | 1 | 1 | 0.471 | 0 | 2 | 0 | 4.5 |
| Wiping Mouth | 0.8 | 1 | 0.422 | 0 | 1 | -1.779 | 1.406 |
| Hand to Eye | 0.7 | 1 | 0.483 | 0 | 1 | -1.035 | -1.224 |
| Hand to | 0.3 | 0 | 0.483 | 0 | 1 | 1.035 | -1.224 |
| Hand Touching | 0.5 | 0 | 0.85 | 0 | 2 | 1.358 | 0.107 |
| Chin Resting in | 0.3 | 0 | 0.675 | 0 | 2 | 2.277 | 4.765 |
| Hand to Ear | 0.1 | 0 | 0.316 | 0 | 1 | 3.162 | 10 |
| Hand to Nose | 1.5 | 2 | 0.85 | 0 | 2 | -1.358 | 0.107 |
| Hand Above | 2.1 | 1.5 | 2.601 | 0 | 9 | 2.42 | 6.675 |
| Hand Below | 1.7 | 1.5 | 1.059 | 0 | 3 | 0.042 | -1.238 |
| Hand Clench | 1 | 0.5 | 1.155 | 0 | 3 | 0.541 | -1.393 |
| Hand Shrug | 5.8 | 4.5 | 6.179 | 0 | 22 | 2.286 | 6.258 |
| Hand Stop | 1 | 0 | 1.491 | 0 | 4 | 1.258 | 0.257 |
| Hand Scratch, | 8.4 | 7.5 | 7.989 | 0 | 25 | 1.05 | 0.881 |
| Hands Together | 33.2 | 34 | 10.942 | 17 | 51 | 0.146 | -0.911 |
| Self Soothe | 20.9 | 12.5 | 17.729 | 5 | 65 | 2.021 | 4.318 |
| Neck Covered | 0.1 | 0 | 0.316 | 0 | 1 | 3.162 | 10 |
| Body Holding | 4.5 | 3.5 | 4.95 | 0 | 15 | 1.33 | 1.139 |
| Hands Hiding | 3 | 3 | 2.494 | 0 | 6 | -0.107 | -1.859 |
| Hand Dismiss | 1.1 | 0.5 | 1.287 | 0 | 3 | 0.556 | -1.576 |
| Shielding Eyes | - | - | - | - | - | - | - |
| Obscene Gesture | - | - | - | - | - | - | - |
| Fidget Object | 7 | 5.5 | 7.04 | 0 | 21 | 1.132 | 0.476 |
| Phase 2B: Subject 2 | | | | | | | |
| Illustrators | 71.8 | 54 | 43.006 | 26 | 140 | 0.863 | -1.029 |
| Hands | - | - | - | - | - | - | - |
| Hands | 10.2 | 2.5 | 16.349 | 0 | 42 | 1.606 | 1.055 |

| Behavior | M | Mdn | SD | Min | Max | Skew | Kurt |
|---|---|---|---|---|---|---|---|
| Covering Mouth | 22.6 | 23.5 | 6.31 | 8 | 29 | -1.472 | 2.451 |
| Wiping Mouth | 2.2 | 2.5 | 1.229 | 0 | 4 | -0.467 | -0.544 |
| Hand to Eye | 1.2 | 1 | 0.789 | 0 | 3 | 1.29 | 2.985 |
| Hand to | 1.6 | 1.5 | 0.966 | 0 | 3 | 0.111 | -0.623 |
| Hand Touching | 1 | 1 | 0.943 | 0 | 3 | 0.994 | 1.185 |
| Chin Resting in | - | - | - | - | - | - | - |
| Hand to Ear | 0.3 | 0 | 0.483 | 0 | 1 | 1.035 | -1.224 |
| Hand to Nose | 12.6 | 13 | 2.119 | 8 | 15 | -1.094 | 1.414 |
| Hand Above | 0.3 | 0 | 0.483 | 0 | 1 | 1.035 | -1.224 |
| Hand Below | 0.5 | 0.5 | 0.527 | 0 | 1 | 0 | -2.571 |
| Hand Clench | 1.1 | 1 | 0.568 | 0 | 2 | 0.091 | 1.498 |
| Hand Shrug | 3.3 | 2 | 4.473 | 0 | 14 | 1.745 | 3.185 |
| Hand Stop | 0.2 | 0 | 0.422 | 0 | 1 | 1.779 | 1.406 |
| Hand Scratch, | 2.5 | 1.5 | 3.659 | 0 | 12 | 2.254 | 5.78 |
| Hands Together | - | - | - | - | - | - | - |
| Self Soothe | 2 | 2 | 1.414 | 0 | 5 | 0.884 | 1.226 |
| Neck Covered | - | - | - | - | - | - | - |
| Body Holding | - | - | - | - | - | - | - |
| Hands Hiding | 0.1 | 0 | 0.316 | 0 | 1 | 3.162 | 10 |
| Hand Dismiss | 0.4 | 0 | 0.699 | 0 | 2 | 1.658 | 2.045 |
| Shielding Eyes | - | - | - | - | - | - | - |
| Obscene Gesture | - | - | - | - | - | - | - |
| Fidget Object | 0.5 | 0 | 0.85 | 0 | 2 | 1.358 | 0.107 |
| Phase 2C: Subject 3 | | | | | | | |
| Illustrators | 101 | 102 | 30.742 | 58 | 145 | 0.018 | -1.358 |
| Hands | 5 | 6 | 2.915 | 1 | 9 | -0.317 | -0.475 |
| Hands | 3 | 3 | 1.727 | 0 | 5 | -0.191 | -0.564 |
| Covering Mouth | 4 | 4 | 1.414 | 1 | 5 | -0.808 | -0.229 |
| Wiping Mouth | 3 | 3 | 0.916 | 2 | 5 | 0.488 | 0.421 |
| Hand to Eye | 4 | 4 | 1.669 | 0 | 5 | -1.936 | 4.175 |
| Hand to | 1 | 1 | 0.756 | 0 | 2 | 0 | -0.7 |
| Hand Touching | 0 | 0 | 0.463 | 0 | 1 | 1.44 | 0 |
| Chin Resting in | 1 | 0 | 0.756 | 0 | 2 | 1.323 | 0.875 |
| Hand to Ear | 1 | 1 | 0.463 | 0 | 1 | -1.44 | 0 |
| Hand to Nose | 2 | 2 | 1.282 | 0 | 4 | 0.611 | -0.021 |
| Hand Above | 1 | 1 | 1.309 | 0 | 3 | 1.018 | -0.7 |

| Behavior | M | Mdn | SD | Min | Max | Skew | Kurt |
|---|---|---|---|---|---|---|---|
| Hand Below | 2 | 2 | 0.744 | 0 | 2 | -1.951 | 3.205 |
| Hand Clench | 4 | 4 | 1.885 | 2 | 7 | 0.275 | -1.483 |
| Hand Shrug | 9 | 9 | 4.432 | 2 | 17 | 0.458 | 1.27 |
| Hand Stop | 4 | 4 | 3.068 | 1 | 11 | 1.547 | 3.213 |
| Hand Scratch, | 3 | 2 | 3.251 | 0 | 10 | 2.128 | 5.078 |
| Hands Together | 37 | 37 | 10.716 | 17 | 51 | -0.458 | 0.715 |
| Self Soothe | 0 | 0 | 0.354 | 0 | 1 | 2.828 | 8 |
| Neck Coverered | - | - | - | - | - | - | - |
| Body Holding | 0 | 0 | 0.354 | 0 | 1 | 2.828 | 8 |
| Hands Hiding | 10 | 3 | 16.848 | 2 | 51 | 2.684 | 7.342 |
| Hand Dismiss | 1 | 0 | 2.504 | 0 | 7 | 2.054 | 4.054 |
| Shielding Eyes | 0 | 0 | 0.463 | 0 | 1 | 1.44 | 0 |
| Obscene Gesture | - | - | - | - | - | - | - |
| Fidget Object | | | | | | | |

Table E.2

*Descriptive Statistics for the Body Behaviours in Each Phase*

| Behavior | M | Mdn | SD | Min | Max | Skew | Kurt |
|---|---|---|---|---|---|---|---|
| Phase 2A: Subject 1 | | | | | | | |
| Lean Forward | 1.5 | 1.5 | 1.08 | 0 | 3 | 0 | -1.032 |
| Lean Back | 1.2 | 1 | 1.135 | 0 | 3 | 0.661 | -0.709 |
| Lean Side | 0.4 | 0 | 0.699 | 0 | 2 | 1.658 | 2.045 |
| Rotate Away | 0.4 | 0 | 0.699 | 0 | 2 | 1.658 | 2.045 |
| Protective Object | 1.8 | 0 | 5.692 | 0 | 18 | 3.162 | 10 |
| Body Fidget | 8.3 | 3.5 | 12.746 | 0 | 42 | 2.472 | 6.47 |
| Fidgeting | | | | | | | |
| Phase 2B: Subject 2 | | | | | | | |
| Lean Forward | 0.4 | 0 | 0.843 | 0 | 2 | 1.779 | 1.406 |
| Lean Back | 0.7 | 0.5 | 0.823 | 0 | 2 | 0.687 | -1.043 |
| Lean Side | 1.1 | 0 | 2.807 | 0 | 9 | 3.038 | 9.395 |
| Rotate Away | - | - | - | - | - | - | - |
| Protective Object | - | - | - | - | - | - | - |
| Body Fidget | 4.5 | 0.5 | 10.835 | 0 | 35 | 3.04 | 9.403 |
| Fidgeting | | | | | | | |
| Phase 2C: Subject 3 | | | | | | | |

| Behavior | M | Mdn | SD | Min | Max | Skew | Kurt |
|---|---|---|---|---|---|---|---|
| Lean Forward | 3 | 1 | 4.324 | 0 | 12 | 1.626 | 2.276 |
| Lean Back | 4 | 4 | 4 | 0 | 11 | 0.732 | -0.523 |
| Lean Side | 6 | 4 | 6.833 | 0 | 20 | 1.373 | 1.888 |
| Rotate Away | 1 | 0 | 1.753 | 0 | 5 | 2.627 | 7.027 |
| Protective Object | 1 | 1 | 1.389 | 0 | 4 | 1.12 | 1.106 |
| Body Fidget | | | | | | | |
| Fidgeting | 54 | 53 | 18.843 | 24 | 82 | -0.081 | -0.567 |

Table E.3

*Descriptive Statistics for the Limbs Behaviours in Each Phase*

| Behavior | M | Mdn | SD | Min | Max | Skew | Kurt |
|---|---|---|---|---|---|---|---|
| Phase 2A: Subject 1 | | | | | | | |
| Arms Crossed | | | | | | | |
| One Arm Crossed | 2.7 | 1.5 | 3.164 | 0 | 8 | 0.791 | -1.137 |
| Two Arms | 1.2 | 1.5 | 0.919 | 0 | 2 | -0.473 | -1.807 |
| Arms Akimbo | 0.5 | 0.5 | 0.527 | 0 | 1 | 0 | -2.571 |
| Arms Akimbo | - | - | - | - | - | - | - |
| Legs Crossed | | | | | | | |
| Legs Open | | | | | | | |
| Leg Resting on | | | | | | | |
| Tucked on Seat | | | | | | | |
| Leg Kicking | - | - | - | - | - | - | - |
| Knees Up | | | | | | | |
| Legs Other | | | | | | | |
| Phase 2B: Subject 2 | | | | | | | |
| Arms Crossed | | | | | | | |
| One Arm Crossed | - | - | - | - | - | - | - |
| Two Arms | - | - | - | - | - | - | - |
| Arms Akimbo | 0.7 | 1 | 0.675 | 0 | 2 | 0.434 | -0.283 |
| Arms Akimbo | - | - | - | - | - | - | - |
| Legs Crossed | | | | | | | |
| Legs Open | | | | | | | |
| Leg Resting on | | | | | | | |
| Tucked on Seat | | | | | | | |
| Leg Kicking | - | - | - | - | - | - | - |

| Behavior | M | Mdn | SD | Min | Max | Skew | Kurt |
|---|---|---|---|---|---|---|---|
| Knees Up | | | | | | | |
| Legs Other | | | | | | | |
| | | | Phase 2C: Subject 3 | | | | |
| Arms Crossed | 3 | 3 | 0.744 | 2 | 4 | 0.824 | -0.152 |
| One Arm Crossed | | | | | | | |
| Two Arms | | | | | | | |
| Arms Akimbo | - | - | - | - | - | - | - |
| Arms Akimbo | - | - | - | - | - | - | - |
| Legs Crossed | 4 | 4 | 1.768 | 0 | 5 | -1.309 | 1.68 |
| Legs Open | 1 | 1 | 0.463 | 0 | 1 | -1.44 | 0 |
| Leg Resting on | 0 | 0 | 1.061 | 0 | 3 | 2.828 | 8 |
| Tucked on Seat | - | - | - | - | - | - | - |
| Leg Kicking | 8 | 2 | 15.25 | 0 | 45 | 2.757 | 7.694 |
| Knees Up | - | - | - | - | - | - | - |
| Legs Other | - | - | - | - | - | - | - |

Table E.4

*Descriptive Statistics for the Head and Shoulders Behaviours in Each Phase*

| Behavior | M | Mdn | SD | Min | Max | Skew | Kurt |
|---|---|---|---|---|---|---|---|
| | | | Phase 2A: Subject 1 | | | | |
| Full Shrug | 0.3 | 0 | 0.675 | 0 | 2 | 2.277 | 4.765 |
| Shrug Without | 2 | 2 | 1.563 | 0 | 4 | 0 | -1.782 |
| Asymmetric Shrug | 0.9 | 0.5 | 1.101 | 0 | 3 | 0.863 | -0.522 |
| Head Side Tilt | 18.4 | 19 | 10.926 | 6 | 42 | 0.989 | 1.207 |
| Head Tilt Up | 5.8 | 5.5 | 3.615 | 1 | 13 | 0.661 | 0.3 |
| Head Tilt Down | 0.5 | 0 | 0.707 | 0 | 2 | 1.179 | 0.571 |
| Head Down Away | 0.6 | 0 | 0.843 | 0 | 2 | 1.001 | -0.665 |
| Head Nod | 39.1 | 34.5 | 19.365 | 13 | 68 | 0.408 | -1.075 |
| Head Shake | 13.6 | 13.5 | 4.742 | 6 | 22 | 0.234 | -0.099 |
| Head Averted | 8.8 | 8.5 | 5.493 | 0 | 18 | 0.1 | -0.474 |
| | | | Phase 2B: Subject 2 | | | | |
| Full Shrug | - | - | - | - | - | - | - |
| Shrug Without | 0.2 | 0 | 0.422 | 0 | 1 | 1.779 | 1.406 |
| Asymmetric Shrug | 0.2 | 0 | 0.422 | 0 | 1 | 1.779 | 1.406 |
| Head Side Tilt | 15.8 | 10.5 | 10.768 | 4 | 34 | 0.65 | -1.169 |

| Behavior | M | Mdn | SD | Min | Max | Skew | Kurt |
|---|---|---|---|---|---|---|---|
| Head Tilt Up | 1 | 0 | 1.333 | 0 | 3 | 0.703 | -1.577 |
| Head Tilt Down | 1.3 | 0.5 | 2.058 | 0 | 6 | 1.793 | 2.425 |
| Head Down Away | - | - | - | - | - | - | - |
| Head Nod | 48.8 | 50.5 | 13.011 | 30 | 67 | -0.204 | -1.144 |
| Head Shake | 26.6 | 27.5 | 9.143 | 13 | 40 | -0.226 | -1.022 |
| Head Averted | 1.7 | 0.5 | 2.584 | 0 | 8 | 1.91 | 3.713 |
| Phase 2C: Subject 3 | | | | | | | |
| Full Shrug | 2 | 3 | 1.642 | 0 | 4 | -0.254 | -1.963 |
| Shrug Without | 2 | 2 | 1.923 | 0 | 6 | 0.897 | 0.59 |
| Asymmetric Shrug | 4 | 2 | 4.291 | 0 | 11 | 0.823 | -0.98 |
| Head Side Tilt | 45 | 50 | 17.204 | 17 | 67 | -0.702 | -0.506 |
| Head Tilt Up | 11 | 2 | 17.203 | 0 | 41 | 1.455 | 0.153 |
| Head Tilt Down | 14 | 4 | 21.299 | 0 | 61 | 1.982 | 3.668 |
| Head Down Away | 5 | 0 | 7.873 | 0 | 20 | 1.546 | 0.981 |
| Head Nod | 38 | 26 | 25.707 | 20 | 98 | 2.333 | 5.774 |
| Head Shake | 23 | 21 | 11.145 | 6 | 43 | 0.528 | 0.838 |
| Head Averted | 17 | 18 | 10.011 | 0 | 30 | -0.455 | -0.488 |

Table E.5

*Descriptive Statistics for the Face and Eyes Behaviours in Each Phase*

| Behavior | M | Mdn | SD | Min | Max | Skew | Kurt |
|---|---|---|---|---|---|---|---|
| Phase 2A: Subject 1 | | | | | | | |
| Lip Bite | 0.1 | 0 | 0.316 | 0 | 1 | 3.162 | 10 |
| Lip Lick | 19.3 | 18 | 11.499 | 6 | 40 | 0.671 | -0.383 |
| Pursed Lips | 10.9 | 10 | 8.225 | 0 | 29 | 1.05 | 1.696 |
| Biting Self | - | - | - | - | - | - | - |
| Biting Other | - | - | - | - | - | - | - |
| Oral Fixation Self | - | - | - | - | - | - | - |
| Oral Fixation | - | - | - | - | - | - | - |
| Gaze Aversion | | | | | | | |
| Eyes Looking | 65.4 | 56.5 | 48.635 | 14 | 166 | 1.198 | 0.861 |
| Eyes Searching | 18.3 | 6.5 | 24.046 | 0 | 73 | 1.44 | 1.918 |
| Blink | 385.6 | 391.5 | 96.918 | 179 | 510 | -0.914 | 1.457 |
| Full Eye Closure | 30.8 | 22.5 | 20.66 | 11 | 76 | 1.293 | 1.312 |
| Eye Rolling | 6.1 | 1.5 | 9.927 | 0 | 31 | 2.107 | 4.536 |

| Behavior | M | Mdn | SD | Min | Max | Skew | Kurt |
|---|---|---|---|---|---|---|---|
| Phase 2B: Subject 2 | | | | | | | |
| Lip Bite | 0.6 | 0 | 1.075 | 0 | 3 | 1.691 | 1.864 |
| Lip Lick | 12.4 | 13 | 7.321 | 0 | 24 | -0.238 | -0.27 |
| Pursed Lips | 2.6 | 1.5 | 2.716 | 0 | 7 | 0.945 | -0.754 |
| Biting Self | - | - | - | - | - | - | - |
| Biting Other | - | - | - | - | - | - | - |
| Oral Fixation Self | 1.6 | 0 | 5.06 | 0 | 16 | 3.162 | 10 |
| Oral Fixation | - | - | - | - | - | - | - |
| Gaze Aversion | | | | | | | |
| Eyes Looking | 37 | 31 | 26.891 | 0 | 77 | 0.206 | -1.411 |
| Eyes Searching | 12.9 | 5 | 13.916 | 0 | 41 | 0.999 | -0.026 |
| Blink | 328.1 | 336 | 70.848 | 151 | 401 | -1.855 | 4.679 |
| Full Eye Closure | 12.9 | 11.5 | 11.06 | 0 | 31 | 0.66 | -0.779 |
| Eye Rolling | 0.9 | 0 | 2.183 | 0 | 7 | 2.961 | 9.005 |
| Phase 2C: Subject 3 | | | | | | | |
| Lip Bite | 1 | 1 | 0 | 1 | 1 | #DIV/0! | #DIV/0! |
| Lip Lick | 18 | 18 | 5.33 | 8 | 24 | -0.639 | 0.416 |
| Pursed Lips | 11 | 8 | 10.809 | 0 | 33 | 1.438 | 2.071 |
| Biting Self | 0 | 0 | 0.354 | 0 | 1 | 2.828 | 8 |
| Biting Other | 0 | 0 | 0 | 0 | 0 | #DIV/0! | #DIV/0! |
| Oral Fixation Self | 1 | 0 | 1.727 | 0 | 5 | 2.472 | 6.375 |
| Oral Fixation | 0 | 0 | 0 | 0 | 0 | #DIV/0! | #DIV/0! |
| Gaze Aversion | 206 | 180 | 118.616 | 44 | 373 | 0.233 | -1.407 |
| Eyes Looking | | | | | | | |
| Eyes Searching | | | | | | | |
| Blink | 328 | 324 | 93.97 | 195 | 451 | -0.109 | -1.601 |
| Full Eye Closure | 29 | 15 | 27.422 | 4 | 83 | 1.268 | 0.764 |
| Eye Rolling | 4 | 3 | 3.546 | 0 | 10 | 0.641 | -0.414 |

Table E.6

*Descriptive Statistics for the Reflexive Behaviours in Each Phase*

| Behavior | M | Mdn | SD | Min | Max | Skew | Kurt |
|---|---|---|---|---|---|---|---|
| Phase 2A: Subject 1 | | | | | | | |
| Deep Inhale | 5.9 | 5 | 5.466 | 0 | 20 | 2.139 | 5.703 |
| Sigh | 3.2 | 3 | 3.048 | 0 | 9 | 0.764 | -0.122 |

| Behavior | M | Mdn | SD | Min | Max | Skew | Kurt |
|---|---|---|---|---|---|---|---|
| Yawn | - | - | - | - | - | - | - |
| Sniff | 2.4 | 1 | 3.307 | 0 | 9 | 1.219 | 0.002 |
| Clear Throat | 3.9 | 4 | 1.663 | 1 | 7 | 0.377 | 0.921 |
| Cough | - | - | - | - | - | - | - |
| Hard Swallow | 1 | 0 | 1.7 | 0 | 5 | 1.867 | 2.931 |
| Phase 2B: Subject 2 | | | | | | | |
| Deep Inhale | 2.9 | 2 | 2.846 | 0 | 9 | 1.536 | 1.537 |
| Sigh | 0.5 | 0 | 0.972 | 0 | 3 | 2.27 | 5.356 |
| Yawn | - | - | - | - | - | - | - |
| Sniff | 0.3 | 0 | 0.675 | 0 | 2 | 2.277 | 4.765 |
| Clear Throat | 0.2 | 0 | 0.422 | 0 | 1 | 1.779 | 1.406 |
| Cough | - | - | - | - | - | - | - |
| Hard Swallow | - | - | - | - | - | - | - |
| Phase 2C: Subject 3 | | | | | | | |
| Deep Inhale | 26 | 23 | 22.123 | 1 | 72 | 1.341 | 2.106 |
| Sigh | 3 | 2 | 2.563 | 0 | 8 | 1.56 | 3.028 |
| Yawn | 0 | 0 | 0.744 | 0 | 2 | 1.951 | 3.205 |
| Sniff | 1 | 0 | 1.069 | 0 | 3 | 2.339 | 5.469 |
| Clear Throat | 0 | 0 | 0.354 | 0 | 1 | 2.828 | 8 |
| Cough | - | - | - | - | - | - | - |
| Hard Swallow | 10 | 10 | 7.249 | 0 | 22 | 0.269 | -0.111 |

Table E.7

*Descriptive Statistics for the Paralinguistics I Behaviours in Each Phase*

| Behavior | M | Mdn | SD | Min | Max | Skew | Kurt |
|---|---|---|---|---|---|---|---|
| Phase 2A: Subject 1 | | | | | | | |
| Pause | 20.6 | 18.5 | 12.322 | 6 | 52 | 1.998 | 5.356 |
| Latency | 4.6 | 4.5 | 2.591 | 1 | 10 | 0.744 | 0.937 |
| Dissociation | - | - | - | - | - | - | - |
| Fillers | 70.9 | 58 | 37.96 | 31 | 162 | 1.702 | 3.355 |
| Speech Errors | 10 | 6.5 | 12.009 | 0 | 36 | 1.211 | 1.063 |
| Laughter | 22.2 | 23 | 6.812 | 11 | 33 | -0.376 | -0.035 |
| Vocal Raises | 12.3 | 7 | 12.41 | 0 | 29 | 0.452 | -1.953 |
| Interruptions | 2.8 | 2 | 3.084 | 0 | 11 | 2.45 | 6.777 |
| Scoff | - | - | - | - | - | - | - |

| Behavior | M | Mdn | SD | Min | Max | Skew | Kurt |
|---|---|---|---|---|---|---|---|
| | | | Phase 2B: Subject 2 | | | | |
| Pause | 13.7 | 12 | 8.706 | 3 | 30 | 0.993 | 0.224 |
| Latency | 5.4 | 6 | 3.627 | 1 | 12 | 0.347 | -0.562 |
| Dissociation | - | - | - | - | - | - | - |
| Fillers | 60.7 | 58.5 | 19.664 | 37 | 102 | 0.986 | 1.202 |
| Speech Errors | 11.4 | 7.5 | 10.501 | 1 | 32 | 1.099 | 0.154 |
| Laughter | 0.1 | 0 | 0.316 | 0 | 1 | 3.162 | 10 |
| Vocal Raises | 9.5 | 7 | 9.478 | 0 | 24 | 0.671 | -1.324 |
| Interruptions | 0.6 | 0 | 1.075 | 0 | 3 | 1.691 | 1.864 |
| Scoff | - | - | - | - | - | - | - |
| | | | Phase 2C: Subject 3 | | | | |
| Pause | 17 | 10 | 15.104 | 6 | 51 | 1.951 | 4.037 |
| Latency | 2 | 1 | 2.712 | 0 | 8 | 2.183 | 4.965 |
| Dissociation | - | - | - | - | - | - | - |
| Fillers | 99 | 92 | 43.494 | 55 | 174 | 0.669 | -0.686 |
| Speech Errors | 57 | 54 | 33.093 | 19 | 120 | 0.958 | 0.92 |
| Laughter | 10 | 11 | 1.768 | 7 | 12 | -0.967 | 0.522 |
| Vocal Raises | 8 | 3 | 13.435 | 0 | 40 | 2.511 | 6.604 |
| Interruptions | 11 | 8 | 9.015 | 2 | 27 | 0.904 | -0.46 |
| Scoff | | | | | | | |

Table E.8

*Descriptive Statistics for the Paralinguistics II Behaviours in Each Phase*

| Behavior | M | Mdn | SD | Min | Max | Skew | Kurt |
|---|---|---|---|---|---|---|---|
| | | | Phase 2A: Subject 1 | | | | |
| Corrections | 1.6 | 1.5 | 1.897 | 0 | 6 | 1.454 | 2.482 |
| Verbal Diversions | 1.7 | 0 | 2.791 | 0 | 7 | 1.191 | -0.437 |
| Generalizations | 5.1 | 5.5 | 3.635 | 0 | 11 | 0.125 | -1.145 |
| Lack of Memory | 5.3 | 6.5 | 3.129 | 0 | 9 | -0.973 | -0.133 |
| Recovered | 0.7 | 0 | 1.337 | 0 | 4 | 2.076 | 4.059 |
| Reasons | 1.7 | 1 | 1.494 | 0 | 4 | 0.639 | -0.992 |
| Dismissiveness | 1.5 | 0 | 2.273 | 0 | 6 | 1.348 | 0.49 |
| Humor | 3.4 | 3 | 2.503 | 0 | 7 | 0.015 | -1.333 |
| Extreme | 1.1 | 0 | 1.663 | 0 | 4 | 1.253 | -0.037 |
| Pardon Self | 0.3 | 0 | 0.949 | 0 | 3 | 3.162 | 10 |

| Behavior | M | Mdn | SD | Min | Max | Skew | Kurt |
|---|---|---|---|---|---|---|---|
| Pardon Other | 0.2 | 0 | 0.422 | 0 | 1 | 1.779 | 1.406 |
| Sarcasm | 0.2 | 0 | 0.632 | 0 | 2 | 3.162 | 10 |
| Phase 2B: Subject 2 | | | | | | | |
| Corrections | 1.5 | 2 | 1.179 | 0 | 3 | -0.255 | -1.44 |
| Verbal Diversions | 1.7 | 0 | 2.359 | 0 | 6 | 0.956 | -0.703 |
| Generalizations | 1.9 | 2 | 1.663 | 0 | 5 | 0.377 | -0.447 |
| Lack of Memory | 4.1 | 3.5 | 2.079 | 2 | 8 | 0.67 | -0.609 |
| Recovered | 1 | 0.5 | 1.155 | 0 | 3 | 0.541 | -1.393 |
| Reasons | 1.7 | 1 | 1.889 | 0 | 5 | 0.663 | -1.145 |
| Dismissiveness | 0.8 | 0 | 1.317 | 0 | 3 | 1.183 | -0.577 |
| Humor | 0.5 | 0 | 0.972 | 0 | 3 | 2.27 | 5.356 |
| Extreme | 0.1 | 0 | 0.316 | 0 | 1 | 3.162 | 10 |
| Pardon Self | - | - | - | - | - | - | - |
| Pardon Other | 0.8 | 0.5 | 1.229 | 0 | 4 | 2.261 | 5.879 |
| Sarcasm | - | - | - | - | - | - | - |
| Phase 2C: Subject 3 | | | | | | | |
| Corrections | 1 | 1 | 1.753 | 0 | 4 | 1.194 | -0.388 |
| Verbal Diversions | 7 | 6 | 6.886 | 1 | 18 | 1.058 | -0.5 |
| Generalizations | 10 | 7 | 8.084 | 2 | 24 | 0.807 | -0.875 |
| Lack of Memory | 5 | 5 | 2.138 | 1 | 7 | -0.292 | -0.905 |
| Recovered | 1 | 1 | 0.744 | 0 | 2 | 0.824 | -0.152 |
| Reasons | 1 | 1 | 0.991 | 0 | 3 | 1.486 | 2.973 |
| Dismissiveness | 3 | 2 | 2.56 | 0 | 8 | 1.374 | 2.553 |
| Humor | 2 | 1 | 1.246 | 1 | 4 | 0.895 | -1.132 |
| Extreme | 8 | 5 | 8.502 | 0 | 22 | 0.718 | -1.199 |
| Pardon Self | - | - | - | - | - | - | - |
| Pardon Other | 6 | 6 | 4.027 | 0 | 13 | 0.23 | 0.208 |
| Sarcasm | | | | | | | |

Table E.9

*Descriptive Statistics for the Incongruent Behaviours in Each Phase*

| Behavior | M | Mdn | SD | Min | Max | Skew | Kurt |
|---|---|---|---|---|---|---|---|
| Phase 2A: Subject 1 | | | | | | | |
| Full Shrug | 0.1 | 0 | 0.316 | 0 | 1 | 3.162 | 10 |
| Shrug Without | 0.4 | 0 | 0.516 | 0 | 1 | 0.484 | -2.277 |

| Behavior | M | Mdn | SD | Min | Max | Skew | Kurt |
|---|---|---|---|---|---|---|---|
| Asymmetric Shrug | 0.1 | 0 | 0.316 | 0 | 1 | 3.162 | 10 |
| Head Nod | 0.4 | 0 | 0.516 | 0 | 1 | 0.484 | -2.277 |
| Head Shake | 5.4 | 6 | 2.319 | 1 | 8 | -0.944 | 0.24 |
| Humor | 0.8 | 0.5 | 0.919 | 0 | 2 | 0.473 | -1.807 |
| Hand Shrug | 1.1 | 0 | 1.729 | 0 | 5 | 1.584 | 1.862 |
| Hand Stop | 0.4 | 0 | 0.516 | 0 | 1 | 0.484 | -2.277 |
| Laughter | 7.2 | 6 | 7.177 | 0 | 21 | 0.789 | -0.372 |
| Head Down and | 0.2 | 0 | 0.422 | 0 | 1 | 1.779 | 1.406 |
| Hand Dismiss | 0.6 | 0 | 0.966 | 0 | 3 | 1.959 | 4.187 |
| Obscene Gesture | - | - | - | - | - | - | - |
| Phase 2B: Subject 2 | | | | | | | |
| Full Shrug | - | - | - | - | - | - | - |
| Shrug Without | 0.1 | 0 | 0.316 | 0 | 1 | 3.162 | 10 |
| Asymmetric Shrug | - | - | - | - | - | - | - |
| Head Nod | 0.1 | 0 | 0.316 | 0 | 1 | 3.162 | 10 |
| Head Shake | 6.9 | 6 | 6.027 | 0 | 16 | 0.374 | -1.494 |
| Humor | - | - | - | - | - | - | - |
| Hand Shrug | 0.3 | 0 | 0.483 | 0 | 1 | 1.035 | -1.224 |
| Hand Stop | 0.1 | 0 | 0.316 | 0 | 1 | 3.162 | 10 |
| Laughter | - | - | - | - | - | - | - |
| Head Down and | - | - | - | - | - | - | - |
| Hand Dismiss | - | - | - | - | - | - | - |
| Obscene Gesture | - | - | - | - | - | - | - |
| Phase 2C: Subject 3 | | | | | | | |
| Full Shrug | 1 | 2 | 1.302 | 0 | 3 | 0.105 | -1.922 |
| Shrug Without | 2 | 2 | 2.07 | 0 | 6 | 1.159 | 0.812 |
| Asymmetric Shrug | 3 | 2 | 3.926 | 0 | 10 | 0.835 | -0.84 |
| Head Nod | 3 | 3 | 3.227 | 0 | 9 | 0.701 | -0.141 |
| Head Shake | 13 | 13 | 8.28 | 0 | 26 | 0.17 | -0.102 |
| Humor | 1 | 1 | 1.488 | 0 | 3 | 0.477 | -2.249 |
| Hand Shrug | 8 | 9 | 4.853 | 0 | 17 | 0.219 | 1.708 |
| Hand Stop | | | | | | | |
| Laughter | 8 | 9 | 2.825 | 3 | 12 | -0.763 | 0.853 |
| Head Down and | | | | | | | |
| Hand Dismiss | | | | | | | |
| Obscene Gesture | | | | | | | |

Table E.10

*Descriptive Statistics for the Duration Behaviours in Each Phase*

| Behavior | M | Mdn | SD | Min | Max | Skew | Kurt |
|---|---|---|---|---|---|---|---|
| | | | Phase 2A: Subject 1 | | | | |
| Self Soothe | 291.8 | 238.5 | 202.467 | 21 | 733 | 1.041 | 1.508 |
| Protective Object | 9.1 | 0 | 28.777 | 0 | 91 | 3.162 | 10 |
| Arms Crossed | | | | | | | |
| One Arm Crossed | 101 | 5 | 212.277 | 0 | 674 | 2.648 | 7.328 |
| Two Arms | 76.1 | 20 | 134.973 | 0 | 435 | 2.508 | 6.674 |
| Legs Crossed | | | | | | | |
| Leg Open Stance | | | | | | | |
| Leg Resting on | | | | | | | |
| Tucked on Seat | | | | | | | |
| Knees Up | | | | | | | |
| Leg Other | | | | | | | |
| Pause | 179.3 | 168.5 | 79.089 | 68 | 376 | 1.693 | 4.941 |
| Latency | 52.6 | 15 | 68.998 | 2 | 188 | 1.194 | -0.165 |
| Dissociation | - | - | - | - | - | - | - |
| | | | Phase 2B: Subject 2 | | | | |
| Self Soothe | 20.3 | 18.5 | 11.166 | 0 | 44 | 0.511 | 2.545 |
| Protective Object | - | - | - | - | - | - | - |
| Arms Crossed | | | | | | | |
| One Arm Crossed | - | - | - | - | - | - | - |
| Two Arms | - | - | - | - | - | - | - |
| Legs Crossed | | | | | | | |
| Leg Open Stance | | | | | | | |
| Leg Resting on | | | | | | | |
| Tucked on Seat | | | | | | | |
| Knees Up | | | | | | | |
| Leg Other | | | | | | | |
| Pause | 52.3 | 52.5 | 21.97 | 12 | 86 | -0.38 | -0.125 |
| Latency | 32.4 | 32 | 26.613 | 2 | 81 | 0.531 | -0.57 |
| Dissociation | - | - | - | - | - | - | - |
| | | | Phase 2C: Subject 3 | | | | |
| Self Soothe | 20 | 0 | 54.749 | 0 | 155.16 | 2.827 | 7.995 |
| Protective Object | 5 | 1 | 6.345 | 0 | 16.997 | 1.245 | 0.576 |

| Behavior | M | Mdn | SD | Min | Max | Skew | Kurt |
|---|---|---|---|---|---|---|---|
| Arms Crossed | 154 | 172 | 56.224 | 17.109 | 194 | -2.632 | 7.205 |
| One Arm Crossed | | | | | | | |
| Two Arms | | | | | | | |
| Legs Crossed | 1126 | 1315 | 460.638 | 0 | 1319 | -2.705 | 7.419 |
| Leg Open Stance | 26 | 1 | 71.443 | 0 | 202.99 | 2.828 | 7.997 |
| Leg Resting on | 164 | 0 | 465.142 | 0 | 1315.62 | 2.828 | 8 |
| Tucked on Seat | - | - | - | - | - | - | - |
| Knees Up | - | - | - | - | - | - | - |
| Leg Other | - | - | - | - | - | - | - |
| Pause | 86 | 72 | 45.874 | 50 | 192.754 | 2.246 | 5.398 |
| Latency | 12 | 15 | 10.429 | 0 | 27.482 | -0.045 | -1.373 |
| Dissociation | - | - | - | - | - | - | - |

Table E.11

*Descriptive Statistics for the Affect I Behaviours in Phase 2C*

| Behavior | M | Mdn | SD | Min | Max | Skew | Kurt |
|---|---|---|---|---|---|---|---|
| Micro Expression | | | | | | | |
| Happiness | 3 | 1 | 5.167 | 0 | 15 | 2.349 | 5.624 |
| Sadness | 2 | 1 | 2.642 | 0 | 8 | 1.865 | 3.807 |
| Anger | 3 | 3 | 3.044 | 0 | 8 | 0.522 | -1.11 |
| Fear | 1 | 1 | 2.387 | 0 | 7 | 2.35 | 5.831 |
| Surprise | 2 | 1 | 4.027 | 0 | 12 | 2.593 | 6.992 |
| Contempt | 3 | 3 | 3.012 | 0 | 8 | 0.559 | -0.917 |
| Disgust | 3 | 1 | 3.703 | 0 | 10 | 1.238 | 0.328 |
| ME Mask | | | | | | | |
| Happiness | 2 | 1 | 2.726 | 0 | 7 | 1.129 | -0.104 |
| Sadness | 0 | 0 | 0.744 | 0 | 2 | 1.951 | 3.205 |
| Anger | 0 | 0 | 0.354 | 0 | 1 | 2.828 | 8 |
| Fear | 0 | 0 | 0.463 | 0 | 1 | 1.44 | 0 |
| Surprise | - | - | - | - | - | - | - |
| Contempt | 0 | 0 | 0.354 | 0 | 1 | 2.828 | 8 |
| Disgust | 0 | 0 | 0.354 | 0 | 1 | 2.828 | 8 |
| Neutral | 15 | 12 | 11.285 | 1 | 38 | 1.249 | 2.165 |

Table E.12

*Descriptive Statistics for the Affect II Behaviours in Phase 2C*

| Behavior | M | Mdn | SD | Min | Max | Skew | Kurt |
|---|---|---|---|---|---|---|---|
| Incongruent | | | | | | | |
| Happiness | 4 | 4 | 2.878 | 0 | 7 | -0.096 | -1.681 |
| Sadness | 0 | 0 | 0.463 | 0 | 1 | 1.44 | 0 |
| Anger | 0 | 0 | 0.463 | 0 | 1 | 1.44 | 0 |
| Fear | 0 | 0 | 0.354 | 0 | 1 | 2.828 | 8 |
| Surprise | 0 | 0 | 0.518 | 0 | 1 | 0.644 | -2.24 |
| Contempt | 1 | 0 | 1.061 | 0 | 3 | 1.96 | 3.937 |
| Disgust | 0 | 0 | 0.744 | 0 | 2 | 1.951 | 3.205 |
| Neutral | - | - | - | - | - | - | - |
| Expected | | | | | | | |
| Happiness | 1 | 1 | 0.886 | 0 | 2 | 0.615 | -1.481 |
| Sadness | 1 | 1 | 1.356 | 0 | 4 | 1.539 | 2.571 |
| Anger | 1 | 0 | 1.188 | 0 | 3 | 1.652 | 1.355 |
| Fear | 0 | 0 | 0.354 | 0 | 1 | 2.828 | 8 |
| Surprise | - | - | - | - | - | - | - |
| Contempt | 1 | 0 | 1.069 | 0 | 3 | 2.339 | 5.469 |
| Disgust | - | - | - | - | - | - | - |
| Neutral | 2 | 1 | 2.8 | 0 | 8 | 1.784 | 3.249 |

# Appendix E

Individual Coder Reliability Estimates for Behavioral Categories Across Three Subjects

Table F.1

*Individual Coder Reliability Estimates for Behavioural Categories Across Three Subjects*

| | Coder 1 | Coder 2 | Coder 3 | Coder 4 | Coder 5 | Coder 6 | Coder 7 | Coder 8 |
|---|---|---|---|---|---|---|---|---|
| Hands | | | | | | | | |
| Subject 1 | 0.875 | 0.825 | 0.812 | 0.757 | 0.773 | 0.881 | 0.858 | 0.884 |
| Subject 2 | 0.745 | 0.731 | 0.670 | 0.728 | 0.747 | 0.742 | 0.801 | 0.797 |
| Subject 3 | 0.937 | 0.889 | 0.912 | 0.890 | 0.821 | 0.877 | 0.854 | 0.874 |
| Body | | | | | | | | |
| Subject 1 | -0.030 | 0.329 | 0.281 | 0.243 | 0.273 | 0.305 | -0.010 | 0.133 |
| Subject 2 | 0.025 | 0.195 | --- | -0.036 | 0.241 | -0.114 | 0.045 | 0.143 |
| Subject 3 | 0.926 | 0.878 | 0.863 | 0.933 | 0.893 | 0.869 | 0.662 | 0.858 |
| Limbs | | | | | | | | |
| Subject 1 | -0.082 | 0.340 | -0.012 | 0.274 | 0.242 | 0.290 | 0.207 | 0.107 |
| Subject 2 | 0.543 | NA | 0.543 | 0.543 | 0.543 | 0.543 | 0.457 | --- |
| Subject 3 | 0.679 | 0.614 | 0.733 | 0.072 | 0.689 | 0.602 | 0.689 | 0.113 |
| H&S | | | | | | | | |
| Subject 1 | 0.802 | 0.718 | 0.824 | 0.742 | 0.684 | 0.802 | 0.817 | 0.790 |
| Subject 2 | 0.891 | 0.841 | 0.925 | 0.881 | 0.911 | 0.880 | 0.901 | 0.804 |
| Subject 3 | 0.697 | 0.351 | 0.436 | 0.338 | 0.714 | 0.521 | 0.585 | 0.649 |
| Eyes and Face | | | | | | | | |
| Subject 1 | 0.709 | 0.933 | 0.916 | 0.944 | 0.957 | 0.966 | 0.964 | 0.968 |
| Subject 2 | 0.714 | 0.970 | 0.970 | 0.978 | 0.976 | 0.986 | 0.970 | 0.981 |
| Subject 3 | 0.790 | 0.887 | 0.834 | 0.807 | 0.820 | 0.875 | 0.865 | 0.882 |
| Reflexive | | | | | | | | |
| Subject 1 | 0.283 | 0.404 | 0.179 | 0.309 | 0.596 | 0.377 | 0.488 | 0.603 |
| Subject 2 | -0.055 | 0.647 | 0.295 | 0.615 | 0.640 | 0.367 | 0.402 | 0.463 |
| Subject 3 | 0.745 | 0.763 | 0.606 | 0.667 | 0.663 | 0.602 | -0.036 | 0.455 |
| Paralinguistic I | | | | | | | | |
| Subject 1 | 0.815 | 0.524 | 0.801 | 0.834 | 0.749 | 0.746 | 0.822 | 0.847 |
| Subject 2 | 0.904 | 0.731 | 0.834 | 0.848 | 0.910 | 0.910 | 0.823 | 0.903 |
| Subject 3 | 0.790 | 0.643 | 0.752 | 0.526 | 0.728 | 0.672 | 0.621 | 0.731 |
| Paralinguistic II | | | | | | | | |
| Subject 1 | 0.559 | 0.283 | -0.198 | 0.529 | 0.237 | 0.286 | 0.558 | 0.372 |
| Subject 2 | 0.484 | 0.056 | 0.166 | 0.351 | 0.385 | 0.323 | 0.354 | 0.393 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Subject 3 | 0.305 | 0.191 | 0.200 | 0.414 | 0.320 | 0.256 | 0.132 | 0.361 |
| Incongruence | | | | | | | | |
| Subject 1 | 0.279 | 0.360 | 0.457 | 0.432 | 0.403 | 0.335 | 0.349 | 0.312 |
| Subject 2 | 0.494 | 0.177 | 0.408 | 0.177 | 0.541 | 0.500 | 0.278 | -0.013 |
| Subject 3 | 0.591 | 0.584 | 0.526 | 0.629 | 0.586 | 0.624 | -0.192 | 0.387 |
| Duration | | | | | | | | |
| Subject 1 | 0.572 | 0.595 | 0.539 | 0.301 | -0.049 | 0.485 | 0.473 | 0.602 |
| Subject 2 | 0.646 | 0.704 | 0.724 | 0.838 | 0.713 | 0.674 | 0.833 | 0.838 |
| Subject 3 | 0.990 | 0.989 | 0.984 | 0.986 | 0.990 | 0.990 | 0.960 | -0.064 |
| Affect I | | | | | | | | |
| Subject 1 | | | | | | | | |
| Subject 2 | | | | | | | | |
| Subject 3 | 0.456 | -0.007 | 0.570 | 0.621 | 0.666 | 0.576 | 0.586 | 0.340 |
| Affect II | | | | | | | | |
| Subject 1 | | | | | | | | |
| Subject 2 | | | | | | | | |
| Subject 3 | 0.378 | 0.265 | 0.278 | 0.410 | 0.177 | -0.094 | -0.139 | 0.093 |

## Appendix F

Literature to Support Selection of Relevant Behaviors

Abe, N., Suzuki, M., Tsukiura, T., Mori, E., Yamaguchi, K., Itoh, M., & Fujii, T. (2006). Dissociable roles of prefrontal and anterior cingulate cortices in deception. Cerebral Cortex, 16(2), 192-199. doi:http://dx.doi.org.ezproxy.library.yorku.ca/10.1093/cercor/bhi097

Akehurst, L., & Vrij, A. (1999). Creating suspects in police interviews. Journal of Applied Social Psychology, 29, 192–210.

Allen, J. J., Iacono, W. G., & Danielson, K. D. (1992). The identification of concealed memories using the event-related potential and implicit behavioral measures: A methodology for prediction in the face of individual differences. Psychophysiology, 29(5), 504-522. doi:http://dx.doi.org.ezproxy.library.yorku.ca/10.1111/j.1469-8986.1992.tb02024.x

Alonso-Quecuty, M. (1992). Deception detection and reality monitoring: A new answer to an old question? In F. Losel, D. Bender, & T. Bliesener (Eds.), Psychology and law: International perspectives (pp. 328–332). New York: de Gruyter.

Anolli, L., & Ciceri, R. (1997). The voice of deception: Vocal strategies of naive and able liars. Journal of Nonverbal Behavior, 21(4), 259-284. doi:http://dx.doi.org.ezproxy.library.yorku.ca/10.1023/A:1024916214403

Anolli, L., Balconi, M., & Ciceri, R. (2003). Linguistic styles in deceptive communication: Dubitative ambiguity and elliptic eluding in packaged lies. Social Behavior and Personality, 31(7), 687-710.

Bello, R. (2006). Causes and paralinguistic correlates of interpersonal equivocation. Journal of Pragmatics, 38(9), 1430-1441. doi:http://dx.doi.org.ezproxy.library.yorku.ca/10.1016/j.pragma.2005.09.001

Bond, C. F., Kahler, K. N., & Paolicelli, L. M. (1985). The miscommunication of deception: An adaptive perspective. Journal of Experimental Social Psychology, 21(4), 331-345. doi:http://dx.doi.org.ezproxy.library.yorku.ca/10.1016/0022-1031(85)90034-4

Bond, C. F., Omar, A., Mahmoud, A., & Bonser, R. N. (1990). Lie detection across cultures. Journal of Nonverbal Behavior, 14(3), 189-204. doi:http://dx.doi.org.ezproxy.library.yorku.ca/10.1007/BF00996226

Buller, D. B., & Aune, R. K. (1987). Nonverbal cues to deception among intimates, friends, and strangers. Journal of Nonverbal Behavior, 11, 269–290.

Buller, D. B., Burgoon, J. K., Buslig, A. L. S., & Roiger, J. F. (1994). Interpersonal deception: VIII. further analysis of nonverbal and verbal correlates of equivocation from the bavelas et al. (1990) research. Journal of Language and Social Psychology, 13(4), 396-417.

Buller, D. B., Burgoon, J. K., Buslig, A., & Roiger, J. (1996). Testing interpersonal deception theory: The language of interpersonal deception. Communication Theory, 6, 268–289.

Buller, D. B., Comstock, J., Aune, R. K., & Strzyzewski, K. D. (1989). The effect of probing on deceivers and truthtellers. Journal of Nonverbal Behavior, 13, 155–170.

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.

Burgoon, J. K., Buller, D. B., Afifi, W., White, C., & Buslig, A. (1996, May). The role of immediacy in deceptive interpersonal interactions. Paper presented at the annual meeting of the International Communication Association, Chicago, IL.

Burgoon, J. K., Buller, D. B., Floyd, K., & Grandpre, J. (1996). Deceptive realities: Sender, receiver, and observer perspectives in deceptive conversations. Communication Research, 23, 724–748.

Burgoon, J. K., Buller, D. B., Guerrero, L. K., Afifi, W. A., & Feldman, C. M. (1996). Interpersonal deception: XII. Information management dimensions underlying deceptive and truthful messages. Communication Monographs, 63, 50–69.

Burns, J. A., & Kintz, B. L. (1976). Eye contact while lying during an interview. Bulletin of the Psychonomic Society, 7, 87–89.

Caso, L., Maricchiolo, F., Bonaiuto, M., Vrij, A., & Mann, S. (2006a). The impact of deception and suspicion on different hand movements. Journal of Nonverbal Behavior, 30, 1–19.

Caso, L., Vrij, A., Mann, S., & DeLeo, G. (2006b). Deceptive responses: The impact of verbal and nonverbal countermeasures. Legal and Criminological Psychology, 11, 99–111.

Chiba, H. (1985). Analysis of controlling facial expression when experiencing negative affect on an anatomical basis. Journal of Human Development, 21, 22–29.

Christensen, D. (1980). Decoding of intended versus unintended nonverbal messages as a function of social skill and anxiety. Unpublished doctoral dissertation, University of Connecticut, Storrs.

Ciofu, I. (1974). Audiospectral analysis in lie detection. Archiv fur Psychologie, 126, 170–180.

Cody, M. J., & O'Hair, H. D. (1983). Nonverbal communication and deception: Differences in deception cues due to gender and communicator dominance. Communication Monographs, 50, 175–192.

Cody, M. J., Lee, W. S., & Chao, E. Y. (1989). Telling lies: Correlates of deception among Chinese. In J. P. Forgas & J. M. Innes (Eds.), Recent advances in social psychology: An international perspective (pp. 359– 368). Amsterdam: North-Holland.

Cody, M. J., Marston, P. J., & Foster, M. (1984). Deception: Paralinguistic and verbal leakage. In R. N. Bostrom & B. H. Westley (Eds.), Communication yearbook 8 (pp. 464–490). Beverly Hills, CA: Sage.

Craig, K. D., Hyde, S. A., & Patrick, C. J. (1991). Genuine, suppressed and faked facial behavior during exacerbation of chronic low back pain. Pain, 46, 161–171.

Cutrow, R. J., Parks, A., Lucas, N., & Thomas, K. (1972). The objective use of multiple physiological indices in the detection of deception. Psychophysiology, 9, 578–588.

Darwin, C., & Prodger, P. (1998). The expression of the emotions in man and animals. Oxford University Press, USA.

Davatzikos, C., Ruparel, K., Fan, Y., Shen, D. G., Acharyya, M., Loughead, J. W., Gur, R. C., & Langleben, D. D. (2005). Classifying spatial patterns of brain activity with machine learning methods: Application to lie detection. NeuroImage, 28, 663–668.

Davis, M., & Hadiks, D. (1995). Demeanor and credibility. Semiotica, 106, 5– 54.

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.

Dente, E., Barath, A. A., Ng, J., Vrij, A., Mann, S., & Bull, A. (2006). Tracking hand and finger movements for behaviour analysis. Pattern Recognition Letter, 27, 1797–1808.

DePaulo, B. M., Blank, A. L., Swain, G. W., & Hairfield, J. G. (1992). Expressiveness and expressive control. Personality and Social Psychology Bulletin, 18, 276–285.

DePaulo, B. M., Epstein, J. A., & LeMay, C. S. (1990). Responses of the socially anxious to the prospect of interpersonal evaluation. Journal of Personality, 58, 623–640.

DePaulo, B. M., Lanier, K., & Davis, T. (1983). Detecting the deceit of the motivated liar. Journal of Personality and Social Psychology, 45, 1096–1103.

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological bulletin*, *129*(1), 74.

DePaulo, B. M., Rosenthal, R., Rosenkrantz, & Green, C. R. (1982b). Actual and perceived cues to deception: A closer look at speech. Basic and Applied Social Psychology, 3, 291–312.

DePaulo, P. J., & DePaulo, B. M. (1989). Can deception by salespersons and customers be detected through nonverbal behavioral cues? Journal of Applied Social Psychology, 19, 1552–1577.

deTurck, M. A., & Miller, G. R. (1985). Deception and arousal: Isolating the behavioral correlates of deception. Human Communication Research, 12, 181–201

Driver, J. & Aalst, M. (2012). *You can't lie to me : the revolutionary program to supercharge your inner lie detector and get to the truth*. New York: HarperOne.

Dulaney, E. F. (1982). Changes in language behavior as a function of veracity. Human Communication Research, 9, 75–82.

Ebesu, A. S., & Miller, M. D. (1994). Verbal and nonverbal behaviors as a function of deception type. Journal of Language and Social Psychology, 13, 418–442.

Ekman, P. & Friesen, W. V. (1969). Nonverbal Leakage and Clues to Deception. Psychiatry, 32(1), 88-106.

Ekman, P. & Friesen, W. V. (1969). The Repertoire of Nonverbal Behavior: Categories, Origins, Usage, and Coding. Semiotica, 1(1), 49-98.

Ekman, P. & Rosenberg, E. (2005). *What the face reveals: basic and applied studies of spontaneous expression using the facial action coding system (FACS*. Oxford New York: Oxford University Press.

Ekman, P. (1988). Lying and nonverbal behavior: Theoretical issues and new findings. Journal of Nonverbal Behavior, 12, 163–176.

Ekman, P. (2007). *Emotions revealed : recognizing faces and feelings to improve communication and emotional life*. New York: Owl Books.

Ekman, P. (2009). *Telling lies : clues to deceit in the marketplace, politics, and marriage*. New York, NY: W.W. Norton.

Ekman, P., & Friesen, W. V. (1972). Hand movements. *Journal of communication, 22*(4), 353-374.

Ekman, P., & O'Sullivan, M. (1991). Who can catch a liar? American Psychologist, 46, 913–920.

Ekman, P., Friesen, W. V., & O'Sullivan, M. (1988). Smiles when lying. Journal of Personality and Social Psychology, 54, 414–420.

Ekman, P., Friesen, W. V., & Scherer, K. R. (1976). Body movement and voice pitch in deceptive interaction. Semiotica, 16, 23–27.

Ekman, P., Friesen, W. V., & Simons, R. C. (1985). Is the startle reaction an emotion? Journal of Personality and Social Psychology, 49, 1416– 1426.

Ekman, P., O'Sullivan, M., Friesen, W. V., & Scherer, K. R. (1991). Face, voice, and body in detecting deceit. Journal of Nonverbal Behavior, 15, 125–135.

Elliott, G. L. (1979). Some effects of deception and level of selfmonitoring on planning and reacting to a self-presentation. Journal of Personality and Social Psychology, 37, 1282–1292.

Engelhard, I. M., Merckelbach, H., & van den Hout, M. A. (2003). The guilty knowledge test and the modified stroop task in detection of deception: An exploratory study. Psychological Reports, 92, 683–691.

Exline, R. V., Thibaut, J., Hickey, C. B., & Gumpert, P. (1970). Visual interaction in relation to Machiavellianism and an unethical act. In R. Christie & F. Geis (Eds.), Studies in Machiavellianism (pp. 53–75). New York: Academic Press.

Farrow, T. F. D., Reilly, R., Rahman, T. A., Herford, A. E., Woodruff, P. W. R., & Spence, S. A. (2003). Sex and personality traits influence the difference between time taken to tell the truth or lie. Perceptual and Motor Skills, 97, 451–460.

Feeley, T. H., & deTurck, M. A. (1998). The behavioral correlates of sanctioned and unsanctioned deceptive communication. Journal of Nonverbal Behavior, 22, 189–204.

Fiedler, K., & Walka, I. (1993). Training lie detectors to use nonverbal cues instead of global heuristics. Human Communication Research, 20, 199–223.

Fiedler, K., Schmid, J., Kurzenhauser, S., & Schroter, V. (1997). Lie detection as an attribution process: The anchoring effect revisited. Unpublished manuscript.

Finkelstein, S. (1978). The relationship between physical attractiveness and nonverbal behaviors. Unpublished honors thesis, Hampshire College, Amherst, MA.

Frank, M. G. (1989). Human lie detection ability as a function of the liar's motivation. Unpublished doctoral dissertation, Cornell University, Ithaca, NY.

Fukuda, K. (2001). Eye blinks: New indices for the detection of deception. International Journal of Psychophysiology, 40, 239–245.

Gagnon, L. R. (1975). The encoding and decoding of cues to deception. Unpublished doctoral dissertation, Arizona State University, Tempe.

Galin, K. E., & Thorn, B. E. (1993). Unmasking pain: Detection of deception in facial expressions. Journal of Social and Clinical Psychology, 12, 182–197.

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

Gladwell, M. (2007). *Blink: The power of thinking without thinking*. Back Bay Books.

Goldstein, E. R. (1923). Reaction times and the consciousness of deception. American Journal of Psychology, 34, 562–581.

Gozna, L., & Babooram, N. (2004). Non-traditional interviews: Deception in a simulated customs baggage search. In A. Czerederecka, T. Jaskiewicz- Obydzinska, R. Roesch, & J. Wojcikiewicz (Eds,), Forensic psychology and law (pp. 153–161). Krakow, Poland: Institute of Forensic Research Publishers.

Granhag, P. A., & Stromwall, L. A. (2002). Repeated interrogations: Verbal and nonverbal cues to deception. Applied Cognitive Psychology, 16, 243–257.

Greene, J. O., O'Hair, H. D., Cody, M. J., & Yen, C. (1985). Planning and control of behavior during deception. Human Communication Research, 11, 335–364.

Gregg, A. P. (2007). When vying reveals lying: The timed antagonistic response alethiometer. Applied Cognitive Psychology, 21, 621–648

Gross, J. J., & Levenson, R. W. (1993). Emotional suppression: Physiology, self-report, and expressive behavior. Journal of Personality and Social Psychology, 64, 970–986.4

Hadjistavropoulos, H. D., & Craig, K. D. (1994). Acute and chronic low back pain: Cognitive, affective, and behavioral dimensions. Journal of Consulting and Clinical Psychology, 62, 341–349.

Hall, M. E. (1986). Detecting deception in the voice: An analysis of fundamental frequency, syllabic duration, and amplitude of the human voice. Unpublished doctoral dissertation, Michigan State University, East Lansing.

Harrison, A. A., Hwalek, M., Raney, D. F., & Fritz, J. G. (1978). Cues to deception in an interview situation. Social Psychology, 41, 156–161.

Hartley, G. & Karinch, M. (2007). *I can read you like a book : how to spot the messages and emotions people are really sending with their body language*. Franklin Lakes, NJ: Career Press.

Heilveil, I., & Muehleman, J. T. (1981). Nonverbal clues to deception in a psychotherapy analogue. Psychotherapy: Theory, Research and Practice, 18, 329–335.

Hemsley, G. D. (1977). Experimental studies in the behavioral indicants of deception. Unpublished doctoral dissertation, University of Toronto, Toronto, Ontario, Canada.

Hernandez-Fernaud, E., & Alonso-Quecuty, M. (1997). The cognitive interview and lie detection: A new magnifying glass for Sherlock Holmes? Applied Cognitive Psychology, 11, 55–68.

Hess, U. (1989). On the dynamics of emotional facial expression. Unpublished doctoral dissertation, Dartmouth College, Hanover, NH.

Hess, U., & Kleck, R. E. (1994). The cues decoders use in attempting to differentiate emotion-elicited and posed facial expressions. European Journal of Social Psychology, 24, 367–381.

Hill, M. L., & Craig, K. D. (2002). Detecting deception in pain expressions: The structure of genuine and deceptive facial displays. Pain, 98, 135–144.

Hocking, J. E., & Leathers, D. G. (1980). Nonverbal indicators of deception: A new theoretical perspective. Communication Monographs, 47, 119–131.

Hofer, E., K¨ohnken, G., Hanewinkel, R., & Bruhn, C. (1993). Diagnostik und attribution von glaubw¨ urdigkeit. Kiel: final report to the Deutsche Forschungsgemeinschaft, KO 882/4-2.

Horvath, F. S. (1973). Verbal and nonverbal clues to truth and deception during polygraph examinations. Journal of Police Science and Administration, 1, 138–152.

Horvath, F., Jayne, B., & Buckley, J. (1994). Differentiation of truthful and deceptive criminal suspects in behavior analysis interviews. Journal of Forensic Sciences, 39, 793–807.

Johnson, R., Barnhardt, J., & Zhu, J. (2003). The deceptive response: Effects of response conflict and strategic monitoring on the late positive component and episodic memory-related brain activity. Biological Psychology, 64, 217– 253.

Kalma, A., Witte, M., & Zaalberg, R. (1996). Authenticity: operationalization, manipulation, and behavioural components: An exploration. Medium Psychologie, 8, 49–65.

Kennedy, J., & Coe, W. C. (1994). Nonverbal signs of deception during posthypnotic amnesia. The International Journal of Clinical and Experimental Hypnosis, 17, 13–19.

Klaver, J. R., Lee, Z., & Hart, S. D. (2007). Psychopathy and nonverbal indicators of deception in offenders. Law and Human Behavior, 31, 337–351.

Knapp, M. L., Hart, R. P., & Dennis, H. S. (1974). An exploration of deception as a communication construct. Human Communication Research, 1, 15–29.

Kohnken, G. (1985). Speech and deception of eyewitnesses: An information processing approach. In F. L. Denmark (Ed.), The psychology of women (pp. 117–139), North Holland, the Netherlands: Elsevier Science Publishers.

Kohnken, G., Schimossek, E., Aschermann, E., & Hofer, E. (1995). The cognitive interview and the assessment of the credibility of adults' statements. Journal of Applied Psychology, 80, 671–684.

Koper, R. J., & Sahlman, J. M. (2001). The behavioral correlates of naturally occurring, high motivation deceptive communication. Manuscript submitted for publication.

Krauss, R. M. (1981). Impression formation, impression management, and nonverbal behaviors. In E. T. Higgins, C. P. Herman, & M. P. Zanna (Eds.), Social cognition: Vol. 1. The Ontario Symposium (pp. 323–341). Hillsdale, NJ: Erlbaum.

Kraut, R. E. (1978). Verbal and nonverbal cues in the perception of lying. Journal of Personality and Social Psychology, 36, 380–391

Kraut, R. E., & Poe, D. (1980). Behavioral roots of person perception: The deception judgments of customs inspectors and laymen. Journal of Personality and Social Psychology, 39, 784–798.

Landry, K. L., & Brigham, J. C. (1992). The effect of training in Criteria-Based Content Analysis on the ability to detect deception in adults. Law and Human Behavior, 16, 663–676.

Langleben, D. D., Loughead, J. W., Bilker, W. B., Ruparel, K., Childress, A. R., Busch, S. I., & Gur, R. C. (2005). Telling the truth from lie in individual subjects with fast event-related fMRI. Human Brain Mapping, 26, 262–272.

Leal, S., Vrij, A., Fisher, R., & van Hooff (2006). Cognitively induced arousal suppression when lying. Manuscript submitted for publication.

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 training control group. Western Journal of Communication, 69, 203–217.

Mak, E. G., Y., & Lee, T. M. C. (2006). Detection of feigned memory impairments using a Chinese word task. Psychological Reports, 98, 779–788.

Manaugh, T. S., Wiens, A. N., & Matarazzo, J. D. (1970). Content saliency and interviewee speech behavior. Journal of Clinical Psychology, 26, 17–24.

Mann, S., Vrij, A., & Bull, R. (2002). Suspects, lies and videotape: An analysis of authentic high-stakes liars. Law and Human Behavior, 26, 365–376.

Marston, W. M. (1920). Reaction-time symptoms of deception. Journal of Experimental Psychology, 3, 72–87.

Matarazzo, J. D., Wiens, A. N., Jackson, R. H., & Manaugh, T. S. (1970). Interviewee speech behavior under conditions of endogenously present and exogenously-induced motivational states. Journal of Clinical Psychology, 26, 141–148.

McClintock, C. C., & Hunt, R. G. (1975). Nonverbal indicators of affect and deception in an interview setting. Journal of Applied Social Psychology, 5, 54–67.

Mehrabian, A. (1971). Nonverbal betraying of feeling. Journal of Experimental Research in Personality, 5, 64–73.

Mehrabian, A. (1972). Nonverbal communication. Chicago, IL: Aldine-Atherton. Meijer, E. (2004). Leugendetectie en wetenschap: Vriend of vijand? De Psycholoog, 39, 174–179.

Miller, G. R., deTurck, M. A., & Kalbfleisch, P. J. (1983). Self-Monitoring, rehearsal, and deceptive communication. Human Communication Research, 10, 97–117.

Morris, D. (1994). *Bodytalk : a world guide to gestures*. London: Cape.

Morris, D. (2002). Peoplewatching. London: Vintage.

Motley, M. T. (1974). Acoustic correlates of lies. Western Speech, 38, 81–87.

Navarro, J. & Karlins, M. (2008). *What every BODY is saying: an ex-FBI agent's guide to speed reading people*. New York, NY: Collins.

Nunez, J. M., Casey, B. J., Egner, T., Hare, T., & Hirsch, J. (2005). Intentional false responding shares neutral substrates with response conflict and cognitive control. NeuroImage, 25, 267–277.

O'Hair, H. D., Cody, M., & McLaughlin, M. L. (1981). Prepared lies, spontaneous lies, Machiavellianism and nonverbal communication. Human Communication Research, 7, 325–339.

Parliament, L., & Yarmey, A. D. (2002). Deception in eyewitness identification. Criminal Justice and Behavior, 29, 734–746.

Pennebaker, J. W., & Chew, C. H. (1985). Behavioral inhibition and electrodermal activity during deception. Journal of Personality and Social Psychology, 49, 1427–1433.

Porter, S., & Yuille, J. C. (1996). The language of deceit: An investigation of the verbal clues to deception in the interrogation context. Law and Human Behavior, 20, 443–459.

Porter, S., Doucette, N., Earle, J., & MacNeil, B. (2008). Half the world knows not how the other half lies': Investigation of cues to deception exhibited by criminal offenders and non-offenders. *Legal and Criminological Psychology, 13*, 27-38.

Potamkin, G. G. (1982). Heroin addicts and nonaddicts: The use and detection of nonverbal deception cues. Unpublished doctoral dissertation, California School of Professional Psychology, Los Angeles.

Riggio, R. E., & Friedman, H. S. (1983). Individual differences and cues to deception. Journal of Personality and Social Psychology, 45, 899–915.

Rockwell, P., Buller, D. B., & Burgoon, J. K. (1997). Measurement of deceptive voices: Comparing acoustic and perceptual data. Applied Psycholinguistics, 18, 471–484.

Rosenfeld, P. J., Biroschak, J. R., & Furedy, J. J. (2006). P300-based detection of concealed autobiographical versus incidentally acquired information in target and non-target paradigms. International Journal of Psychophysiology, 60, 251–259.

Ruby, C. L., & Brigham, J. C. (1998). Can Criteria-Based Content Analysis distinguish between true and false statements of African- American speakers? Law and Human Behavior, 22, 369–388.

Rybold, V. S. (1994). Paralinguistic cue leakage during deception: A comparison between Asians and Euroamericans. Unpublished master's thesis, California State University, Fullerton. SAS Institute. (1985). SAS user's guide: Statistics (Version 5). Cary, NC: Author.

Sayenga, E. R. (1983). Linguistic and paralinguistic indices of deception. Unpublished doctoral dissertation, University of Michigan, Ann Arbor.

Scheflen, A. E. (1964). THE SIGNIFICANCE OF POSTURE IN COMMUNICATION SYSTEMS. Psychiatry, 27(4), 316-331.

Schneider, S. M., & Kintz, B. L. (1977). The effect of lying upon foot and leg movement. Bulletin of the Psychonomic Society, 10, 451–453.

Seymour, T. L., Seifert, C. M., Shafto, M. G., & Mosmann, A. L. (2000). Using response time measures to assess 'guilty knowledge'. Journal of Applied Psychology, 85, 30–37.

Sitton, S. C., & Griffin, S. T. (1981). Detection of deception from clients' eye contact patterns. Journal of Counseling Psychology, 28, 269–271.

Spence, S. A., Farrow, T. F. D., Herford, A. E., Wilkinson, I. D., Zheng, Y., & Woodruff, P. W. R. (2001). Behavioural and functional anatomical correlates of deception in humans. Neuroreport: For Rapid Communication of Neuroscience Research, 12, 2849–2853

Spence, S. A., Farrow, T. F. D., Leung, D. Shah, S., Reilly, B., Rahman, A., & Herford, A. (2003). Lying as an executive function. In P. W. Halligan, C. Bass, & D. A. Oakley (Eds.), Malingering and illness deception (pp. 255–266). Oxford, England: Oxford University Press.

Sporer, S. L. (1997). The less traveled road to truth: Verbal cues in deception detection in accounts of fabricated and self-experienced events. Applied Cognitive Psychology, 11, 373–397.

Steele, H., & Steele, M. (Eds.). (2008). Clinical applications of the adult attachment interview. Guilford Press.

Stiff, J. B., & Miller, G. R. (1986). "Come to think of it . . . ": Interrogative probes, deceptive communication, and deception detection. Human Communication Research, 12, 339–357.

Stiff, J. B., Corman, S., Krizek, B., & Snider, E. (1994). Individual differences and changes in nonverbal behavior: Unmasking the changing faces of deception. Communication Research, 21, 555–581.

Streeter, L. A., Krauss, R. M., Geller, V., Olson, C., & Apple, W. (1977). Pitch changes during attempted deception. Journal of Personality and Social Psychology, 24, 12–21.

Stromwall, L. A., Hartwig, M., & Granhag, P. A. (2006). To act truthfully: Nonverbal behaviour and strategies during a police interrogation. Psychology, Crime, & Law, 12, 207–219.

Todd-Mancillas, W. R., & Kibler, R. J. (1979). A test of concurrent validity for linguistic indices of deception. Western Journal of Speech Communication, 43, 108–122.

Vendemia, J. M. C., Buzan, R. F., & Green, E. P. (2005a). Practice effects, workload, and reaction time in deception. American Journal of Psychology, 118, 413–429.

Vendemia, J. M. C., Buzan, R. F., & Simon-Dack, S. L. (2005b). Reaction time of motor responses in two-stimulus paradigms involving deception and congruity with varying levels of difficulty. Behavioural Neurology, 16, 25–36.

Verschuere, B., Crombez, G., & Koster, E. H. W. (2004b). Orienting to guilty knowledge. Cognition and Emotion, 18, 265–279.

Verschuere, B., Crombez, G., De Clercq, A., & Koster, E. H. W. (2004a). Autonamic and behavioural responding to concealed information: Differentiating orienting and defensive responses. Psychophysiology, 41, 461–466.

Vrij, A. (1993). Credibility judgments of detectives: The impact of nonverbal behavior, social skills, and physical characteristics on impression formation. Journal of Social Psychology, 133, 601–610.

Vrij, A. (1995). Behavioral correlates of deception in a simulated police interview. Journal of Psychology, 129, 15–28.

Vrij, A. (2000). Detecting Lies and Deceit: The Psychology of Lying and Implications for Professional Practice. (Wiley Series on the Psychology of Crime, Policing and Law).

Vrij, A. (2006a). Challenging interviewees during interviews: The potential effects on lie detection. Psychology, Crime, & Law, 12, 193–206.

Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities*. John Wiley & Sons.

Vrij, A., & Heaven, S. (1999). Vocal and verbal indicators of deception as a function of lie complexity. Psychology, Crime, & Law, 5, 203–315.

Vrij, A., & Mann, S. (2001a). Telling and detecting lies in a high-stake situation: The case of a convicted murderer. Applied Cognitive Psychology, 15, 187–203.

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.

Vrij, A., Akehurst, L. Soukara, S., & Bull, R. (2004a). Detecting deceit via analyses of verbal and nonverbal behavior in children and adults. Human Communication Research, 30, 8–41.

Vrij, A., Edward, K., & Bull, R. (2001a). People's insight into their own behaviour and speech content while lying. British Journal of Psychology, 92, 373–389.

Vrij, A., Edward, K., & Bull, R. (2001c). Stereotypical verbal and nonverbal responses while deceiving others. Personality and Social Psychology Bulletin, 27, 899–909.

Vrij, A., Edward, K., Roberts, K. P., & Bull, R. (2000). Detecting deceit via analysis of verbal and nonverbal behavior. Journal of Nonverbal Behavior, 24, 239–263.

Vrij, A., Mann, S., & Fisher, R. (2006a). An empirical test of the Behaviour Analysis Interview. Law and Human Behavior, 30, 329–345.

Vrij, A., Mann, S. A., Fisher, R. P., Leal, S., Milne, R., & Bull, R. (2008). Increasing cognitive load to facilitate lie detection: The benefit of recalling an event in reverse order. *Law and human behavior, 32*(3), 253-265.

Vrij, A., Semin, G. R., & Bull, R. (1996). Insight in behavior displayed during deception. Human Communication Research, 22, 544–562.

Vrij, A.,& Mann, S. (2001a). Telling and detecting lies in a high-stake situation: The case of a convicted murderer. Applied Cognitive Psychology, 15, 187–203.

Walczyk, J. J., Roper, K. S., Seemann, E., & Humphrey, A. M. (2003). Cognitive mechanisms underlying lying to questions: Response time as a cue to deception. Applied Cognitive Psychology, 17, 755–744.

Walczyk, J. J., Schwartz, J. P., Clifton, R., Adams, B., Wei, M., & Zha, P. (2005). Lying person-to-person about live events: A cognitive framework for lie detection. Personnel Psychology, 58, 141–170.

Weiler, J., & Weinstein, E. (1972). Honesty, fabrication, and the enhancement of credibility. Sociometry, 35, 316–331.

Winkel, F. W., & Vrij, A. (1995). Verklaringen van kinderen in interviews: Een experimenteel onderzoek naar de diagnostische waarde van Criteria Based Content Analysis. Tijdschrift voor Ontwikkelingspsychologie, 22, 61–74.

Zaparniuk, J., Yuille, J. C., & Taylor, S. (1995). Assessing the credibility of true and false statements. International Journal of Law and Psychiatry, 18, 343–352.

Zhou, L., & Zang, D. (2006). A comparison of deception behavior in dyad and triadic group decision making in synchronous computer-mediated communication. Small Group Research, 37, 140–164.

Zuckerman, M., DeFrank, R. S., Hall, J. A., Larrance, D. T., & Rosenthal, R. (1979). Facial and vocal cues of deception and honesty. Journal of Experimental Social Psychology, 15, 378–396.