Development and Validation of Ostensive Communication Coding System

Jeffy Chun Yu Ho


University College London
UCL Doctorate in Clinical Psychology

Thesis declaration form

I confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signature: [signature]

Name: Jeffy Chun Yu Ho

Date: 27th July 2020
Overview

This three-part thesis reviews, develops and attempts to establish the psychometric properties of the Ostensive Communication Coding System (OCCS).

Part One: Conceptual Introduction – The conceptual introduction considers the key concepts which are addressed in the empirical paper. It does so by reviewing literature which has addressed the theory of natural pedagogy and ostensive communication while at the same time looking at the relevant empirical evidence which is presented in the body of literature. A critical evaluation of the literature suggests that there are major limitations regarding the empirical evidence. The author therefore calls for the development of a new instrument which is to collect the data needed to address these limitations.

Part Two: Empirical Paper – The empirical paper reviews and integrates the existing literature on the four ostensive cues which have been identified to be part of ostensive communication. Behavioural operational definitions and measuring methods in regard to each of the ostensive cues are derived from that foundation. This paper further reports how the OCCS has been applied to two learning paradigms to test the psychometric properties of the instrument. Results are reported and discussed. Also, some directions for further development of this instrument are recommended. This was a joint project with Shruit Jamnadass (DClipsy, 2020). The contribution of the authors is summarised in Appendix A.

Part Three: Critical Appraisal - The critical appraisal uses a reflective stance to discuss experiences learned in the research process.
Impact Statement

The study presented here has developed and undertaken an initial examination of the psychometric properties of a new instrument which has been designed to capture the ostensive communication of mother-infant dyads in naturalistic interaction. The results of the study have provided initial support of the reliability and validity of the ostensive communication coding system (OCCS). That has paved the way for a further validation of the instrument. The findings discussed, when replicated with a larger sample, the successful development of the instrument will make it possible to collect data that can reflect the comprehensive essence of ostensive communication, which then can help to address the major limitations of the empirical evidence supporting the theory of Natural Pedagogy. The OCCS also provides reliable means to study and investigate the dynamic interactions among the four identified ostensive cues and therefore contributes to the understanding of the role natural pedagogy has played regarding the phenomenon of cultural evolution.

In addition, for clinical purposes, the OCCS can potentially be used as a screening tool to identify individual who might have impairment in the capacity of ostensive communication. In future study, when replicate with a clinical sample (e.g. Mothers with major depression) who is likely to have higher risk of impaired ostensive communication, the data generated by OCCS can be compared with the normative sample and therefore to establish the benchmark of a clinical cut-off. Comparison of an individual’s OCCS scores against the clinical threshold not only can help clinician to identify individual who might need support to enhance their ostensive communication, information obtained from the evaluation of individual ostensive cue can also be used as a guide for focused intervention.
# Table of Contents

**List of Tables and Figures** 10

**Acknowledgements** 12

**Part One: Conceptual Introduction** 13

**Re-examination of Empirical Evidence of Natural Pedagogy** 13

**and Ostensive communication** 13

**Abstract** 14

**Introduction** 15

Communication and Ostension .......................................................................................................................... 16

**Code-based ostensive communicative system and natural pedagogy** 18

Eye contact .......................................................................................................................................................... 19

Infant Directed Speech (IDS) .............................................................................................................................. 20

Contingent response ......................................................................................................................................... 22

**Ostensive cues and the state of epistemic trust** 25

Presumption of referent .................................................................................................................................... 26

Presumption of shared knowledge and genericity .............................................................................................. 27

Presumption of relevance and over-imitation ...................................................................................................... 28

**Ostensive cues and learning** 31

Eye contact .......................................................................................................................................................... 33

IDS ...................................................................................................................................................................... 34

Contingent response .......................................................................................................................................... 36

**Critical evaluation and discussion** 37

Lack of unifying behavioural operational definition .......................................................................................... 38

Real-life phenomenon or laboratory artefact? ..................................................................................................... 41

Developmental quality of ostensive communication and more questions ......................................................... 43

**Proposal for future research** 44

**References** 47

**Part Two: Empirical Paper** 61

**Development and validation of Ostensive Communication Coding System** 61

**Abstract** 62

**Introduction** 63

Lack of unifying operational definition ........................................................................................................... 64

Real-life phenomenon or laboratory artefact? ..................................................................................................... 65

Co-occurrence of ostensive cues and its developmental quality ........................................................................ 66
Present study 70

Step 1. Empirical basis of OCCR & Step 2. Behaviour operational definition 71
Eye contact ........................................................................................................ 71
Behavioural operational definition of Eye contact in present study .................. 71
IDS ..................................................................................................................... 72
Behaviour operational definition of IDS in present study ................................ 76
Contingent responsivity ................................................................................... 79
The origin of contingent response ..................................................................... 79
Operational definition of contingent responses in empirical studies ............... 80
Behavioural operational definition of contingent responsivity in present study 87
Name addressing ............................................................................................. 98
Behavioural operational definition of name addressing in present study ....... 98
Co-occurrence .................................................................................................. 99
Behaviour operational definition of Co-occurrence in present study .............. 99

Step 3. Considerations of learning paradigms .................................................. 100
Task A ............................................................................................................... 100
Task B ............................................................................................................... 100

Step 4. Considerations of the initial examination of OCCS ............................. 102
The capacity to evaluate the performance of ostensive communication ......... 102
Mentalisation and reflective functioning ......................................................... 102
Parenting stress level ....................................................................................... 103

Method 105

Ostensive communication (OC) project .......................................................... 105
Ethical approval ............................................................................................... 105

Participants ..................................................................................................... 105
Power analysis ................................................................................................. 105
Recruitment ..................................................................................................... 106

Study design and learning tasks ..................................................................... 106
Learning Task A: the imitation paradigm (switching on the lamp by using the forehead) ............. 107
Learning task B: word-object association ....................................................... 108

Procedure ....................................................................................................... 109
Setting .............................................................................................................. 109

Measures ......................................................................................................... 119
Ostensive communication coding system (OCCS) .......................................... 119
Demographics Questionnaire (Appendix I) .................................................. 119
Time ................................................................................................................ 121
Attention ....................................................................................................... 123
IDS ................................................................................................................... 123
Contingent responsivity ................................................................................. 125
Co-occurrence ............................................................................................... 127
The rate of using ostensive cue ...................................................................... 130
Measures used for validity analyse ................................................................. 131

Step 5. Generating data of ostensive communication via OCCS ..................... 134
Coding of ostensive communication ............................................................... 134

Step 6. Establish the psychometric property of OCCS .................................... 135
Covid-19 .......................................................................................................... 138
Result  142
Demographic characteristics .............................................................................. 142
Descriptive statistics .......................................................................................... 144
Hypothesis 1: OCCS achieves high to excellent level of inter-rater reliability in all sub-scale. …146
Hypothesis 2: Participants with high frequency of ostensive cues tend to have high score of co-occurrence. ................................................................. 151
Hypothesis 3: The performance of different dimensions of OCCS in Task A are considerably different from Task B. .............................................................. 157
Hypothesis 4: There is no observable pattern between the performance of frequency items and the rate items .................................................................................. 166
Hypothesis 5: The constructs of mentalisation show observable pattern of relationships with components of contingent responsivity and ostensive communication in general. .................. 167
Hypothesis 5 continue, and hypothesis 6: Score of PSI-SF shows observable pattern of negative relationship with the different dimensions of OCCS .................................................. 189
Discussion  217
Inter-rater reliability and its implication .............................................................. 217
The dimension of rate ........................................................................................ 220
Co-occurrence and the relationships among the three dimensions .................. 223
Differential validity ............................................................................................. 223
Construct validity ............................................................................................... 225
  Reflective functioning and Contingent responsivity (CR) .................................. 225
  Parenting stress (PSI-SF) and ostensive communication ................................... 229
  Applying OCCS in two learning paradigms and its implications ....................... 230
Clinical implication ............................................................................................. 232
  The Case study ................................................................................................. 234
Limitation and further development of OCCS .................................................... 235
  Measuring eye contact and name addressing .................................................. 235
  Measuring and validating IDS ......................................................................... 236
  Measuring contingent responsivity ................................................................. 236
Conclusion  237
References  239
Part 3: Critical Appraisal  256
Introduction  257
Influence of prior experience and theoretical orientation on my approach ............ 257
Major challenge – recruitment .......................................................................... 258
Major challenge - testing ................................................................................... 261
Adaption to COVID 19 ...................................................................................... 263
Conclusion  266
Reference 267
Appendices

Appendix A: Contributions to the Research Project................................................................. 287
Appendix B: Ethical Approval.................................................................................................. 288
Appendix C: Participant Information Sheet ........................................................................... 289
Appendix D: Participant Consent Form .................................................................................. 294
Appendix E: Participant Debrief Form ................................................................................... 296
Appendix F: Recruitment Poster ............................................................................................ 298
Appendix G: Appointment Letter to Participants .................................................................... 300
Appendix H: Instruction Sheet for Experiment ........................................................................ 302
Appendix I: Demographics Questionnaire ............................................................................ 305
Appendix J: Case Study .......................................................................................................... 313
Appendix K: Coding system.................................................................................................... 320
Appendix L: Baby Bib .............................................................................................................. 321
List of Tables and Figures

Table 1. Operational definition of Eye contact and Name addressing 71
Table 2. Operational definition of Infant Directed Speech 76
Table 3. Operational definition of Contingent responsivity 88
Table 4. Operational definition of Name addressing 98
Table 5. Operational definition of co-occurrence 99
Figure 1. Study design 109
Table 6. Summary of the experiment procedure 114
Figure 2. Overview of ostensive communication coding system 120
Table 7. Definition of the start and the end of a teaching episode in Task A and Task B 121
Table 8. Reversing score of the scale of Directive and Intrusive behaviours 126
Table 9. Computation of co-occurrence score 129
Table 10. Sample Demographic information (N=14) 142
Table 11. Mean (SD) of Ostensive cue items in Task A and Task B (N=14) 144
Table 12. Interrater reliability of ostensive communication coding system base of Task B (N=14) 146
Table 13. Composition of ostensive communication performance score 149
Table 14. Ostensive cue performance score 149
Table 15. Correlation of ZF (A) and ZR (A), ZGR (A) and Co-occurrence (A) 151
Table 16. Comparison of OCCS dimensions in Task A 151
Figure 3a and 3b. Observation of OCCS dimensions in Task A. 152
Table 17. Correlation of ZF (B) and ZR (B), ZGR (B) and Co-occurrence (B) 153
Table 18. Comparison of OCCS dimensions in Task B 154
Figure 4a and 4b. Observation of OCCS dimensions in Task A. 155
Table 19. Correlation between ZF, ZR, ZGR and Co-occurrence between Task A and Task B 157
Table 20. Comparison of ZF, ZR, ZGR, Co-occurrence and Rate of Co-occurrence scores of Task A and Task 158
Figure 5a and 5b. Observation of the ZF dimension in Task A and Task B. 160
Figure 6a and 6b. Observation of the ZR dimension in Task A and Task B. 162
Figure 7a and 7b. Observation of the ZGR dimension in Task A and Task B. 163
Figure 8a and 8b. Observation of the ZF dimension in Task A and Task B. 164
Table 21. Frequency, rate and time 166
Table 22. Mean (SD) of Parents Reflective Functioning Questionnaire and Parenting Stress Index 167
Table 23. Correlation between Pre-Mentalizing modes (PM) with Scaffolding, Quality of contingent behaviours and emotional support in Task A & Task B 168
Table 24. Pre-mentalising state (PM) and Contingent responsivity (CR) 169
Figure 9a and 9b. Observation of the relationship between PM and Scaffolding in Task A and Task B. 171
Figure 10a and 10b. Observation of the relationship between PM and CB quality in Task A and Task B. 172
Figure 11a and 11b. Observation of the relationship between PM and emotional support in Task A and Task B. 173
Table 25. Correlations between Certainty of Mental state (CM) with Scaffolding, Quality of contingent behaviours and emotional support in Task A & Task B 176
Table 26. Certainty of Mental State (CM) and Contingent responsivity (CR) 177
Figure 12a and 12b. Observation of the relationship between CM and Scaffolding in Task A and Task B. 178
Figure 13a and 13b. Observation of the relationship between CM and CB quality in Task A and Task B. 179
Figure 14a and 14b. Observation of the relationship between CM and emotional support in Task A and Task B. 180
Table 27. Correlations between Interest and Curiosity in mental state (IC) with Scaffolding. Quality of contingent behaviours and emotional support in Task A & Task B 182
Table 28. Interest and Curiosity in mental state and Contingent responsivity

Figure 15a and 15b. Observation of the relationship between IC and Scaffolding in Task A and Task B.

Figure 16a and 16b. Observation of the relationship between IC and CB quality in Task A and Task B.

Figure 17a and 17b. Observation of the relationship between IC and Emotional support in Task A and Task B.

Table 29. Mean (SD) of Parenting Stress Index – short form (PSI-SF)

Table 30. Summary of observable patterns of scores among OCCS, PRFQ and PSI-SF

Table 31. Correlations between Interest and Curiosity in mental state (IC) with Scaffolding, Quality of contingent behaviours and emotional support in Task A & Task B

Table 32. The relationship between ZF (A), Parenting reflect functioning and Parenting Stress Index

Figure 18a and 18b. Observation of the relationship between PRFQ and ZF(A).

Figure 19a and 19b. Observation of the relationship between PSI and ZF (A).

Table 33. The relationship between ZR (A), Parenting reflect functioning and Parenting Stress Index

Figure 20a and 20b. Observation of the relationship between PRFQ and ZR(A).

Figure 21a and 21b. Observation of the relationship between PSI and ZR(A).

Table 34. The relationship between ZGR (A), Parenting reflect functioning and Parenting Stress Index

Figure 22a and 22b. Observation of the relationship between PRFQ and ZGR(A).

Figure 23a and 23b. Observation of the relationship between PSI and ZGR(A).

Table 35. The relationship between Co-occurrence (A), Parenting reflect functioning and Parenting Stress Index

Figure 24a and 24b. Observation of the relationship between PRFQ and Co-occurrence (A).

Figure 25a and 25b. Observation of the relationship between PSI and Co-occurrence (A).

Table 36. The relationship between ZF (B), Parenting reflect functioning and Parenting Stress Index

Figure 26a and 26b. Observation of the relationship between PRFQ and ZF(B).

Figure 27a and 27b. Observation of the relationship between PSI and ZF(B).

Table 37. The relationship between ZR (B), Parenting reflect functioning and Parenting Stress Index

Figure 28a and 28b. Observation of the relationship between PRFQ and ZR(B).

Figure 29a and 29b. Observation of the relationship between PSI and ZR(B).

Table 38. The relationship between ZGR (B), Parenting reflect functioning and Parenting Stress Index

Figure 30a and 30b. Observation of the relationship between PRFQ and ZGR(B).

Figure 31a and 31b. Observation of the relationship between PSI and ZGR(B).

Table 39. The relationship between Co-occurrence (A), Parenting reflect functioning and Parenting Stress Index

Figure 32a and 32b. Observation of the relationship between PRFQ and Co-occurrence (B).

Figure 32a and 32b. Observation of the relationship between PSI and Co-occurrence (B).

Table 40. Comparison of participants with lowest and highest no. of name addressing

Table 41. ZF, ZR and ZGR total of P24 & P49

Table 42. Comparison of ostensive communication performance between P24 and P49

Table 43. Comparison of PRFQ & PSI between P24 & P49
Acknowledgements

This thesis represents the culmination of a 10-year journey of my intellectual and clinical adventure in London. I would like to express my deep gratitude to everyone who has been part of this amazing and challenging journey.

Firstly, I would like to thank my parents for their unconditional support of my pursuit to become a psychotherapist and clinical psychologist. It is an eternal sorrow that my father passed away in 2018. He cannot be present to celebrate my accomplishment of this training and to accept my deepest gratitude to his devotion to support me. I would also like to thank my dearest mother for her unconditional love and care, without her support, I would never come to where I am as a professional and as a person.

Secondly, I would like to thank all my friends who have continuously supported me. It would be an impossible journey to travel without them. I would especially like to thank Simon Marsh and Stephanie Grimm for their helpful support with writing this paper.

Thirdly I would like to thank my research partner, Shruti, who has been the major support throughout this challenging process.

Finally, I would like to thank my two my research supervisors, Prof. Peter Fonagy and Tobias Nolte MD for their wisdom, patience, encouragement and support.
Part One: Conceptual Introduction

Re-examination of Empirical Evidence of Natural Pedagogy

and Ostensive communication
Abstract

The theory of Natural Pedagogy as proposed by Csibra and Gergely (2006, 2010) has provided a comprehensive account of the phenomenon of cultural learning. The theory suggested that ostensive communication in early interaction between the mother-infant dyad plays a crucial role in facilitating the transmission of cultural knowledge. The three ostensive cues identified in that theory are eye contact, being addressed by name and contingent responsivity. This conceptual introduction provides a detailed review of the theory as well as the relevant literature. A critical evaluation of the empirical evidence presented in the body of literature suggests there are major limitations within that evidence, which as a consequence makes the empirical evidence fall short of offering robust support to the theory. This paper concludes by calling for development of a new instrument which allows for collecting data that is necessary for addressing these limitations.
Introduction

What makes humans unique as a species? Why are we, unlike any other animal, capable of developing extremely sophisticated social systems and technological infrastructures and have therefore made advancements on many fronts, ranging from art or sports to agriculture and political systems? This continuous refinement and the dispersion and development of skills and knowledge have made human civilisation more prosperous and capable of surviving. Heyes (2016) considers this process as evolutionary and refers to it as cultural evolution. What makes humans truly distinctive may depend on their capacity to learn from their ancestors (Campbell, 1965; Richerson & Boyd, 2005; Tomasello, 2014). Whilst most other species mainly pass on information to their offspring through genetic mechanisms, humans inherit them through cultural learning (Heyes, 2016). The obvious challenge is that each generation of humans has to learn and manage more complex cultural practices, technological skills and object functioning - things that often are impossible to absorb through mere observation. A lot of cultural knowledge and information therefore is epistemically and cognitively opaque. For example, a modern individual would not be able to fully utilise the complex workings of a car merely by observing how his or her parents drive it. It would take extensive teaching and learning for an individual to acquire the skills needed for driving. The increasing complexity of information creates problems of learnability. Insufficient transparency and a lack of observable teleological and causal properties of cultural knowledge might therefore endanger the successful intergenerational inheritance (Csibra & Gergely, 2006; Csibra & Gergely, 2011; Gergely & Csibra, 2005). To understand the phenomenon of cultural evolution, we need to examine how knowledge is transmitted from one generation to the next: In essence, how human communication works in the context of pedagogy.
Some developmental researchers believed that human infants have inherited an innate capacity to be receptive to the communication of knowledge and information (e.g., Batki et al., 2000; Csibra & Gergely, 2009; Mason et al., 2019). Among different theories, natural pedagogy, as proposed by Csibra & Gergely (2006) and Gergely & Csibra (2005), provides a ground-breaking and comprehensive account of this phenomenon called cultural learning. They maintain that human infants have a “well-organised package of biases, tendencies and skills” (Csibra & Gergely, 2006, p. 8) that allows them to be highly sensitive to communication from adults. Therefore they can learn the information presented in an exceptionally effective manner. Csibra and Gergely (2011) also consider it a unique human genetic and biological adaptation which enables them to teach and learn. In support of this theory, they reviewed the evidence from various studies. They conclude that this package of biases can reflect the essence of ostensive cueing in naturalistic interactions between mother and infant. To shed light on this knowledge gap, this paper will at first review the theory and then critically evaluate the relevant empirical studies and evidence of past decades. The ultimate aim is to enrich the theory by identifying appropriate behavioural operationalisations of ostensive cues therefore allowing examination of natural ostension in future research.

Communication and Ostension

According to the classical code model of communication, communication is merely a coding-decoding mechanism. In the context of human communication, this means the addresser encodes the intended message into signals, which then get transmitted through a medium. The addressee receives and decodes these signals, so that the meaning of the message can be comprehended and understood. This approach can explain various communication phenomena; however, it cannot give a full account of features which make human social communication unique and in which information is often transformed, distorted
and lost (Sperber & Wilson, 1997). Taking a different approach from the code model, Grice (1989) believes that human communication is a more complicated matter, in the sense that is an expression and recognition of intention. Building on this work, Sperber (1986, 1995); Sperber and Wilson (1997) took a different approach and proposed the relevance theory and the inferential model of communication. They argued that human communication is an inferential process; the addresser provides evidence of the intention to communicate certain information. This is then inferred by the addressee and an expectation of relevance is induced. This expectation will guide the audience towards decoding the meaning of the addressee’s utterances. Therefore, any act of human communication has two intentions: an informative intention and a communicative intention. The former is referring to the intention to transmit information and thereby to modify the addressee’s cognitive environment. This goal is achieved by using ostensive communication, therefore getting the addressee to recognise the communicative intention and be receptive to the information presented.

Ostension is broadly defined as signals that manifest the communicator’s communicative intention, with the effect of drawing the addressee’s attention to the source of information. As much as Csibra (2010) agreed on Sperber and Wilson’s claim, he believed that the communication between mother and young infant is different from those between adults. For the inferential process to take place would be necessary that an individual has prior background knowledge, for example about the historical context of this conversation; or shared information about the relevance of this particular semantic event. All of these things are not usually available to new-borns. They have insufficient linguistic skills and cognitive recourse; therefore, when interpreting communication, it is less likely that young infants are employing a comparable higher-level cognitive process as their adult counterpart would.
In his article published in 2010, Csibra argued that in early communications, parents and young infants are using a code-based communicative system. The operation of this system does not rely on an inferential process which would require a higher-level cognitive capacity. Instead, infants would “decode” specific ostensive signals, like a reflex reaction. They would switch into a particular state of mind that makes reception and processing of new information possible. This system seems to be genetically and biologically adaptive. Evidence from various studies indicated that new-borns demonstrate innate preparedness to recognize these signals or cues (E.g. Batki et al. 2000; Teresa et al., 2002; Masataka, 2003; Marcela et al., 2003; Vouloumanos, 2007; cited in Csibra, 2010) - not only from their parents but also from unfamiliar people or non-human sources (Watson, 1972).

In general, to classify a behavioural signal or cue as ostensive, it has to fulfil two main criteria: first, it needs to demonstrate the informative intention from the communicator and secondly it has to indicate who is being addressed by this intention (Dan Sperber & Wilson, 1997). However, in this specific context of early development and pedagogy, the signal has to be innately specified. Csibra proposed three additional features: “These signals (1) must unambiguously specify that the infant is the addressee of a communicative act, (2) must be discriminable by new-borns, and (3) must induce preferential orientation towards their source.” (Csibra, 2010, p. 144). Three types of cues can match with this definition:

1. Eye contact
2. Infant-directed speech (IDS) or Motherese
3. Contingent response
Eye contact

In everyday communication, direct gazing into another person’s eyes to form eye contact is always a powerful signal to initiate or establish a communicative link. When a person directs his gaze to another person, there is a natural tendency for the individual at the receiving end to assume themselves to be target audience of the information that the addressee (who initiates the eye contact) intends to communicate. Humans also use eye contact to express intimacy and regulate social interaction (Kleinke, 1986). Current evidence in developmental, social and cognitive neuroscience give an idea of the effect of eye contact in adult humans. For example, there is a network in the brain, usually termed “the social brain” (Adolphs, 2009; Brothers, 1990; Johnson et al., 2005) which is specialised in processing social information and interaction. Eye contact has been proven to modulate the development and activation of this area (Senju & Johnson, 2009).

Human infants seem to acquire the capacity of eye-detection very soon after birth. In their study in 2000, Batki et al. showed two female adults photographs to new-borns, one with eyes open, the other with her eyes closed. They reported that, on average at age of 36 hours human infants already showed a preference to look at the face with open eyes. Teresa et al. (2002) came to similar conclusions when they showed photographs to 3-day-old infants. Not a single infant among the 105 participants chose to look at the face where the eyes were averted, over those where someone appeared to make eye contact with them. Several eye-tracking studies found that 5-6 months infants followed the gaze of a person on a computer screen (Gredebäck et al., 2008; Joanna et al., 2014; Senju & Cisbra, 2008).

Eye-contact also tends to elicit positive affective responses in young infants. Various studies found new-borns smiled significantly more to a person who looked into their eyes and the smile diminished very soon after the eye contact was broken (Caron et al., 1997; Hains & Muir, 1996; Symons et al., 1998). Besides, similar to what goes on with an adult counterpart,
eye contact tends to enhance infants’ attention to the face. They showed better accuracy in memorizing faces when eye contact took place (Hood et al. 2003).

Not only that infants demonstrate innate sensitivity towards eye contact, they also expect to learn referential information after they are addressed. Evidence of such process was reported in various studies. For example, 6-month-old infants were found significantly more likely to follow the actor’s gaze to the target object in ostensive condition (direct eye contact) than non-ostensive condition (Senju & Csibra, 2008; Szufnarowska et al., 2014). However, several researchers offered alternative explanation as they have also demonstrated that infants are sensitive to actor’s head and eye motions, they therefore argued that infants’ expectation to learn referential information can also be cued by these movements (e.g. Csibra (2010); C. Moore et al. (1997). Further review of these experiments suggested that procedures in these experiments actually involved face-to-face interaction with or direct gazing towards the infant prior to the head or eyes movements, neither of these studies therefore can tease apart if these movements alone can induce gaze following without eye contacts.

Overall speaking, there are strong evidence suggest that eye contact meets the proposed criteria and therefore should classify as an ostensive cue.

**Infant Directed Speech (IDS)**

Speech is, without doubt, the most frequently used signal and medium for communication. It is a universal phenomenon across languages and cultures that mothers instinctively use IDS, also referred as the “baby talk or motherese” (Newport, 1975), when having verbal communication with young infants and children (Broesch & Bryant, 2015; Bryant & Barrett, 2007; Fais et al., 2010; Ferguson, 1964; A. Fernald, 1989; Grieser & Kuhl, 1988; Kitamura et al., 2001; Kline, 2013; Lee et al., 2008; Liu et al., 2007; Papoušek et al., 1991; Saint-Georges et al., 2013; Spinelli et al., 2017). Compared with adult directed speech
(ADS), many studies have shown that IDS is significantly different in its prosodic contours properties and linguistic structure. Electrolaryngographic measures indicated that IDS usually comes with the higher fundamental frequency (pitch, F0), and it is broader in range (F0 variability), slower in pace and with longer pauses (Fisher & Tokura, 1995; Grieser & Kuhl, 1988). IDS also tends to come as short sentences or phrases with more straightforward vocabulary, grammar and high frequency of repetition (Anne Fernald & Simon, 1984). It is very explicit in meaning, and very often comes in the form of a question (Soderstrom et al., 2008).

As early as in the third trimester, the human foetus shows responsiveness to auditory information (Hirnholz & Benacerraf, 1983). Decasper et al. (1994) found that foetuses that at the age of 37 weeks reacted differently when exposed to a recording of their mother’s nursery rhymes in comparison to hearing an unfamiliar one. That indicates human infants are able to distinguish maternal speech from other audio stimuli even before birth. Furthermore, 2-day-old infants have been observed to have early attentional preferences for IDS, and they looked significantly longer to faces producing IDS than to those speaking ADS (Cooper & Aslin, 1990; Anne Fernald, 1985; Anne Fernald & Kuhl, 1987; Mason et al., 2019). Csibra and Gergely (2011) argued that the infants’ auditory system is adapted to discriminate human speech, especially their mother’s, from other noise. Neurocognitive studies using neurological and brain imaging methods have shown evidence that supports this proposal. For example, despite the new-born being asleep, increased blood flow in the frontal cerebral region was identified when they heard IDS. Another research showed that IDS brought more brain activation in infants’ (6-13 months-old) left and right temporal areas (Naoi et al., 2012). Furthermore the hypothesis was confirmed that they have an innate sensitivity to the acoustic properties of IDS (Saito et al., 2007).
As mentioned before by no means infants at this stage can recognise the communicative and informative intent via the content of the speech. Masataka, (2003) conducted a study and found that mothers would naturally exaggerate the prosodic features of IDS when not being able to get the infant’s attention. Although findings are less consistent regarding the relationship between IDS and infants’ attention. For example, some researchers found that higher F0 mean and wider F0 variability are associated with increased attention in newborns (Butler & Markman, 2014; Spinelli et al., 2017). 6- to 8-month old infants (R. S. C. Phillips, 1994), while other author found opposite association 6- to 9-month-old show no preference to utterance with a higher F0 mean, 3-month old infants pay better attention to utterances with a lower F0 mean and F0 variability (Kitamura & Lam, 2009). IDS was found as a strong cue that guides infant attention towards referential information. Senju & Csibra (2008) demonstrated that 6-month-old are more likely to follow an adult gaze towards the targeted object if the adult addressed the object in IDS. Overall, evidences in the literature are mostly consistent with the theory and demonstrate that not only infants are able to distinguish IDS, but that IDS makes them recognise instantly that a communicative intent is directed at them. This allows the infant to unambiguously register themselves as the addressee, orientate towards the speaker and eventually be biased to receive the incoming information.

**Contingent response**

The most commonly observed contingent response is the turn-taking pattern in everyday dialogue between adults: one person speaks whilst the other party remains silent and wait for their turn to respond. Csibra (2010) suggested that the natural synchronization of this process – the communication going back and forth - is ostensive because the individual learns that the response from the other party is related to what he or she is doing. Therefore,
they can assume that there is a communicative intention in this exchange even without knowing precisely what the information is.

Contingent turn-taking is also observable in early mother and infant interaction from birth. The primary form of turn taking occurs on the most elementary level – sucking the mother’s nipple. Masataka (2003) reported that infants start to insert breaks into their sucking pattern for no physiological benefit. Mothers tend to respond to these pauses with jiggling the infant. Not only does the sucking tend to be resumed as soon as the mother stops jiggling, but the study also found that the infants tend to pause more frequently when the mother was not responsive during those pauses. Therefore, it seemed to demonstrate infants’ innate sensitivity toward contingent reactions. Besides, Csibra (2010) suggested this phenomenon implied that infants were expecting a reaction (jiggling) from their mother whenever they produced a signal (a pause). It is the most fundamental way for an infant to experiment and practice contingent responsivity in their environment. In situations where a child plays, it is not uncommon to see mother remain silent and still while the infant engages in an activity, but as soon as the infant takes a pause, the mother tends to engage promptly by utterance or physical touch. These engagements are halted soon after the infant resume their play. Infants might also initiate turn-taking by offering toys to their mother. The mother then plays with it, describes its features to expand the cognitive awareness and soon returns it to the infant for further exploration. This process often repeats itself for several times. More supportive evidence suggested that infants prefer contingent interactions over random stimulation (A. E. Bigelow, 2001; Ann E. Bigelow & Rochat, 2006; Nadel et al., 1999; Stormark & Braarud, 2004). Goldstein and Schwade (2008) found that infants learned information more effectively from their parents when their responses come in a contingent manner (within 2-5s).

In fact, Watson (1972, 1979, 1985, 1994) reported similar findings. In his study conducted in 1972, he tied an object to infants’ leg so this object moved contingently with the
infants’ leg movement. Watson not only observed an increased rate of leg kicking in the contingent condition; he also found that the infants exhibited positive social responses like smiling and cooing. These results provided evidence that very young infants have an innate capacity to detect contingency in their actions and external stimuli, and that the experience of causally controlling external events arouse positive emotions. Keller et al. (1999) confirmed that contingent responses result in positive affect, where violation of contingency expectations is associated with negative affect. Clearly, contingent response meets the proposed criteria and should class as ostensive cue.

As infants continue to develop, contingent responses are not limited to this single-modal of turn taking but emerge in other forms and media. For example, the infant might test maternal response with their facial expression, vocalisation and physical touching (Bateson, 1979), while the mother also contingently responds to the behaviour displayed by the infant and corresponds to their cognitive and emotional state in similar ways.

**Contingent responsivity and mentalisation**

Watson (1979, 1994) and Gergely and Watson (1996) provided further evidence of young infants’ sophisticated ability to detect and analyse contingency. They investigated the mechanism of contingency detection and the parental capacity of contingent response, especially in regard to parental sensitivity to their infants’ internal state. They proposed that this process of contingent response is vitally important for infant’s psychological development. Building on the infant learning literature and adult biofeedback training studies (e.g. Miller, 1969, 1978), Gergely and Watson suggested that while young infants are incapable of distinguishing their internal emotional state or subjective experience in early life, they learn to detect and categorise them through the repeated encounters of experience of their carers being sensitive to their displayed behaviour in correspondence to their subjective
state and producing “the state-contingent external biofeedback in the form of the empathic reflection” (Gergely & Watson, 1996, p. 1193), which later known as the “marker mirroring interactions”. And it is this process that enables the development of second-order representations of the infants’ subjective experiences and lays the cornerstone of the capacity to mentalise (Fonagy, 1998; Fonagy et al., 2018). Mentalising is referring to the individual’s capacity to understand their own and other’s behaviours by taking mental states, such as thoughts, feelings, wishes and desires, into consideration (Bateman & Fonagy, 2012; Fonagy, 2004; Fonagy & Allison, 2012; Fonagy & Target, 2006). This capacity is pivotal for any individual to be able to predict and interpret other’s actions efficiently, and to navigate the complexity of intra- and interpersonal relationships (Fonagy & Allison, 2012). More importantly, this proposition suggests that the development of the infants capacity to mentalise is largely subjected to the parental sensitivity in caregiving, and this process is mediated by the use of appropriate contingent response or ostensive cues in general (Fonagy et al., 2015). This notion also gives rise to new questions; for example, if parental sensitivity plays a significant role in the capacity to use contingent response. There are individual differences in reflective functioning and other related qualities, therefore intuitively, contingent response should not be merely defined as a binary construct but as an element on a continuous spectrum. This will be further considered in the discussion section of this paper.

Ostensive cues and the state of epistemic trust

What impact do ostensive cues have on infants, and how does it facilitate intergeneration cultural learning with its unique epistemic functions? Csibra and Gergely (2011) and Gergely et al. (2007) discussed this notion extensively. They proposed that, when information is demonstrated or communicated with ostensive cues, human infants tend to switch into a specialised epistemic state – “basic epistemic trust” (Gergely et al., 2007,
p.145), which allows them to attend to and learn about presented information selectively. This state encompasses three essential cognitive biases: (1) referential expectation, (2) an assumption that the knowledge is universally shared by other people in the cultural community and bias to attribute or encode the presented information is generalisable knowledge, and (3) presumption of relevance.

**Presumption of referent**

In the study paradigm of violation-of-expectation looking time conducted by Csibra and Volein (2008), a group of infants (8-12 months old) watched a short video on a computer screen. Under two conditions, a female experimenter in the video appeared to be addressing the infants with ostensive cues (looking at the infant to form eye contact and greet them) before shifting her gaze to look behind one of two barriers. In the consistent condition, an object that appeared to be behind the barrier was revealed, whilst in the inconsistent condition the infants found no object behind the barrier. Infants looked significantly longer at the empty space in the inconsistent condition indicating that they expected to find an object where the female experimenter’s gaze was directed and were surprised nothing was there. Those infants kept going back to the empty side to look for the object. Furthermore, slightly older infants (13 months), who had a certain level of vocabulary, expected not only to find an object but an object the experimenter had named, (e.g. a spoon) in addition to ostensive and deictic gesture (e.g. pointing) (Gliga & Csibra, 2009). Ostensive cues do not just simply capture the infants’ attention but induce an expectation to learn something. Therefore they were surprised when the information was absent.
**Presumption of shared knowledge and genericity**

Various studies have shown, when information is presented via ostensive cues, infants are more likely to attribute it as knowledge that can be generalised to another relevant situation beyond the immediate episodic event. Furthermore they seem to assume that this particular knowledge is available to and shared by other people. This cognitive bias is not limited to enduring object properties, but also human’s subjective experiences or attitude. For example, in an experiment conducted by Egyed et al. (2013), object-directed emotions (Positive for object A, Negative for object B) were presented by one experimenter to infants in ostensive communication condition (such as eye contact, smiling, saying “Hi [baby’s name]” and “Hi, look!”) or non-ostensive communication condition. It was hypothesised that, when this cognitive bias was activated by ostensive cues, the infant would make an object-centred interpretation of the demonstrated emotions (i.e. this kind of object is nice and desirable, everyone knows this). In contrast, without the cognitive bias, the infant would make a person-centre interpretation (i.e. This particular person likes this particular object). In the testing phase, either the same (experimenter 1) or a different experimenter (experimenter 2) came in and requested “Give me one of them” to test infants’ object choice. In the non-ostensive communicative-same person context, 87% of infants chose object A, which indicates that the infant was able to learn about the object-directed emotion and acted in an emotion-congruent manner. In the non-ostensive different person context, only 31% of infants chose object A. So, the experiment confirmed the hypothesis that the infant made a person-centred interpretation. It did not see this knowledge as generalisable and available to other people, but rather as the object preference of one experimenter which was independent from what other people might think. On contrast, in the ostensive commutation-different person context, significantly more (almost 70%) infants chose to give object A. This result
supports the proposal that when presented with ostensive cues, infants made an object-centre interpretation, generalised the learned knowledge and assumed that the attitude is shared between both experimenters. Király et al. (2013) believes this biased attribution induces infants’ social motivation and readiness to acquire shared knowledge in the community, such as social conventions and cultural practices.

In another study, Jennifer et al. (2008) investigated memory biases induced by ostensive communication. When nine-months-old infants were presented with a novel object, study result suggested that infants memorised different information depending on the presence or absence of ostensive cues. In the communicative context, infants tend to retain information of the object’s appearance (permanent property) whilst they retain information of the object’s location (transient episodic property). Various studies have reported similar evidence (e.g. Xu, 2002).

**Presumption of relevance and over-imitation**

The natural pedagogy theory predicted that ostensive communication could also activate the presumption of relevance. Under the influence of this bias, the infant would treat the communicator as a benevolent and reliable source of information. Therefore, the information presented would be attributed with personal relevance rather than taken merely as instrumental action (Csibra & Gergely, 2006; Gergely, 2007; Gergely & Csibra, 2006). The epistemic vigilance, a natural tendency to critically evaluate the truthfulness of the information (D. Sperber et al., 2010) would be relaxed. Therefore, infants can learn cultural knowledge fast and thoroughly without further verification of its validity, even when the knowledge appears to be cognitively opaque or violating the rational principle. Csibra and Gergely argued that this cognitive bias is observable in the phenomenon of over-imitation.
Imitation has a unique position in culture evolution (Legare & Nielsen, 2015; Mark Nielsen et al., 2012). It is described as the “cultural analogue of the mechanisms of genetic inheritance” (Heyes, 2016, p. 290) because it allows cultural information and knowledge to be transmitted with sufficient fidelity and therefore makes space for cumulative development. The ability of infants to reproduce skills or actions of other people is ubiquitous. Heyes (2016) defines imitation as a faithful and intentional reproduction of an observed action. This phenomenon was demonstrated in the classical study conducted by Meltzoff (1988) in his imitation paradigm; 14-months-old infants watched an experimenter’s switching on a lamp by touching her head to a box; a week later, almost 70% of these infants imitated the head action. This result has inspired a lot of empirical research in the past decades in which imitation is considered one of the essential vehicles for culturally and cognitively opaque knowledge and for its vital role in cultural evolution.

There are different types of imitation. *Over-imitation* for example refers to the phenomenon of an individual “to imitate perceivably casual, unnecessary actions in relation to the goal of an action sequence performed by a model” (Hoehl et al., 2019, p. 91) This phenomenon is commonly observable in interactions between infants and adults (Derek et al., 2007; Király et al., 2013).

However, a further study found that infants can show some capacity to be a rational agent and expect other agents to act rationally. Infants of the same age were found to be reluctant to do full imitations in order to accomplish a given objective if there is a plausible explanation that the model acts inefficiently. Gergely et al. (2002) called that “*rational imitation*”. In their study conducted in 2002, they modified the Meltzoff’s (1988) imitation paradigm and tested if efficiency evaluation would change the infants’ choices. In the study an experimenter demonstrated a novel head action. When it was visible that her hands were unavailable, only 21% of infants imitated. It was suggested that when the infants watched the
demonstrator using a less rational means only because her hands were not available, they
assumed they could use a more efficient way when they themselves were unconstrained. Such
findings had been further replicated in other studies looking at various tasks (Buttelmann et
al., 2008; Király, 2009; Tomasello, 2006; Zmyj et al., 2009).

This result posed an interesting question. If infants at this age had the capacity to
evaluate the efficiency in the first place, they could have known what a better way is to
handle the task, whether they see the free hand or not. So what made them reproduce the
novel head action in the original paradigm? Although there are different underlying factors
contributing to over-imitation (see Hoehl et al., 2019 for a review), in their study in 2013,
Kiraly et al. proved that ostensive communication plays a critical role. With a similar
paradigm, the same novel head action was demonstrated to infants in different
communicative contexts. No difference in imitation was found between hands-free and
hands-occupied conditions when ostensive communication was absent. Imitation only took
place when the infants were addressed with ostensive signals. The activation of the
presumption of relevance must have led the infants to over-imitation and make them believe
that it is important to acquire the presented information regardless of the efficiency or the
opacity of the model’s choice of action. In fact, further empirical evidence indicates that not
only infants (Brugger et al., 2007; M. Nielsen, 2006) but also adults (Wang et al., 2011) have
higher tendency to imitate when they are addressed by ostensive cues.

These studies demonstrate the mechanisms how ostensive communication facilitates
the process of cultural evolution through the activation of three essential cognitive biases and
the epistemic trust.
Ostensive cues and learning

The sections above summarised the theory of natural pedagogy, supportive empirical evidence and the proposed mechanism, explaining how the foundation of knowledge and information transmission is established in early communicative exchanges between infant and caregiver and has therefore made cultural evolution possible. It is reasonable to hypothesize that the presence of ostensive cues and the state of epistemic trust have a positive effect on infants and children’s learning in general.

Over the past decades, a large number of empirical studies had been conducted to investigate the effect of ostensive cueing and epistemic trust on infant learning. In order to bring the empirical studies together to give a unifying illustration, in her doctoral thesis, Fillingham (2018) conducted a meta-analysis based on a systematic literature search on the PSYCinfo database. She replicated the search strategies from Shneidman and Woodward (2016)’s systematic review, and searched terms related epistemic trust and studies focused on children 5 years old and younger with the following inclusion criteria (Fillingham, 2018, p. 18):

1) Ostensive cueing was experimentally manipulated
2) A non-ostensive cueing control condition was included, where the demonstration was still intentional and not accidental
3) A dependent variable that measured learning quantitatively (i.e. correct vs incorrect) was used
4) Participants were typically developing children aged between 2 months and 5 Years

34 out of 783 articles met the inclusion criteria and were reviewed in the meta-analysis. It was reported that these ostensive cues studies investigated children across various
ages and developmental abilities. A majority of these researches were cross-sectional studies conducted in a laboratory setting where ostensive cues were experimentally manipulated under different learning paradigms.

Two types of learning have these paradigms investigated – 1. Concept formation (e.g. Novel words & object association, category learning), and 2. Imitation of novel action. Attempts to pass on these two types of knowledge are most commonly found in parent-infant pedagogical situation. These studies aimed at tapping into different mechanism of information processing or cognitive bias that might be influenced by ostensive communication.

In concept formation paradigms, for example, through activation of the presumptions (cognitive biases) of referent and relevance, the theory suggests ostensive cues might have an impact on attention, hence the likelihood of retaining the novel concept. The learning outcomes are therefore measured by the level of attention (e.g. Duration or number of looking at the target object (Senju & Csibra (2000), or the success rate of correctly associating the novel label with the novel object (e.g. Floor & Akhtar, 2006)).

Through activation of the presumption of genericity, the use of ostensive cues might also have an impact on the choice of information (e.g. permeant property vs transient episodic property) processed by infant. In such case, the learning outcome is measured by how likely the infant generalises the property to the other kind member or bases on what property (e.g. Function vs colour) does the infant categorise objects (e.g. Butler & Markman, 2014).

In paradigms of novel action, through activation of the presumption of relevance, ostensive cue might cause infant to bypass the efficiency evaluation, hence, to increase the likelihood of reproducing the novel action despite it looks irrational. The learning outcome is therefore measured by whether the infant demonstrates over-imitation or not (e.g. Shneidman, Gaskins & Woodward, 2016).
Fillingham analysed the data using a random effects model and used version 3 of Comprehensive meta-Analysis (CMA; 2014) to calculate the Hedge’s $g$. She concluded that there is a significant positive small-to-medium effect (Cohen, 1977) of ostensive cueing of $g = .407$ (95% confidence interval [CI] = .261-.544).

For the purpose of exploring the impact of each individual ostensive cue on infant learning, this paper will review some studies which were included in Fillingham’s meta-analysis in further detail, as well as looking at other relevant and more recent literature that builds on but is beyond the scope of her study.

**Eye contact**

De Bordes et al. (2013)’s study was included in the meta-analysis that investigated the impact of eye contact on 20-month-old infant’s gaze following behaviour. An actor was presented in a video with one of the four conditions, which were 1. Initial eye contact (IEC), 2. Eye salience (ES), 3. Mother salience (MS) and 4. No eye contact (NEC). The actor then shifted her gaze to the object next to her. The eyes in ES and the mouth in MS condition were covered with moving and blinking dots and therefore not visible to infant. Those animations were designed to draw infant’s attention to that particular area of the face. The face of the model in NES was covered by a colourful amination of a flower, and therefore also not visible to the infant. IES was the only condition that was considered as ostensive. The results suggested that infants were much less likely to follow the model’s gaze in the MS and NEC. Consistent with the theory, the occurrence of gaze following in the IES condition was significantly higher than the MS and NEC condition. The study also reported that the ES condition did not differ from IES in that respect. De Bordes et al. (2013) concluded that gaze following by infants might not be driven by the referential expectation induced by the mutual eye contact, but simply by some general attention mechanism. Although this study was less conclusive as to whether eye contact induced referential expectation can lead to gaze
following, it was not robust enough to fully reject the theory. After all, attentional mechanism and referential expectation are not mutually exclusive.

Striano et al. (2006) was another study included in the meta-analysis. It investigated the effect of eye contact and joint attention on 9 and 12-months-old infants’ learning in relation of their attention and object processing. In the joint attention condition, the experimenter made eye contact with the infants, showed a positive facial expression and said, “Oh, nice!” before speaking about a novel toy. When introducing the toy, the experimenter moved her gaze back and forth between the infants’ eyes and the toy. In the non-ostensive condition, on the contrary, the experimenter formed no eye contact with the infant and only looked at the toy when speaking about it. The result reported was consistent with the proposal of the theory: In the presence of ostensive cues, infants tended to follow the experiment’s gaze and look at the toy. 9-months-old infants looked at the novel toy significantly longer than those in the non-ostensive condition, although only an insignificant difference in looking time was found in 12-months-old infants.

IDS

Ma et al. (2011)’s study was also included in the meta-analysis. It was one of the experimental studies that examined the effect of IDS on 21-months-old toddlers’ word learning. In this study, word learning was measured by recording the time the toddlers spent on looking at the correct object after being taught what it is called. Infants learned the word-object association through watching a video with the pictures of the objects and a voice from the experimenter saying: “Look here, it is a modi (made up name for the purpose of the study), see the modi, that’s the modi. Look what the modi is doing? Now the modi is going over there. Where is the modi going? ...” (Ma et al., 2011, p. 190) either in IDS or ADS. 21-months-olds infants were found to only learn words reliably when learning them in the IDS
condition. The effect was particularly salient for those who still had a relatively small vocabulary. That indicates that IDS in fact facilitates word mapping at the early stages of lexical acquisition, though this effect is diminished as the vocabulary size grow. 27-month-old toddlers who were tested with only ADS under the same paradigm were found to be already able to learn novel word-object association reliably from ADS.

Going beyond Fillingham’s review, Zangl and Mills (2007) reported a similar finding in a neurocognitive study, where familiar words spoken in IDS triggered increased activity in infants’ brain, compared to when words were spoken in ADS. This effect gradually diminished as infants grew older and became more experienced in word learning. These results suggest that the role of ostensive cues might change during an infants’ development.

In addition, in 2015, Hernik and Csibra investigated the effect of IDS on infants’ learning regarding artefact functions and the acquisition of generic knowledge about enduring kind-relevant properties of the referents. In experiment one, they employed the paradigm of violation-of-expectation looking time. In their study, infants watched a video that demonstrated the functions of two novel artefacts (1. banana-peeler that peels the skin of a banana and 2. banana-healer that “unpeels” a banana after it has been peeled). The video came with an audio track in which a female voice said “Hello baby, hello. I will show you something. Look! See? Done.”(Hernik & Csibra, 2015, p. 180) in IDS. In the testing phase, infants were showed an outcome that’s either congruent or incongruent with the function of the artefact. Infants were found looking significantly longer at the incongruent outcome.

That suggested that infants were able to form a representation of the function of the artefacts and were surprised when the outcome was inconsistent with the knowledge they had acquired. In a second experiment, an group of infants in a non-ostensive condition was compared to the first group. These infants watched the same demonstration videos, but it came with ADS prosody features and language content. The looking pattern in the non-
ostensive group was found to be significantly different from that of infants who were ostensively addressed. That suggests that the representations formed in those two conditions were qualitatively different. To understand this difference, in a last experiment, infants from each condition in addition to the original demonstration videos watched a video presenting counterevidence (i.e. banana-peeler unpeels the skin whilst banana-healer peels off the skin) before the testing phase. The result indicated that the knowledge acquired by infants who were exposed to IDS was much more resilient and enduring than of those in the non-ostensive group. They, therefore, were looking much longer at the incongruent outcome despite presented with information that challenged their acquired knowledge.

**Contingent response**

Within Fillingham’s review, none of the studies included focused solely and explicitly on contingent response. This ostensive cue might have been implicitly included in the interaction with infant participants. For example, in Williamson and Brand (2014)’s study, the experimenter was described as “emotionally engaging”, which is very likely to have implied contingent responses to infant’s emotional state.

Shimpi et al. (2013) conducted a study to examine the effect of familiarity of the person modelling, as well as ostensive communication on 18 to 24-months-infant’s imitation behaviours. In one of the experiments concerning the sole effect of familiarity, infant participants observed the experimenter interacted with a third person with a variety of familiarity (Unfamiliar experimenter, infant’s mother, unfamiliar toddler). The result suggested that infants in all conditions were able to learn the novel actions. Of note, the result provided counter evidence to the theory that ostensive communication does not play a significant role in infant’s imitation learning if they can learn from observing a demonstration of the experimenter to another person. However, further review of the experiment procedure provided a potential explanation. Prior to the training phase, the very experimenter had
extensive interaction with the infant in the warm-up phase: “During the warm-up, the experimenter engaged in joint attention with the toddler, alternating gaze between the toddler and the shape sorter, engaging in turn taking with the toddler by offering and taking the shapes and helping the toddler to put the shapes in the sorter…” (Shimpi et al., 2013, p. 313). Although these ostensive cues were not used in the training phase but might have well “contaminated” the non-ostensive condition. On the contrary of disproving the theory, the result might have demonstrated the potential enduring effect of ostensive cues on infant learning.

Going beyond Fillingham’s review, in a recent study, Tsuji et al. (2020) tested the effect of contingent response on infants’ words learning. 12-months-old infants were randomly assigned to the contingent and referential (CR+) condition and the noncontingent and nonreferential (CR-) condition. In the CR+ condition, infants watched a video showing a virtual agent who tried to teach them a novel word-object association. This video came with a gaze-contingent eye-tracking system; the movement of the presented object was contingently responsive to the infants’ gaze. The object stopped moving if the infant looked away. Whilst in the CR- condition, the movement of the object was pre-programmed and not responsive to infants’ gaze. The findings indicated that infants in the CR+ condition had a better learning outcome. This result not only supports the proposal in the theory of natural pedagogy that infants are sensitive to contingent responsivity even from a non-human source. It also provided evidence that contingent response facilitates an infant’s learning.

**Critical evaluation and discussion**

Although those studies mentioned above and Fillingham (2018)’s review reported encouraging evidence that confirmed the hypothesises proposed by the theory, it is important
to take a critical perspective towards the existing literature in order to develop this theory further. The following sections of this paper will, therefore, further review studies that were included in the meta-analysis and look beyond those, hoping to inspire direction for potential new developments and research.

**Lack of unifying behavioural operational definition**

Sage and Baldwin (2011), a study included in Fillingham (2018)’s meta-analysis, tested whether ostensive cues had an effect on infants’ imitation learning of a causally relevant tool-use demonstration. They found that infants who had been exposed to ostensive cues displayed a superior reproduction of the tool-use sequence over those in the non-ostensive condition despite the fact that they displayed an equivalent level of attention toward the demonstration. In the illustration of the procedures, they reported: “…she directed pedagogical cues to infants whilst demonstrating the tool-use sequence. These cues included eye contact, gaze shifting, name referral, infant-directed speech and an engaging emotional expression...” (Sage & Baldwin, 2011, p. 837). Although there might be some degree of consensus of how these cues might look like, it is considered as insufficient for meeting the standard of a behavioural operational definition for empirical scientific research, especially when it comes to behaviours like IDS. There could be significant variations in its very quality (e.g. pitch, pace, number of repetition) from individual to individual.

In another study that was included in the meta-analysis, Seehagen and Herbert (2010) reported a similar result in which ostensive cues significantly facilitated infants’ imitation learning. Ostensive cues were manipulated and recorded in a video. Illustrating the procedures, they explained that “…mothers were asked to narrate the event in a style they would normally use at home when showing their infant something novel, with the exception that they were asked not to use their child’s name...” (Seehagen & Herbert, 2010, p. 170). It
was not clear how the “mothers’ style” was defined, and whether essential ostensive cues (e.g. eye contact, contingent responses, IDS) stated in theory were included. There were also various studies which employed verbal referential instruction or narration towards the infants like “Look! I am going to show you something interesting, we are going to play a tennis ball!” (e.g. (Kupán et al., 2017, p. 189)). Despite reporting consistent positive results, there was no clear indication of whether these instructions were said in IDS or ADS, and whether some contingent response was being part of the interaction. Most of the other studies included in the analysis have demonstrated similar limitation.

This observation creates to concern and raises the question, whether the lack of unifying behavioural operational definition of ostensive cues is a common issue in the existing literature. In order to explore this notion, some further reviews have been conducted beyond the literatures included in Fillingham (2018)’s meta-analysis. For example, in a longitudinal study focusing on exploring the effect of IDS on infants’ language learning, Dave (2014) found evidence to be consistent that IDS played a critical role in infants’ language development. A significant positive association was identified between the exposure of IDS and toddlers’ expressive vocabulary at 18 months. However, in this study, IDS was not defined with its qualitative features in regard to the fundamental frequency, pace and grammatical structure, but focused on repetition and speech content. That included guiding questions, commands, as well as guiding speech. The latter directed the infants’ attention to objects and utterances used to direct the child’s attention to the mother. As the infants continue to develop their capacity to understand language content, it would make sense to not only consider the prosodic features of IDS matter but also look at what being said in IDS as an important component. Yet that aspect is rarely included in IDS research. While being mindful of the differences in the IDS operational definition, another interesting finding of this study was that IDS content and repetition tend to change across the infant’s
development, and this change has a positive correlation with the infants learning. This finding, echoed with Ma et al. (2011) and Zangl and Mills (2007)’s study mentioned in previous sections, brought attention to the potential developmental nature of IDS and other ostensive cues. This aspect might have often been overlooked in the cross-sectional studies.

The issue that a unifying operational definition is lacking seems to have also appeared in contingent responsivity studies. As much as contingent responsivity is one of the most important ostensive cues in communication and it often co-occurs with other cues (e.g. joint attention, turn-taking in utterance, IDS, facial expression), in most of the studies reviewed in Fillingham’s analysis, it was either not included explicitly as part of the ostensive condition, or it was only very loosely defined. Although it reported supporting evidence, Tsuji et al. (2020), a study mentioned previously, studied contingency only in the narrow context of visual stimulation. In fact, instead of defining contingent responsivity loosely as merely turn taking, Klein et al. (1987) conceptualised it with five discrete caregiver behaviours by the caregiver: (1) intentionality and reciprocity, (2) mediation of meaning, (3) transcendence, (4) mediated feelings of competence, and (5) mediated regulation of behaviour (Klein et al., 1987, p. 117). Landry et al. measured relatively recently, in 2006, contingent responsivity in another four distinct behaviours: “(1) contingent responding, (2) emotional-affective support, (3) support for infant foci of attention, and (4) language input that supports developmental needs” (Landry et al., 2006, pp. 627-628). There were also the differences in how these behaviours were measured, for example. Mahoney (1999), for example, proposed twelve essential of the caregiver behaviours that would capture contingent responsivity, and these individual behaviours were rated in a 5-point Likert scale. In Klein and Landry’s studies, discrete behaviours were instead captured in a frequency count. Clearly, in the current body of literature, there is a lack of consensus in the operational definition of contingent responsivity and even in broader terms, ostensive communication.
**Real-life phenomenon or laboratory artefact?**

Considering the very compelling evidence reviewed earlier in this paper, there can be no doubt that ostensive communication and cultural learning exist in real-life interactions between caregivers and infants. However, when ostensive cues are studied in cross-sectional laboratory design, how comparable are they with the ostensive cues used by parents? Therefore how valid are they in regard to the cultural learning taking place in real life situation? For example, it is very common to find prescribed speech content or gestures in ostensive communication studies (e.g. Egyed et al. in 2013; Ma et al., 2011; Sage and Baldwin, 2011; Shneidman et al., 2014). Let alone that these cues were not very standardised - some studies included verbal referential cues (e.g. “Look!”, “Watch this”, “Do you want to see this?”) or non-verbal referential cues (e.g. pointing) but others did not - most of these cues were shown by the experimenter, a video or audio soundtrack or even by an virtual agent rather than the caregivers. In fact, studies have shown that it is the natural parental speech that makes all the difference for infants. The specific features of natural DIS produced spontaneously by parents cannot be reproduced in a laboratory context by experimenters and actors (Knoll & Costall, 2015). In their metanalysis review, Spinelli et al. (2017) concluded that IDS produced by actors or non-parental adults did not exactly correspond to the IDS infants encountered in actual social exchange with their parents. That was echoed by Knoll and Costall (2015) who stressed the importance of investigating IDS in naturalistic observational studies.

This concern is not only limited to studies that focused on IDS but is also valid for how to look at other ostensive cues. We might also wonder in how far the contingent responses created artificially in a laboratory corresponds to naturalistic parent-infant
interaction. Multiple studies concerning the role of IDS in early language perception and acquisition found that IDS rarely occurs in isolation but is often accompanied by synchronisation of other cues, including exaggerated visual facial expressions (Chong et al., 2003; Green et al., 2010; Shepard et al., 2012) and referential gestures (Gogate et al., 2000; Gogate et al., 2006). This multimodal stimulation provides intersensory redundancies that facilitate the learning process (Bahrick & Lickliter, 2002; Gogate & Bahrick, 1998; Gogate et al., 2006). These studies had not examined contingent responsivity explicitly, but it is reasonable to assume some of these cues were natural responses to the infant’s immediate state in the context of pedagogy, and therefore contingent in nature. In fact, we did already learn about the complex quality of contingent responsivity from Klein et al., (1987), Mahoney (1999) and Landry et al. (2006) in their efforts to capture various aspects of such complex construct. Contingent responses experienced by infants in real-life are usually co-occurring with other cues. Or they take place in a multimodal context which includes responses to vocalisation, behaviours and affective state via eye contact, joint attention, facial expression, IDS and physical touching.

In addition, as discussed in the previous section, conceptually speaking, the use of contingent response, or ostensive cues in general, is likely to be influenced by parental sensitivity and the capacity to mentalise. However, relevant concepts, such as parental reflective functioning (Fonagy et al., 1991) or mind-mindedness (Meins et al., 2003) were rarely taken into account when considering the operational definition of ostensive cues. For example, contingent responsivity is likely to be applied differently by individual and therefore manifests itself as a naturally occurring continuous variable in normal interaction. In experimental designs, however, it is often operationalised in a binary manner.
Therefore, experimental studies in a laboratory setting are insufficient when it comes to revealing the true essence of contingent responsivity, and therefore when addressing the role of ostensive communication in cultural learning.

**Developmental quality of ostensive communication and more questions**

Multiple studies reviewed above reported on the developmental quality of ostensive communication, such as Dave (2014), Ma et al. (2011) and Zangl and Mills (2007). According to them a mother’s use of ostensive cues changes across the infant’s various developmental stages. Similarly, Gogate et al. (2000) learned that the mother’s use of synchronous IDS and gestures decreases as the infants’ language capacity grows. To be precise, mothers of 5-8 months old “prelexical” infants use multimodal synchrony significantly more often than mothers of 9-17 months old “early lexical” infants, and the frequency drops further for mothers of 21-30 months old “advanced lexical” infants. Intuitively, mothers are usually adaptive to infants’ functioning and development and adjust their strategies accordingly (Saint-Georges et al., 2013). Therefore, ostensive communication is not a static pheromone but dynamic in nature. In their study of contingent responsivity, Keller et al. (1999) found that, although mothers tend to respond to their infant in a contingent manner in general, there were observable individual differences, where mothers who reacted with high contingency toward infants’ verbal expression did not respond contingently toward their nonverbal communication, or vice versa. This is then posing some interesting questions that are yet to be answered by the existing literature. For example, what ostensive communication as a whole in the context of pedagogy like in real life? How are different ostensive cues are interacting with each other and in which way do the mechanism of cultural learning changes through the course of this process? Are there individual differences in terms of how ostensive cues are used? In addition to frequency, is there a
difference in the quality of ostensive cueing? If so, what constitutes a better IDS, contingent response or ostensive communication as a whole?

Proposal for future research

This paper has discussed the proposition that the capacity to learn effectively from our ancestors is the core of what makes human distinct from other species. In human society, knowledge and information have been efficiently passed on from one generation to the next through cultural learning. This process has allowed our civilisation to continue developing, advancing and evolving. However, the increased complexity and lack of transparency of information in modern age have created the problem of learnability, which might endanger this process of cultural evolution. It is therefore important to develop a deeper understanding on this matter.

Among many other models, Csibra and Gergely (2006, 2010) have proposed a comprehensive account of cultural learning in their theory of natural pedagogy. In order to understand what unique qualities of human beings have prepared us as a species to become so receptive to new and complex knowledge, Csibra and Gergely brought our attention to early infancy. They proposed that human infants are genetically and biologically programmed to be sensitive to ostensive cues, such as eye contact, infant-directed speech and contingent response, used by their parents in early communication. This paper reviewed empirical research of the past decades. There is compelling evidence to support the claim that human infants show innate capacity to detect and respond to ostensive communication. Csibra and Gergely have further argued that, in response to ostensive cues, a human infant will switch into the state of basic epistemic trust; a special state that entails essential cognitive biases which allow infants to learn in a very efficient manner. These cognitive biases are (1) referential expectation, (2) the assumption that knowledge is universally shared by other
people of the cultural community and a bias to attribute or encode the presented information as generalisable knowledge, and (3) presumption of relevance. The relationship between ostensive cues and these cognitive biases have been demonstrated in a large number of studies. We also learned about the critical role of these biases in facilitating an infant’s learning of epistemically opaque cultural knowledge.

Encouraging evidence in support of this theory has been reviewed and discussed in this paper. Whilst there is strong empirical basis for natural pedagogy, several issues, however, have also been identified and discussed. These include (1) a lack of unifying behavioural operational definition of ostensive cues in the field, (2) and understanding that ostensive communication studied in the laboratory and with cross-sectional designs might not fully correspond to what infants experience in natural interaction with their parents and (3) a lack of studies focusing of the developmental feature and individual difference of ostensive communication. These issues contribute to the existing evidence falling short of offering robust support to the theory.

It is, therefore, suggested that future research should employ the methodology of a longitudinal study. Also it should investigate ostensive communication between mother and infant in a pure or semi-naturalistic setting (e.g. the mother is tasked with teaching her infant new information in a way that closely resembles real-life experience) in order to address the second and the third issue.

This cannot be achieved without addressing the first issue, the problem of operational definition. It is necessary to have a well-established and experientially validated tool that can capture quantitative and qualitative data of the ostensive communication reliably and thoroughly. However, no such tool is currently available in the field. In light of this research need, this paper calls for a development of ostensive cue rating system that has a behavioural operational definition and is tailored to be used in naturalistic observation.
References


Part Two: Empirical Paper

Development and validation of Ostensive Communication Coding System
Abstract

Aim: (1) To develop the Ostensive Communication Coding System (OCCS). (2) To establish the psychometric properties of the OCCS.

Method: Establish the behavioural operational definition and measurements of four ostensive communication – Eye contact, Name addressing, Infant Directed Speech and Contingent responsivity. Apply the OCCS to code videos of mother-infant interaction in two learning paradigms. Due to the occurrence of Covid-19, data collection was interrupted. The second aim of the study could not be fully achieved because of insufficient sample size. 14 infants aged 16 to 22 months took part in the study with their mother. Their interaction in the two learning paradigms are coded with OCCCS.

Result: Interrater reliability of OCCS was assessed via intraclass correlation coefficients. Interrater reliability was excellent for all the scale components (> .90). Construct validity was tested with Parenting Stress Index, Fourth Edition Short Form (PSI-4-SF) and The Parental Reflective Functioning Questionnaire (PRFQ) via quasi-qualitative method, individual patterns of scores were observed. The results provided supportive evidence of the construct validity of OCCS.

Conclusion: The OCCS is likely to be a reliable and valid instruction to assess ostensive communication in natural pedagogical situation. Implications for future development are discussed.
Introduction

The continuous development of human civilisation very much depends on the process of cultural evolution (Heyes, 2016). Cultural learning in everyday life is the building block of this essential phenomenon. As many developmental researchers suggested, ostensive communication in early life between mother and infant lays the foundation for cultural learning. The theory of natural pedagogy proposed by Csibra and Gergely (2006) and Gergely and Csibra (2005) offers a comprehensive framework to understand this matter. After decades of exploration, the accumulation of supportive evidence that have been reported in a significant number of empirical researches – mostly experimental studies - have established a substantial body of literature about ostensive cues (i.e. eye contact, infant-directed speech (IDS), and contingent response). Experimental manipulations have been essential in demonstrating the causality between them. These studies have successfully shown the powerful effect of ostensive cues on infant learning, and furthermore, how the activation of the state of epistemic trust and three cognitive biases (i.e. presumption of referent, presumption of genericity, and presumption of relevance) has ensured the intergenerational transmission of complex and cognitively opaque cultural knowledge (e.g. Brugger et al., 2007; Csibra & Volein, 2008; Egyed et al., 2013; Gliga & Csibra, 2009; Jennifer et al., 2008; Király et al., 2013; Nielsen, 2006 See volume one for a review). Considering the central position of ostensive communication in infant learning, insufficient or non-optimal use of ostensive cues is likely to hinder this process of cultural evolution. In fact, prior researches provided reliable evidence that mothers with mental health difficulties (e.g. Major depression) use less effective ostensive communication (e.g. IDS) and their infants show poorer learnings (Kaplan, Bachorowski, Smoski, & Hudenko, 2002; Kaplan, Dungan, & Zinser, 2004).
Lack of unifying operational definition

However, there are indispensable limitations within the evidence presented in the literature. That makes them fall short of providing robust support to the theory. The first issue is a lack of unifying operational definition of ostensive cues among these studies. Most studies which have claimed to study ostensive communication tend to have a very loose definition. For example, (Butler & Markman, 2014) operationalised ostensive communication as "E1 (experimenter) looked at the child, called the child by her name and said, "Look!" (Butler & Markman, 2014, p. 480) Sage and Baldwin operationalised it as "… eye contact, gaze shifting, name referral, talking in infant-directed speech about actions she wanted to demonstrate for infants, and an engaging emotional expression." (Sage & Baldwin, 2011, p. 830). Even in studies that focused on one particular ostensive cue, their operational definition still varied significantly. For example, while IDS was conventionally conceptualised with its qualitative features in various studies (i.e. prosodic contours, rhythm and linguistic structure) (e.g. Hernik & Csibra, 2015). (Dave, 2014) defined IDS in terms of guiding speech content (e.g. command and question) and its patterns of repetition.

Similarly, in terms of contingent response, when Tsuji et al. (2020) investigated the effect of contingent response on infant's word learning, they merely operationalised contingent response through visual response, while (Mahoney, 1999) identified twelve relevant behaviours (please see volume one for a detailed review). The lack of a standardised operational definition in the field makes it difficult to compare the evidence of these studies. Hence, the validity of the conclusion drawn from these results is inevitably undermined.
Real-life phenomenon or laboratory artefact?

Besides, investigating early ostensive communication in an experimental and laboratory setup faces a significant challenge. What happens in the laboratory might not sufficiently correspond to real-life. Knoll and Costall (2015) and Spinelli et al. (2017) maintained that IDS pronounced by non-parental adults could not precisely correspond to the IDS infants encounter in naturalistic social exchange with their parents. Indeed, different from eye contact, which is relatively open to experimental manipulation as a binary variable, spontaneous and natural IDS produced by parents has complex and delicate features that are far more challenging to reproduce. For instance, although there is a consensus that prosodic contour in IDS is characterised by a higher pitch, making it a distinctive feature, there is also a qualitative difference of prosodic contour from individual to individual. It is almost impossible for an actor or experimenter to reproduce this subtle difference for every infant. Moreover, the speech content in the experiment is often pre-determined. In their study, Ma et al. (2011) examined the effect of IDS on 21-month-old toddler's word learning. The experimenter used IDS to say to participants: "Look here, it is a modi (made up name for the study), see the modi, That's the modi. Look what the modi is doing? Now the modi is going over there. Where is the modi going? ..." (Ma et al., 2011, p. 190). Without doubt, the speech content used by mother in natural pedagogy situation has far more variety than this mechanical statement.

Contingent responses tested in an experimental setting is facing a similar limitation. As mentioned earlier, Tsuji et al. (2020) only narrowly operationalised contingent response in terms of visual signal and response. In a naturalistic setting, not only do infants tend to produce far more variety of signals (e.g. utterance, body movement, facial expression) (Warren & Yoder, 1998) but mothers also contingently give responses in a multimodal manner. Some mother might respond to the same signal through speech while others through
physical touch (Keller et al., 1999). The diversity of signals and contingent behaviours in ordinary interactions go way beyond what an experimental design can potentially control and manipulate.

**Co-occurrence of ostensive cues and its developmental quality**

Because much understanding about ostensive cues has been derived from studies that focused solely on a single ostensive cue, this literature is likely to only reflect a fraction of the phenomenon, as in naturalistic exchange situations ostensive cues do not occur as isolated phenomenon. They tend to co-occur within a short period of time. It is not uncommon to observe a mother using different ostensive cues together. For example, when teaching the name of a novel object, the mother might first look into her child's eyes to establish an eye contact, then addresses the child by his/her name while she picks up the object and shows it to her child. After the child produces an utterance, the mother might respond by clarifying the child's intention (e.g. you want to play too?) and giving the object to the child to play around. At the same time, she labels the object and expands his/her understanding by further describing different features of the object in IDS. The use of all these ostensive cues is loosely coordinated. Such a process is often described in studies as intersensory redundancies. However, the focus of these studies has been to establish whether relevant information is presented through different sensory stimulations rather than asking whether these presentations are ostensive. Human communication is multisensory. It includes visual information (e.g. facial motions, gestures), tactile information (e.g. physical touching) as well as auditory information (Jouanjean-L'Antoëne, 1997); (see Meltzoff & Kuhl, 1994). Maternal communication makes no exception. Multiple studies have reported that maternal communication to infants typically contains redundant information across senses. For example, Goldring (1991) contended that mothers tend to teach their infants word-referent
relations by naming the object and simultaneously using gestures (e.g. pointing) or touching the infant. Messer (1978) and Gogate et al. (2000) reported similar findings. There was no mention in these studies whether the mothers used eye contacts, name address or other cues in the process, but it would be a surprise if they were not. Young infants also showed preference of multimodal communication over unimodal. Kuhl and Meltzoff (1988) found that 4-month-old infants showed a visual preference for presented vowels when they were accompanied by facial expression (mouth shape) and vowel sound. Intersensory redundancies have been proven to enhance infants' language perception and acquisition (Bahrick & Lickliter, 2002; Gogate & Bahrick, 1998; Gogate et al., 2006). Not every sensory stimulation in the learning process is ostensive. It is therefore a question worth asking: how likely do ostensive cues co-occur? What is the effect of co-occurrence on cultural learning? The impact of each ostensive cue has been studied extensively. Yet, no empirical study investigated the co-occurrence of ostensive cues so far.

Besides, multiple studies reported the developmental quality of ostensive communication (e.g. Dave, 2014; Keller et al., 1999; Ma et al., 2011; Zangl & Mills, 2007). The use and the effect of ostensive cues tends to change across an infant's developmental stage. This is posting another interesting question: how different ostensive cues are interacting with each other in the developmental trajectory?

**Unstandardized paradigms**

Fillingham (2018) reported in her meta-analysis that there is considerable variation in learning paradigms used in testing the effect of ostensive cues across studies. These paradigms included ‘Category learning’, ‘Learning actions – novel objects/actions’, ‘Learning actions – unconventional actions’, ‘Object-directed looking’, ‘Word learning’ and ‘Miscellaneous’ (Fillingham, 2018, p. 23) The reason why many different paradigms were
used is probably reflecting the fact that, in real life, any mothers would need to teach a range of different kind of knowledge to their children, these common knowledges include but no limited to label and function of objects, novel actions, culture practices, etc. Different paradigms, however, are likely to result in different effect of ostensive cues. Evidence drew from unstandardized paradigms are less comparable and therefore it is difficult to derive reliable conclusion concerning the effect of ostensive communication from the body of literature.

**The necessary instrument**

To address these limitations identified above, extracting data of ostensive communication from a natural pedagogical setting is an important step to take. However, there is no such instrument available in the field that allows the researcher to capture the data of all ostensive cues in real-life interaction between the mother-child dyad. Codification of behaviour is a widely used and reliable methodology for naturalistic observational studies, therefore, developing a behaviours coding system with standardised behaviour operational definitions of all ostensive cue can be a viable solution. Picking up from the point in the section above, a mother is likely to teach different kinds of knowledge to her child in day to day interaction. Intuitively, she might use different ostensive cues subject to what knowledge she is teaching. To make universal comparison of the effect of ostensive cues possible, this instrument will need to be applicable to most if not all pedagogical situation and experimental paradigm.

Once the instrument is fully validated, its application can also potentially extend to different clinical and research settings. For example, it can potentially be used as an assessment tool to identify individuals who might have difficulties in using ostensive communication and require intervention. The assessment result can also provide valuable
information that helps clinicians to formulate and operationalise the interventions tailored to the individual, as well as to evaluate its effectiveness. Data accumulated in this process can provide useful guidance for developing services that are customised for this specific population.
Present study

In light of this need for a new instrument, the present study will aim to develop a reliable and validated coding system of ostensive cues through the following six steps:

1. Establish the empirical basis of the ostensive communication coding system (OCCS) through detail revision of how each ostensive cue has been defined and measured in empirical studies.

2. Construct the standardised operational definition of each ostensive cue base on the empirical basis.

3. Consideration of two learning paradigms for the purpose of testing whether the OCCS can be applied for all naturalistic setting.

4. Consideration of initial examination of the coding system

5. Generating data by applying the coding system to the two learning paradigms

6. Use the obtained data for reliability and validity analysis for the purpose of establishing the psychometric property of OC
Step 1. Empirical basis of OCCR & Step 2. Behaviour operational definition

Eye contact

Eye contact is a relatively straightforward variable to be measured in an observational study. It is defined as the overlapping moment of the mother's looking into the infant's eyes and of the infant's looking into the mother's eyes (Lohaus et al., 2001). Eye contact is usually measured in frequency.

Behavioural operational definition of Eye contact in present study

Table 1.

<table>
<thead>
<tr>
<th>Scale</th>
<th>Behavioural operational definition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Eye contact</strong></td>
<td>The overlapping moment of the mother's looking into the infant's eyes and of the infant's looking into the mother's eyes</td>
</tr>
</tbody>
</table>
Many studies have described the significant differences in linguistic and paralinguistic characteristics of IDS when comparing with ADS. These differences can be grouped into four main components: (1) prosodic contour, (2) rhythm, (3) linguistic structure and (4) repetition.

1. **Prosodic contours**

   Electrolaryngographic studies have indicated that the prosodic contours of IDS is usually characterised by a higher mean in its fundamental frequency (f0 or pitch) and broader the frequency range (Fisher & Tokura, 1995; Grieser & Kuhl, 1988). The prosodic contours are very informative for an infant. It is a reliable cue of the communicative intentions, sentences type, emotions as well as the focus of a given sentence.

   In their study, Katz et al. (1996), for example, they found mothers tend to use rising contours when they try to seize the infant's attention. In contrast, they use sinusoidal-bell shaped contours when they try to maintain the already established focus and positive affect. Studies have also shown that IDS contours vary according to sentences types are a universal phenomenon across cultures and languages. For instance, mothers tend to use elevated contours for yes-no questions, falling contours for "wh" questions and commands, and sinusoidal-bell shaped contours for declarative communication (Stern et al., 1982).

   Significant difference in prosodic features were also identified across expressions of emotions, such as love, comfort, fear and surprise (Trainor et al., 2000). When teaching novel words to an infant, mothers tend to use hyper-articulated, lengthened vowels and exaggerated pitch (Fernald & Mazzie, 1991).

   In sum, there are very detectable variation in pitch and amplitude in IDS.
2. **Rhythm**

IDS usually present in a much slower tempo and rhythm. There is also longer pause between words, phrases or sentences. Given that there is a notable short-term memory limit in infant, a slower rate of speech can expand the time for the infant to process and encode information communicate by mothers (K. S. Moore et al., 2013).

3. **Linguistic structure**

Linguistically, IDS patterns are usually shorter with much simpler grammatical structure (Brodsky & Waterfall, 2007; Fernald, 1992). There also consist of more isolated, redundant and modified words and phrases (Soderstrom et al., 2008).

Linguistically, IDS is usually characterised by lexical and syntactic simplification. There is a general reduction in speech quantity and meanings (Bornstein et al., 1992; D'Oodorico & Franco, 1990; Furrow et al., 1979; Henning et al., 2005; J. R. Phillips, 1973), and the lexicon used by mothers are more straightforward, specific, concrete, repetitive, and has less variability (Brodsky & Waterfall, 2007; Broen, 1972; Ferguson, 1964; Fernald, 1992; Fernald & Morikawa, 1993).

When naming a particular object, for example, instead of using a sophisticated label, such as "truck", "taxi", mothers tend to use more general terms like "car". Mothers are also inclined to use learned syllables and duplicated short sounds to create words or phrases that are easier for infants to pronounce, such as "goo goo ga ga" (Falk, 2009; Matychuk, 2005; Richman et al., 1992).

IDS vocabulary also presented with specific lexical features, often mothers use onomatopoeias like "bye-bye", "peekaboo", "am" for food, or animal sounds (e.g. meow for the cat, woof for the dog) (Fernald & Morikawa, 1993; Gleitman & Gleitman, 1972). It is also ubiquitous to see mothers using morphological derivations and nonstandard diminutives.
that convey affection and smallness, such as "da-da" for father, "ma-ma" for mother, "tummy" for the stomach, etc. (Bornstein et al., 1992; Eva et al., 2007; Kitamura & Burnham, 2003).

Moreover, studies found that IDS contains more lexicons which concern the present rather than the past or future (Gopnik & Meltzoff, 1993; Nelson, 1973) and verbs which express concrete intention, command or permissions (e.g. to go, to look, come, let's do this) (Adi-Bensaid et al., 2015; Broen, 1972; Ravid, 2010). Also the adjectives or adverbs convey more straightforward essential qualities (e.g. good, bad, dirty) rather than complex and abstract concepts (Adi-Bensaid et al., 2015; Ferguson, 1964; Rowe, 2012).

In terms of syntactic structure, IDS is found to be shorter and syntactically simpler. There is frequent use of isolated and redundant words and phrases (Lew-Williams et al., 2011). There are fewer phrasal elements and subordinate clauses but a more frequent use of incomplete sentence and utterances that do not comply with subject-verb agreement (Broen, 1972; Snow, 1972).

All these features make IDS more perpetually simpler for infants to understand by accommodating to the immature cognitive ability of infants.

4. Repetition

Frequent self-repetition is commonly found in IDS. Snow (1972) suggested that there are two types of repetition: full repetition and partial repetition. Syntactically, full repetition is referring to the complete copy of the original speech, while partial repetition refers to an utterance in which meaning is being repeated without using identical grammatical units. Studies found that mothers tend to use a new frame for repeated phrases or rearrange basic units to form new utterances (Bard & Aylett, 2001; Curl et al., 2006; M. Tomasello et al., 2005). Mothers conveying the same intent with varying form is also defined as a variation set
There are three essential components in a given variation set: (1) lexical substitution and rephrasing, (2) addition and deletion of referential terms, and (3) reordering of constituents. For example, a mother might first say "Can you switch on the light?" followed by the partial repetition of "turn it on". In this case, the verb "switch on" has been substituted with "turn on", the "light" has deleted and replaced by "it". The form of speech has changed from a question to a command, while the overall meaning of the utterance remains unchanged.

Johnson (2013) demonstrated in his study that while a reiteration of the focus of the phrase or sentence could allow for more processing time, repeated exposure also helps the infant to map the object name with the object, enhance infant's understanding of the word, and facilitate the referent mapping. Similar findings were reported by (Brodsky & Waterfall, 2007). Partial repetition facilitates infant comprehension by familiarising them with the variability in syntax and expanding their vocabulary. Besides, when one framework of articulation does not work, mothers rearrange the components in her speech to create a different utterance with a similar meaning, a strategy that has proven helpful for infant to achieve learning goals (Hengst et al., 2010; Snow, 1972).
**Behaviour operational definition of IDS in present study**

Based on the review above, the operational definition of IDS should cover all four characteristics as shown in table 2.

Table 2.  
**Operational definition of Infant Directed Speech**

<table>
<thead>
<tr>
<th>Ostensive cue</th>
<th>Components</th>
<th>Behavioural operational definition</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>IDS</td>
<td>Integrated evaluation of the four IDS components.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>i.e. Prosodic contours, rhythm, linguistic structure, and repetition</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prosodic</td>
<td>Speaking unambiguously in a higher pitch with significant variation in pitch range.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>contours</td>
<td>Observable change in pitch subject to sentence type and expressed emotions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rhythm</td>
<td>Speaking unambiguously in a slower tempo and rhythm. There is a longer pause in between words, phrases and sentences</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linguistic structure</td>
<td>Speaking in a shorter and simpler grammatical structure with isolated words or phrases.</td>
<td>When describing Task A:</td>
<td></td>
</tr>
<tr>
<td>----------------------</td>
<td>---------------------------------------------------------------------------------</td>
<td>-------------------------</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>&quot;light, use head, touch here.&quot;</td>
<td></td>
</tr>
</tbody>
</table>

**Demonstration of using modified words or onomatopoeias**

- "car" rather than "truck".
- "bye-bye."  
- "woof" for dog

| Repetition | Clear demonstration of repetition in speech. | "Can you switch on the light?"
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Rephrase to &quot;Turn it on.&quot;</td>
</tr>
</tbody>
</table>

- **Full repetition**: a complete copy of the original speech.
- **Partial repetition**: repeat meaning but rephrase with an observable change in choices of words and constituents.
Contingent responsivity

The origin of contingent response

Compelling evidence shows that infants and young children have the innate capacity to detect contingent responses from their environment, and that they prefer contingent response over random stimulation (A. E. Bigelow, 2001; Ann E. Bigelow & Rochat, 2006; Masataka, 2003; Nadel et al., 1999; Stormark & Braarud, 2004). The earliest manifestation of contingent response from the mother is her jiggling and turn-taking behaviour in response to the infant's pause in sucking during breastfeeding (Masataka, 2003). This natural synchronisation of a mother detecting the signal from her infant and responding appropriately continues to develop throughout the infant’s development. It goes beyond the most elementary level of pausing and jiggling, and emerges in a range of interactions, contexts and mediums. These include facial expression, vocalisation and physical touching in natural playing and pedagogy (Bateson, 1979). (Dunham & Dunham, 1990) broadly defined contingent responsiveness as the maternal ability to identify an infant's cue or signal, interpret it correctly, and respond accordingly. Contingent response is therefore referring to behaviours that are temporally and functionally related to the infant's signal. This definition is useful, but the wide variety of behaviours adds to the difficulty of formulating a unifying and comprehensive operational definition of contingent responsivity. To address this challenge, we need to consider the origins of maternal contingent responsivity.

The premise for a mother to be able to respond appropriately to the signal of her infant is that she can recognise the signal, and hence, interpret what it means. This ability is closely related to maternal sensitivity and her capacity to mentalise (Gergely & Watson, 1996). A mother is considered to be sensitive if she can pick up subtle cues of her infant and interpret these cues in terms of the infant's internal state (Ainsworth et al., 1974; Murray, 1998). This requires the mother to mentalise and identify where the infant is mentally at, i.e.
to treat her infant as a conscious agent with thoughts, needs, feelings and desires different from her own (Sharp et al., 2020). For example, when a preverbal infant produces an utterance, a sensitive mother would be able first to recognise the signal and then speculate about its meaning; whether, for example, the infant is expressing excitement and desire to explore. Or is the infant is upset because of hunger or distress? Or confused and therefore asking for assistance? She will then act based on her conclusion to meet the infant's need. This process might be happening in a split moment and an automatic manner, but the timing and the choice of action entail the contingent responsivity. Therefore, for a set of behaviours to be considered a contingent response, it needs to demonstrate the quality of sensitivity and mentalisation; also the action needs to take place promptly in relation the signal.

**Operational definition of contingent responses in empirical studies**

In their study, Tsuji et al. (2020) investigated the effect of contingent response on infant's word learning. As mentioned in a previous section, they operationalised contingent response by designing a virtual agent that would act according to the infant's gaze (signal). The action of the virtual agent was contingent in the sense that it took place immediately after the infant's cue. Although the result of the study was consistent with the theory and even though infants showed superior performance in regard to word learning under the influence of virtual agent's contingent action, the empirical evidence generated was not robust enough to support the claim of the effect of contingent response in cultural learning. Clearly, the virtual agent was not mentalising what happened in the infant's mind and only replicated a fraction of what happens in naturalistic interaction between mother and infant.

In another empirical study conducted in 2013, Bornstein and Manian investigated how global judgements of maternal sensitivity are measured by marco-analytic tools (e.g. Emotional Availability Scales) (EA EA Scales; Biringen, 2000) which are assigned to levels
of specific maternal contingent response measured micro-analytically. They operationalised contingent response in a very precise way. Four discrete behaviours of the infants were operationally defined as communicative signals: (1) looking at mother, (2) looking at object, (3) non-distress vocalisation, and (4) smile. At the same time three maternal responses within 3 seconds to these signals were operationally defined as a contingent response: (1) encouraged attention to self, (2) encouraged attention to object, and (3) speech to the infant. In turn, 12 potential pairs of signal-response sequences were coded in a 48 minutes video of mother-infant natural interaction. The result of the study was insightful and will be discussed later in this paper. The operational definition, however, was insufficiently correspondent to the variety of contingent response in real life.

In the literature, there are various studies focusing on the development of indirect intervention program for infants with severe disabilities and developmental difficulties. These interventions were designed to teach mothers specific categories of behaviour which were operationalised under the concept of contingent response and also proven to facilitate early communications. These interventions required evaluation and measurement of caregiver behaviours. Although these studies focused on a different population and context, in addition to being based on little consensus regarding measurements as to how to define contingent response operationally, they provided valuable insights to address this challenge.

In their studies, Yoder and Warren (1998; 2002; 2001) developed the Responsivity Education/Prelinguistic Milieu Teaching (RE/PMT) program to provide training to parents who have children with severe disability and special communication needs. They measured contingent responsivity as a broad construct that encompassed a range of behaviours that parents showed in responding to children's displayed behaviours, including "complying with the presumed meaning of the child's immediately prior communicative message", "imitating some component of the child's immediately prior communicative act", "attempts to clarify the
meaning of the preceding utterance", and "utterances that include the main verb, noun, or function word implicit in the child's immediately preceding nonverbal or verbal communication act" (Warren & Yoder, 1998, p. 1211).

Later in 2007, Warren and Brady proposed to consider contingent responsivity as a "singular but multi-level" construct. They describe three levels: general responsivity, molar responsivity, and molecular responsivity. General responsivity is referring to necessary caregiving behaviour that meet the child's basic biological needs, such as feeding, maintaining safety and comfort. Molar responsivity is referring to "macro" maternal adjustments in response to child behaviours and level of development, such as the warm and positive affect mothers hold when interacting with their children, as well as the adjustment they make when speaking to children (i.e. IDS). This level taps into the mother's sensitivity towards the developmental change of her child. Drawing reference from previous studies (e.g. Kim & Mahoney, 2004, 2005; Landry et al., 2000), they suggest measuring the qualities of these interactions through the use of Likert-style rating scale. In contrast, they define the molecular level of responsivity as maternal behaviours that can "directly linked to change in child behaviours (Brady et al., 2009, p. 337). These include "micro-behaviours" like mother's immediate response to the child's shift of attention, utterance, or change of mood. The critical feature at this level is that the maternal response has to take place within seconds of the child's change in behaviour (Bornstein & Tamis-Lemonda, 1989; Kärtner et al., 2008). They suggest measuring behaviours at this level by its frequency, i.e. how often mother respond contingently to the child's behaviour.

In their study, Warren et al. (2010) investigated the relationship between maternal responsivity and language development in young children with fragile X syndrome. They measured contingent responsivity in both molar and molecular level by coding three separated 10-mins segments of mother-child interaction. For the molar level, six components
(1. Positive affect, 2. Warmth, 3. Flexibility/responsiveness, 4. Physical control, 5. Verbal discipline, and 6. Punitive tone) were scored using a five-point rating scale; while for the molecular level, frequencies of the following maternal responses were captured: redirects, recodes, requests for verbal complies, comments and admonishments directed to the child. They found that some aspects of the molar and molecular ratings correlated with each other. The molar scale provided information about the overall maternal interaction style, while the molecular scale shed some light on specific behaviours linked to child language development. These two scales were complementary to each other. They also found that molecular ratings were most closely aligned with children's cognitive and communication outcome.

Landry et al. (2006) developed the playing and learning strategies (PALS) program as an intervention for low-income mothers and infants with very low birth weight. They were interested in whether learning responsive behaviours would facilitate infants' language development. They employed a different approach. Instead of conceptualising contingent responsivity as a singular concept, they described four "distinct but conceptually-related behaviours" (Landry et al., 2006, p.629) that made up of 12 variables:

A. Contingent responding (Measured in 5-point Likert rating). This is referring to a three-term chain of events in which the infant elicits a signal, the mother responds in a timely and sensitive manner; therefore, the infant experiences his/ her need is met. The key features to evaluate here are the degree of the promptness of the maternal response, and its appropriateness as to the infant's abilities.

B. Emotional-affective support (Measured in 5-point Likert rating). This scale has five sub-categories of behaviours which capture the positive and negative sides:

   a. Positive affect. e.g. smiling, laughing, facial animation
b. Warm sensitivity. e.g. physical affection, enthusiasm, positive tone

c. Restrictiveness. e.g. physical or verbal interruption of activity in which infant has already engaged

d. Physical intrusiveness. e.g. abruptly taking objects away, physical expression of impatience

e. Harsh voice tone

C. Support for infant foci of attention (Measured in Frequency). This is referring to behaviours that provide a structure or scaffold for an infant's immature attention and cognitive capacity, which is thought to facilitate higher levels of learning and self-regulation. These include:

   a. Positive behaviours: encouraging joint engagement, maintaining attentional focus, and allowing the infant to take the lead in the interaction
   b. Negative behaviour: redirecting infant's attention.

D. Language input (Measured in Frequency). This is referring to verbal inputs that facilitate early language development. This scale has four sub-categories of behaviours:

   a. Verbal scaffolding. Verbal prompts that provide conceptual links between object, person, actives or functions.
   b. Labelling objects
   c. Labelling actions
   d. Verbal encouragement. e.g. praising.

The interactions between the mother and the infant were videotaped in a naturalistic living room situation for 15 minutes, and in a toy playing situation for 10 minutes. The videos were subsequently coded. Landry et al. (2006) concluded that the estimate of the total
variance across the 12 variables was 51.4% in living room situation, and 54.4% in toy play. The consistency in the structure of factors provided empirical support for four aspects of responsiveness. They also reported that, in general, the PALS increased the quality and frequency of positive behaviours and reduced the intensity and frequency of negative behaviours. That was positively associated to the development of infant's language ability.

Similarly, building on the theory of Cognitive Modifiability and Mediated Learning Experience (MLE) Feuerstein (1980); Feuerstein et al. (1979); Klein et al. (1987) developed the Mediational Intervention for Sensitising Caregivers (MISC). Klein et al. suggested that young children are equipped to experience some direct learning through their senses. Still, learning of cognitively opaque cultural knowledge would be beyond their reach with their immature capacity. Yet, learning of such knowledge becomes possible when it is mediated by an adult who understands the child and is able to proactively make a delicate adjustment or transformation of the knowledge which aligns with the child's intentions, preferences, interests and capacities. There is a significant overlap between the proposition of MLE and the mentalising stance (Sharp et al. 2020), which taps into the crux of maternal contingent responsivity.

MISC was designed to teach mothers, who have children with severe disabilities, five discrete categories of behaviours that were identified to be related to contingent responsivity:

A. Focusing. Any behaviours that are communicating the intention to teach and is directed towards bringing a change in the child's perception or behaviours. (e.g. "Look! Do you see the car?"

B. Affecting/Exciting. Any behaviours that express excitement, appreciation or affect in relation object, concepts or any kinds of knowledge that would in term excite the
child's about the potential learning experience. (e.g. "wow (with facial expression)!
This is a cool car!")

C. Expanding. Broadening of the child's cognitive awareness through behaviours like
describing, explaining, clarifying or comparing that goes beyond what is necessary to
satisfy the immediate needs of the child. (e.g. "Look! This is a green truck! It has four
big wheels! (lift the truck model, pointing at it) We saw it yesterday too. This is what
your daddy drives at work.")

D. Rewarding. Any behaviours that communicate the adult's satisfaction with a child's
behaviours, e.g. Praising.

E. Regulating behaviours. Regulation of the child's behaviours concerning the learning
experience through behaviours like modelling, demonstration, verbal suggestion.

Klein et al. (1996) reported evidence that the children of mothers who learned these
behaviours through MISC demonstrated a significant gain in their cognitive and language
skills compared to children of mothers who were in the control group.

Current literature lacks consensus about which discrete categories of behaviours make up
the concept of contingent responsivity. Some of the differences might be originating from the
fact that each of these scales has a different target population, context, hence specific focus.
For example, although PALS was designed for mothers with low-income family and infants
with very low birth weight, who have no significant developmental difficulties. At the same
time, MISC was targeted at mothers and infants with severe disabilities. Besides, Landry et
al. focused on mother-infant interactions in ordinary play, while Klein et al. Emphasis was on
the context of teaching and learning. Difference in context explains why PALS contains more
affect or emotion-related sub-categories of behaviours, such as smiling, physical affection,
positive tone. In comparison, MISC contains more teaching and cognitive related sub-categories of behaviours, such as expanding, modelling and demonstration.

**Behavioural operational definition of contingent responsivity in present study**

Although there are some differences among their focuses, Yodder and Warren, Landry and Klein's study are complementary. Theoretically speaking, the contingent response should comprise behaviours that reflect the maternal capacity to be sensitive to and mentalise both the cognitive and emotional state of the infant. Therefore, the operational definition of contingent responsivity in the present study should attempt to bring together the taxonomies of behaviours in PALS and MISC with the concept of sensitivity and mentalisation in mind.
<table>
<thead>
<tr>
<th>Components</th>
<th>Sub-categories of behaviours</th>
<th>Behavioural operational definition</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scaffolding</td>
<td></td>
<td>After considering the child's cognitive and affective state concerning the learning episode, the mother's use of any verbal or nonverbal behaviours which scaffold, prepare and support the child to engage in a learning process.</td>
<td>- What is this? - Do you want this? - Is it a car</td>
</tr>
<tr>
<td>Question</td>
<td>A question used to guide the child towards engaging in a learning episode or a behaviour which is relevant to the learning process.</td>
<td>- What is this? - Do you want this? - Is it a car</td>
<td></td>
</tr>
<tr>
<td>Command</td>
<td>A command or a statement used to guide the child towards engaging in a learning episode a behaviour that's relevant to the learning process.</td>
<td>- Watch! - Look over here. - You try this</td>
<td></td>
</tr>
<tr>
<td>Components</td>
<td>Sub-categories of behaviours</td>
<td>Behavioural operational definition</td>
<td>Examples</td>
</tr>
<tr>
<td>---------------------</td>
<td>------------------------------</td>
<td>---------------------------------------------------------------------------------------------------</td>
<td>----------------------------</td>
</tr>
<tr>
<td>Excitement</td>
<td></td>
<td>Any expression that conveys excitement, appreciation or affect concerning the learning episode, concepts or any kinds of knowledge that would, in turn, excite the child's about the potential learning experience ahead.</td>
<td>- wow! Aww!</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Verbally, this could be conveyed through exaggerated utterance or speech content of one's opinions</td>
<td>- Ohh!</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Or</td>
<td>- This is interesting!</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nonverbally through exaggerated facial expression or bodily movement</td>
<td>- This is very beautiful!</td>
</tr>
<tr>
<td>Narration</td>
<td></td>
<td>A short statement that describes what is happening concerning the learning episode that demonstrates an intention to scaffold the child's cognitive state and guide the child towards the learning process.</td>
<td>- There are more object coming.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- I am going to show you how to switch on the light</td>
</tr>
<tr>
<td>Components</td>
<td>Sub-categories of behaviours</td>
<td>Behavioural operational definition</td>
<td>Examples</td>
</tr>
<tr>
<td>---------------------</td>
<td>------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>Labelling</td>
<td>Labelling of object, concept, action or function concerning the learning episode.</td>
<td>- This is a banana.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- We call this an apple.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- This is the switch</td>
<td></td>
</tr>
<tr>
<td>Referential</td>
<td>Any nonverbal behaviours that guide the child towards the learning episode.</td>
<td>- Pointing at an object</td>
<td></td>
</tr>
<tr>
<td>cues</td>
<td></td>
<td>- Holding an object</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Shaking an object</td>
<td></td>
</tr>
<tr>
<td>Regulating</td>
<td>Regulation of the child's behaviours concerning the learning episode through behaviours like planning, modelling, demonstration, verbal suggestion</td>
<td>- You can do this first, and then use that.</td>
<td></td>
</tr>
<tr>
<td>behaviours</td>
<td></td>
<td>- Let me show you how to do this.</td>
<td></td>
</tr>
<tr>
<td>Contingent</td>
<td>The mother responds to the child's signal in a manner that demonstrates the intention or attempt to mentalise the internal state of her child.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Components</td>
<td>Sub-categories of behaviours</td>
<td>Behavioural operational definition</td>
<td>Examples</td>
</tr>
<tr>
<td>------------</td>
<td>-----------------------------</td>
<td>-----------------------------------</td>
<td>----------</td>
</tr>
<tr>
<td></td>
<td>The key features to evaluate here is the degree of appropriateness as to the infant's internal state abilities, and the promptness of the maternal response.</td>
<td>Contingent responses</td>
<td>A mother responds verbally or nonverbally to the child's immediate signal of communication by</td>
</tr>
<tr>
<td></td>
<td>Complying to the presumed meaning</td>
<td></td>
<td>- The child pushes a toy car slightly forward, the mother said, &quot;you want to me to try?&quot; then the mother takes the car and pretends to play</td>
</tr>
<tr>
<td></td>
<td>Attempt to clarify the presumed meaning</td>
<td></td>
<td>- Child: &quot;uhhhhh&quot; Mother: &quot;oh! You want to play with the car?&quot;</td>
</tr>
<tr>
<td>Components</td>
<td>Sub-categories of behaviours</td>
<td>Behavioural operational definition</td>
<td>Examples</td>
</tr>
<tr>
<td>------------</td>
<td>-----------------------------</td>
<td>-----------------------------------</td>
<td>----------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Imitation</td>
<td>- The child is giggling, the mother follows; The child is playing with a toy car, the mother mimics the way the child plays</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Turn-taking in verbal or non-verbal form (Turn-taking is considered as a good demonstration of contingent response as it reflects a continuation of the attempt to mentalise the intended meaning for the child's communication)</td>
<td>- Child: Caaaa; Mother: Yes, this is a car; Child: geeeen; Mother: that's right, the car is green - The child gives the ball to the mother, she takes it, plays for a bit, returns to the child, the child further explores, repeat this process</td>
</tr>
<tr>
<td>Components</td>
<td>Sub-categories of behaviours</td>
<td>Behavioural operational definition</td>
<td>Examples</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-------------------------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Expanding</td>
<td></td>
<td>In responding to the child's interest or intention to learn, the mother directs toward broadening of the child's cognitive awareness concerning the learning episode through behaviours like describing, explaining, clarifying or comparing that goes beyond what is necessary to satisfy the immediate needs of the child.</td>
<td>- This is the switch! There is a white button here. We use it to turn to the light. We have a similar at home, do you remember?&quot;</td>
</tr>
<tr>
<td>Joint attention</td>
<td></td>
<td>In responding to the child's interest or intention to learn, the mother shifting her gaze between having eye contact with her child and the object, action or functions concerning the learning task leading to the child looking at the same target, which demonstrates the mother's continuous mentalisation and a moment of synchronisation of their thinking.</td>
<td></td>
</tr>
<tr>
<td>Components</td>
<td>Sub-categories of behaviours</td>
<td>Behavioural operational definition</td>
<td>Examples</td>
</tr>
<tr>
<td>--------------------</td>
<td>-------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>----------------</td>
</tr>
<tr>
<td>Emotional support</td>
<td></td>
<td>Emotional support encompasses verbal and nonverbal behaviours that demonstrate the sensitivity and mentalisation of the child's affective and emotional state that can facilitate the learning episode. This includes the presence of positive affective input, the absence of negative affect and unnecessary and excessive directive behaviours.</td>
<td></td>
</tr>
<tr>
<td>Warmth</td>
<td></td>
<td>Any behaviours that convey a gentle and warm approach to the child's behaviours or communicative signals</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Laughing and conveying a sense of enjoyment, appreciation, amusement or excitement.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Showing interest</td>
<td>- What are you playing?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Praising, encouragement and affirmation</td>
<td>- Well done!</td>
<td></td>
</tr>
<tr>
<td>Components</td>
<td>Sub-categories of behaviours</td>
<td>Behavioural operational definition</td>
<td>Examples</td>
</tr>
<tr>
<td>------------</td>
<td>-----------------------------</td>
<td>-----------------------------------</td>
<td>----------</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Empathy and support</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Any behaviours that demonstrate that the mother is mentalising and empathetic to the emotional state of the child.</td>
<td>- Good job!</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>You did it!</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Bravo! That's it, almost there!</td>
</tr>
<tr>
<td></td>
<td>Verbal</td>
<td></td>
<td>- You are very excited!</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- I know you are enjoying this.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Are you bored? Are you hungry?</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- You are upset, aren't you?</td>
</tr>
<tr>
<td></td>
<td>Nonverbal</td>
<td></td>
<td>- Hug</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Give the child food/drink/clothes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Gentle touch</td>
</tr>
<tr>
<td>Components</td>
<td>Sub-categories of behaviours</td>
<td>Behavioural operational definition</td>
<td>Examples</td>
</tr>
<tr>
<td>---------------------</td>
<td>-------------------------------</td>
<td>------------------------------------------------------------------------------------------------------</td>
<td>--------------------------------</td>
</tr>
<tr>
<td></td>
<td>Directive/ intrusive behaviours</td>
<td>Any behaviours that demonstrate the mother's intention or attempt to take the lead or control the child's action. These behaviours include but not limited to the follow:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Directive command</td>
<td>- Get that toy out of your mouth!</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Put the car down!</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Don't do this!</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Repeatedly addressing the child's name that conveys a sense of frustration</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Harsh or authoritarian voice tone</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Shouting</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Facial expression</td>
<td>- Frown, staring</td>
</tr>
<tr>
<td>Components</td>
<td>Sub-categories of behaviours</td>
<td>Behavioural operational definition</td>
<td>Examples</td>
</tr>
<tr>
<td>---------------------</td>
<td>------------------------------</td>
<td>------------------------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Restrictiveness</td>
<td>- Withholding object from the child or restricting child's movement</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Intrusion</td>
<td>- Forcefully taking objects from the child</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Forcing the object or action onto the child despite the child being upset or showing distress</td>
</tr>
</tbody>
</table>
Name addressing

In addition to the three ostensive cues proposed by Cisbra (2010), name addressing is another cue commonly found in mother-infant communication. Name addressing seems to meet the three criteria of ostensive cues. No other signal than a person's name can more unambiguously specify who the addressee of a communicative act is. By the age of 4.5 months, infants were found listening longer to their name compared to others (Mandel et al., 2016). Mandel-Emer and Jusczyk (2003) reported that infants show preference of listening to sentences containing their name over those containing a strangers' name. Infants can discriminate their name from other audio information from a very early stage. Neurocognitive studies have indicated that addressing infant by his/her name elicits a similar effect as other ostensive cues. Grossmann et al. (2010) reported that 5-months old infants showed the same neural activation in the prefrontal area of the brain when having eye contact and hearing their name. Similar findings were reported by several other studies (e.g. Enrici et al. (2011); Kampe et al. (2003); Walter et al. (2004)). In the present study, addressing infant by his/her name is therefore measured as one of the ostensive cues.

**Behavioural operational definition of name addressing in present study**

*Table 4.*

<table>
<thead>
<tr>
<th>Scale</th>
<th>Behavioural operational definition</th>
<th>Measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name addressing</td>
<td>The mother unambiguously addresses her infant by his/her name. The addressed name is regularly used in their daily interaction.</td>
<td>Frequency</td>
</tr>
</tbody>
</table>
Co-occurrence

Behavioural operational definition of Co-occurrence in present study

The co-occurrence scales concerns whether multiple ostensive cues are used within a short period. Therefore, the occurrence is operationally defined as in table 5

*Table 5.* Operational definition of co-occurrence

<table>
<thead>
<tr>
<th>Co-occurrence</th>
<th>Operational definition</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Co-occurrence</td>
<td>Number of discrete ostensive cues demonstrated in a given segment of a learning episode</td>
<td></td>
</tr>
<tr>
<td>No co-occurrence</td>
<td>0-1</td>
<td>0</td>
</tr>
<tr>
<td>Minimal co-occurrence</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Good co-occurrence</td>
<td>3-4</td>
<td>2</td>
</tr>
</tbody>
</table>
Step 3. Considerations of learning paradigms

To test whether the OCCS can be used to extract data of ostensive cues reliably from all pedagogical situation, in this initial stage, the OCCS will be applied to two learning tasks that are closely resemble to naturalistic situations.

Task A

Task A is a partial replication of Meltzoff (1988)'s imitation paradigm, this paradigm is deemed as a classical methodology in testing and demonstrating the effect of ostensive communication on infant's learning of cognitively opaque knowledge. In the original study, an experimenter demonstrated the novel and cognitively opaque action of touching a box with the forehead to switch on a lamp. After watching the demonstration, 70% of the infant patients learned the novel action and imitated. To make the paradigm resemble to natural pedagogical situation, in the present study, the novel head action will be demonstrated by the mother instead of the experimenter. The modified paradigm has been tested in a pilot study, it was evidenced that mothers tend to naturally use different ostensive cues in the demonstration of the novel action. Detail of other modifications will be illustrated in the method section.

Task B

Task B is a partial replication of the word learning study conducted Floor and Akhtar (2006). This paradigm is employed because the word-object association is one of the most common pedagogical situations that a mother and her infant encounter throughout the early development. The learning involves in the paradigm is also very different from task A, which potentially require different teaching strategies and tactics from the mother.

In the original design, Floor and Akhtar showed evidence that infant (16-20 month) can learn novel word-object association in both direct-address and overhearing condition. In
the direct-address condition, an experimenter presented eight novel objects (one at a time) and with an unusual name to infants with standard scripts "I am going to show you what's in here. Want to see what's in here? I will show you this one." And "I am going to show you the toma. Want to see the toma? I will show you the toma" (Floor and Akhtar (2006), p.332). The experimenter repeated the latter script for three times. Each object was labelled nine times in total. Before testing the infant's comprehension by asking "Can you show me the toma? Which one is the toma?", a preference control trial ("Show me the one you like?" or "which one do you want to play with?") was included to ensure the infant's choice was not due to salience or preference for the object.

Similarly, the paradigm is modified that the teaching procedures will be administered by the mother rather than the experimenter to make it resemble to natural interaction.
Step 4. Considerations of the initial examination of OCCS

The capacity to evaluate the performance of ostensive communication

The main objective of OCCS is to evaluate the performance of ostensive communication in any natural pedagogical situation. Theoretically speaking, OCCS should be sensitive to the level of performance and able to reflect it on the scores. Therefore, it is important to examine whether the OCCS can reliably distinguish individual who has high performance on ostensive communication from those who have low performance. To test OCCS’s capacity in this aspect, in addition to applying the OCCS on a group of participants who perform spontaneously in the learning paradigms, another group of participants, who are experimentally manipulated to have limited use of ostensive cues, should be added to the research design. Presumably, if the OCCS is effective in evaluating their performance, the OCCS scores of these two groups should show significant difference in the result of statistical analysis.

And of course, another essential aspect to be investigated is whether OCCS can be used by different researchers reliably. Therefore, the OCCS will be applied to the learning paradigms by at least two independent coders, the inter-ratter reliability of the instrument will be examined.

Mentalisation and reflective functioning

As discussed in previous sections, some empirical studies have suggested that the capacity to mentalise is the most important factor contributing to the ability to respond contingently, and potentially the use of other ostensive cues. Therefore, we would expect constructs of mentalisation and reflective functioning measured by validated in instruments to show some significant relationships with components that constitute contingent responsivity
(CR) (i.e. Scaffolding behaviours, contingent behaviours and emotional support), and potentially other ostensive cues measured by OCCS. Fonagy et al. (2016) suggested that there are two broad types of impairment in reflective function – hypomentalisation and hypermentalisation. Hypomentalisation is referring to concrete thinking, which reflects the inability to consider the complexity and lack of transparency of internal state. Hypermentalisation, also called pseudomentalizing or intrusive mentalising, is referring to the tendency to make attribution of mental representation of actions without appropriate evidence (Fonagy et al., 2016).

The validity of OCCS can be verified if mothers with impaired reflective function show a poorer performance on CR, or ostensive communication in general, in comparison with mothers who have good capacity to mentalise.

**Parenting stress level**

Various research suggested that parents who have high parenting stress demonstrate less positive feeling about their infant and are less able to respond contingently to their infant’s communication cues (Adams, 2006; Crnic et al., 1983; Middlebrook & Forehand, 1985). A number of studies have also suggested that parenting stress is negatively associated with children’s developmental outcome (Anthony et al., 2005; Campbell et al., 1991; Crnic et al., 2005; Crnic & Greenberg, 1990; Magill-Evans & Harrison, 2001; Moss et al., 1998; Pianta & Egeland, 1990) and language development (Magill-Evans & Harrison, 2001). Meanwhile, good quality of ostensive communication tends to positively associate with developmental outcomes (Landry et al., 2006; Ma et al., 2011; Niedźwiecka et al., 2018; Spinelli et al., 2017; Michael Tomasello & Kruger, 1992). Although the exact mechanism behind the relationships among these variables not clear, from the overt presentation, we would therefore expect mothers with higher level of parenting stress to show poorer
performance on ostensive communication in comparison with those mothers who have lower parenting stress.
Method

Ostensive communication (OC) project

The development of the ostensive communication coding system (OCCS) was part of a joint project with Shruti Jamnadass, trainee Clinical Psychologist. The primary research involved (see Appendix A for a detailed summary of researcher involvement). The study was designed to establish the psychometric property of the OCCS and test the effect of ostensive communication as a whole on infants’ learning in different kinds of knowledge.

Ethical approval

Ethical approval was obtained from the University College London (UCL) Ethics Committee (Reference 16673/001; Appendix B) and was issued UCL Data Protection Registration (Z6364106/2019/10/66).

Information sheet (Appendix C) was sent to participants at least 24 hours before the study. All participants understood that there would not be any possible risks involved in taking part in the study and they provided their informed consent (Appendix D). After the study, all participants were given a debrief (Appendix E) and they had opportunities to ask questions.

Participants

Power analysis

Bujang and Baharum (2017) used Power Analysis and Sample Size (PASS) software (version 11.0.7; PASS, NCSS, LLC) to calculate the minimum sample size to estimate the value of intraclass correlation coefficient (ICC). They produced several tables (Bujang and Baharum, 2017, p.7-p.9) to illustrate the minimum number of participants required for
estimate the desired effect size. According to the result of their analysis, with a criterion of 0.70 for reliability and 0.40 for validating, 36 participants are needed for the 80% power criterion and 50 participants are needed for 90% power criterion.

**Recruitment**

Mother with infants aged 14-22 months were recruited through volunteering and opportunity sampling procedures. Prior researches indicated that infants in this age have acquired the cognitive capacity to form mental representation of novel object with novel name (Floor & Akhtar, 2006, Senju & Csibra, 2000), as well as being a rational agent capable of making simple effectiveness evaluation (Buttelmann et al., 2008; Király, 2009; Tomasello, 2006). They have also developed sufficient motor capacity and able to imitate simple actions, such as use of novel apparatuses (Matheson, Moore & Akhtar 2013; Williamson & Brand, 2014). Participants received £17 to compensate for their time and travel cost, and an “Anna Freud Centre Bib” as an incentive for the toddler. The Bib was designed by the researchers and has been very helpful to get the mother’s interested in taking part in the study. A poster (Appendix F) was made which advertised the study. The poster was distributed in the local library, baby bounce, and under 5s class, playgroups and social media (e.g. Facebook & Instagram). There was an excellent response rate with the baby bounce class and social media. Over 65 mother-toddler dyads were successfully recruited to take part in the study. Mothers were contacted via phone or email to organise a date to visit the AFC for the study and an appointment letter was sent to confirm the study date (Appendix G).

**Study design and learning tasks**

The study is a 2x2 mixed experimental design. Participants were randomly allocated into two groups: (1) semi-naturalistic ostensive communication (SNOC), and (2) restricted
ostensive communication (ROC). There were two learning paradigms: (A) imitation paradigm (modified from Meltzoff (1998), and (B) word-object learning (modified from Floor and Akhtar (2006)).

In the SNOC group, the mother-infant dyads were asked to accomplish the learning tasks with minimal instruction from the experimenter so that the mother and infant could interact in a pedagogical situation that most closely resembles their interaction in real life.

In the ROC group, the ostensive communication was manipulated. Mothers were given clear instructions to limit their use of ostensive cues. Decision of what ostensive cues to restrict was based on the feedback from two pilot experiments. Two participants (P25 & P55) who had already completed the experiment in the SNOC group were invited to take part in the pilot experiments. According to their feedback, they found it very difficult not to have any eye contact with their infant because it felt most unnatural, and it could cause potential distress to the infant. The wide variety of contingent responses made it too challenging to prohibit all these behaviors. But the participants reported that they found it manageable not to address their infant by their names and not to use IDS.

**Learning Task A: the imitation paradigm (switching on the lamp by using the forehead)**

Task A was a partial replication of Meltzoff (1988)'s imitation paradigm. A modification was implemented to create a pedagogical situation that's most closely resemble real-life interaction. The novel head action is, therefore demonstrated by the mother instead of the experimenter.

Moreover, ideally, the present study should replicate the timing of the test phase, which takes place a week after the demonstration, to test whether the infant could retain newly learned knowledge. A modification had to be made to manage the logistical demand of two visits to the laboratory per participant. The testing phase is therefore taking place after...
the completion of task B. Although it is a shorter period, task B is considered a massive de-
mand to infants' cognitive resources, which made the retention of the newly learned
knowledge equally challenging compared to the original design.

**Learning task B: word-object association**

Task B is a partial replication of the word learning study conducted by Floor and Akhtar (2006). This paradigm is employed because the word-object association is one of the
most common pedagogical situations which a mother and her infant encounter throughout the
early development. The learning involved in the paradigm is also very different from task A,
which potentially require different teaching strategies and tactics on behalf of the mother.
Figure 1.
Study design

Note:
SNOC: Semi-naturalistic ostensive communication
ROC: Restricted ostensive communication

Procedure

Setting

The experiment was conducted in a consultation room generally used for psychotherapy session for parents and children. The room was around 120 sq ft. There was one normal armchair for adults, one smaller chair for children and a table with a lamp attached on it. The lamp could only be turned on and off by an experimenter by using a remote switch. There was a blanket with some interesting cartoon prints on the floor, and a few cushions around for participants to use when needed. The room was designed to give a comfortable and safe feeling; parents and children often relaxed and settled soon after they entered the room. There were four cameras attached to each side of the wall. The experimenters controlled these cameras through a computer placed on a worktop in one of the corners. The worktop was at a higher level where children were not able to see the screen. These cameras recorded videos simultaneously. While recording, the computer screen showed a live preview of cameras, and the experimenter could manually adjust each camera accordingly to make sure the camera was shooting at the right direction to capture the full interaction between the mother and the
children at the best angel. The materials needed for the experiment were also placed on the worktop.

First, an experimenter welcomed the mother and the toddler from the reception and brought them to the waiting area. The information sheet, consent form and self-report questionnaires were sent with a confirmation letter to the mothers beforehand. To avoid any "ostensive cue pollution" and any social relationship established between the experimenter and the toddler, we notified the mother in the confirmation letter that, for this study, the experimenter would maintain minimal interaction with the toddler, and therefore come across as rather cold. The experiencer greeted only the mother and avoided any eye contact or interaction with the toddler throughout the experiment unless necessary. The mother and toddler were led to the waiting area on the same level as the experiment room. It was a comfortable environment that there were some exciting toys and books around for the toddler to settle in quickly.

1. Briefing phase

After the toddler was settled, usually with some toys, the experimenter went through the information sheet and consent form with the mother making sure she understood the terms, conditions and her rights in the study and also to give time to answer any questions or concerns from the mother. The experimenter gave the mother an initial briefing about the flow and the procedure of the experiment. The mother then watched three instruction videos by herself (not with the toddler), in which she was told what she was expected to do in each phase. After the experimenter checked with the mother that she understood everything, to ease the performance anxiety, they were also reminded that they did not need to remember every step of the task as the instruction video would be shown again right before the task. Plus, there was instruction sheet lying around for her to refer to at any time. The mother was
also explicitly reassured that we did not expect her child to be able to accomplish all the task as we have set the task to be slightly challenging for a toddler at this age.

2. **Free play**

When the mother and toddler were ready, they were brought into the experiment room. They were then given three minutes to do whatever they wanted and to feel familiar and relaxed in the room.

3. **Playing with toys**

The mother and toddler were given a set of familiar toys to play with for another three minutes. It was a multi-coloured, shape sorter bus. With the assistance of the mother, the toddler could sort different shapes and push them through the holes that matched the different shapes of the blocks. It naturally evoked a pedagogy interchange between the mother toddler dyad. This set of toys were not removed before Task A to avoid unnecessary interaction with the toddler unless there was evidence that the toddler was obsessed with it and unable to engage with the following tasks at all.

4. **Task A – Imitation paradigm – Demonstration phase**

In the instruction video, the mother was told that she would be given a box. The task was to teach her child to turn on the lamp by touching their head onto the box. The experimenter demonstrated the novel action in the video. The experimenter placed an instruction sheet (Appendix H) for the mother to reference from. The mother was allowed to use whatever strategies she liked, except physically making the toddler to do the act. The mother could demonstrate up to three times, and the toddler was not allowed to touch the box or take any trail before the testing phase. When the mother touched her head onto the box, the
experimenter turned on the lamp by using the remote switch to create an impression of a causal relationship between the two.

5. Task B – Word-object association - Familiar objects

In task B, familiar objects phase, after watching the instruction video, the mother was first given four objects that toddlers often come across in their day to day life. These included a yellow plastic ball, a purple toy cup, a dog figure and a yellow plastic banana.

In the instruction video, the mother was told this task aimed at teaching their toddler the name of these objects. This phase was designed to allow the mother and the toddler to get familiar with the mechanism of this learning task. Therefore, the difficulty of this phase was reduced by using objects that toddlers find familiar.

6. Preparation

Before the teaching started, the mother was instructed to let the toddler play with those objects for around 30 seconds before and then to ask "Which is your favourite object?"

In the piloting testing, some toddlers showed excessive interest in one or two of the objects and became unable to engage with the task at all. This step was to identify this potential object and remove it. Once the toddler selected their favourite object, the mother gave it to the experimenter and started the teaching.

7. Teaching and immediate testing

The mother was allowed to use whatever strategies she liked, but she could only name the object up to five times, and she was to teach one object at a time.

Once the mother believed her toddler learned the name of the first object, she mixed it with the other two and tested whether the toddler learned it successfully by asking "Can you show
me the (name)". The mother was not allowed to give any prompt or feedback to the toddler if they gave a wrong answer. She could only ask the same question twice. She had to move onto the next object no matter what response the toddler gave. She repeated this process for the other two objects.

8. Delay recall testing

At the end of the familiar phase, she tested if the toddler was able to retain the learning by asking the toddler once again "Can you show me the (name)" for each object.


The procedures of the novel object phase were the same as with the familiar object phase except this time the mother was given two rounds of four unusual novel objects and made-up names (8 objects all together). The toddler was less likely to have seen them in daily life. These objects included a black digital recorder (Saka), a blue badge holder (Toma), a classic wooden thin pick game set (fully sealed) (Yossi), a black cycling computer (Sifa), a pink hair roller (Modi), a wooden puzzle (Loport), a blue plastic triangle (Biku) and a grey scarf hanger (Viko). These objects were presented with various colours and functions that gave a range of features.

10. Task A – Imitation paradigm – testing phase

In this phase, the mother was given the white box so she could place it in front of the toddler. She could say "Can you turn on the lamp?". But she was not allowed to perform any further demonstration, to give the toddler any verbal or non-verbal prompts, or correct the toddler if they did not imitate.
### Table 6.
*Summary of the experiment procedure*

<table>
<thead>
<tr>
<th>Phase</th>
<th>Aims</th>
<th>Instructions to mother</th>
<th>Materials</th>
</tr>
</thead>
<tbody>
<tr>
<td>Briefing</td>
<td>To obtain consent and introduction</td>
<td>1. To watch instruction videos</td>
<td>1. Information sheet</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Consent form</td>
<td>2. N/A</td>
</tr>
<tr>
<td>Free play</td>
<td>To feel relaxed and settled</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Playing with Toys</td>
<td>To feel relaxed and settled</td>
<td>N/A</td>
<td>1. multi-coloured, shape sort ter bus</td>
</tr>
<tr>
<td>Task A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teaching</td>
<td>1. To teach the toddler</td>
<td>1. To watch instruction video</td>
<td>1. Lamp</td>
</tr>
<tr>
<td></td>
<td>to switch on the lamp</td>
<td>2. To use whatever strategies she liked</td>
<td>2. White box</td>
</tr>
<tr>
<td></td>
<td>by touch their head</td>
<td>3. To demonstrate the novel action up to 3 times</td>
<td>3. Instruction sheet</td>
</tr>
<tr>
<td></td>
<td>onto the box</td>
<td>4. To not physically make the toddler performing the action</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5. To not let the toddler try to or touch the box</td>
<td></td>
</tr>
<tr>
<td>Phase</td>
<td>Aims</td>
<td>Instructions to mother</td>
<td>Materials</td>
</tr>
<tr>
<td>--------------------------</td>
<td>-------------------------------------------</td>
<td>----------------------------------------------------------------------------------------</td>
<td>------------------------------------------------</td>
</tr>
<tr>
<td>Task B</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Familiar object –</td>
<td>1. Identify the toddler's favourite object and remove it.</td>
<td>1. To watch introduction video</td>
<td>1. A yellow plastic ball</td>
</tr>
<tr>
<td>Preparation</td>
<td></td>
<td>2. To let the toddler play with four objects for 30s</td>
<td>2. A purple toy cup</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. To ask toddler &quot;Which is your favourite.&quot;</td>
<td>3. A dog figure</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. To remove the object</td>
<td>4. A yellow plastic banana</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5. Instruction sheet</td>
</tr>
<tr>
<td>Familiar objects -</td>
<td>1. To teach the toddler the name of the remaining three everyday objects</td>
<td>2. To use whatever strategies she liked</td>
<td></td>
</tr>
<tr>
<td>Teaching and immediate</td>
<td></td>
<td>3. To teach one object at a time</td>
<td></td>
</tr>
<tr>
<td>testing</td>
<td></td>
<td>4. To name the object up to 5 times</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5. To test the toddler's learning by saying &quot;Can you show me (name)&quot; up to 2 times for each object</td>
<td></td>
</tr>
<tr>
<td>Family objects -</td>
<td>1. To test if the toddler can delay recall of can delay recall of</td>
<td>1. To ask the toddler once again &quot;Can you show me (name)?&quot;</td>
<td></td>
</tr>
<tr>
<td>Delay recall testing</td>
<td></td>
<td>2. To let the toddler play with four objects for 30s</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. To ask toddler &quot;Which is your favourite.&quot;</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. To remove the object</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phase</td>
<td>Aims</td>
<td>Instructions to mother</td>
<td>Materials</td>
</tr>
<tr>
<td>---------------</td>
<td>-----------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Novel objects</td>
<td>1. To identify the toddler's favourite novel and remove it.</td>
<td>1. To watch introduction video</td>
<td>1. A black digital recorder (Saka)</td>
</tr>
<tr>
<td>Group 1 -preparation</td>
<td></td>
<td>2. To let the toddler play with four objects for 30s</td>
<td>2. A blue badge holder (Toma)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. To ask toddler &quot;Which is your favourite.&quot;</td>
<td>3. A classic wooden thin pick game set (fully sealed) (Yossi)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. To remove the object</td>
<td>4. A black cycling computer (Sifa)</td>
</tr>
<tr>
<td>Novel objects</td>
<td>1. To teach the toddler the name of the remaining three novel objects</td>
<td>1. To use whatever strategies she liked</td>
<td>5. Instruction sheet</td>
</tr>
<tr>
<td>Group 1 -Teaching and immediate testing</td>
<td></td>
<td>2. To teach one object at a time</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. To name the object up to 5 times</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. To test the toddler's learning by saying &quot;Can you show me (name)?&quot; up to 2 times for each object</td>
<td></td>
</tr>
<tr>
<td>Novel objects</td>
<td>1. To test if the toddler can delay recall of</td>
<td>1. To ask the toddler once again &quot;Can you show (name)?&quot; and go through each object</td>
<td></td>
</tr>
<tr>
<td>Phase</td>
<td>Aims</td>
<td>Instructions to mother</td>
<td>Materials</td>
</tr>
<tr>
<td>------------------------</td>
<td>-----------------------------------------------</td>
<td>---------------------------------------------------</td>
<td>------------------------------------------------</td>
</tr>
<tr>
<td>Group 1 - Delay recall testing</td>
<td>the name of the novel objects</td>
<td></td>
<td>1. A pink hair roller (Modi)</td>
</tr>
<tr>
<td>Novel objects</td>
<td>1. To identify the toddler's favourite novel object and remove it.</td>
<td>1. To watch introduction video</td>
<td>2. A wooden (Loport)</td>
</tr>
<tr>
<td>Group 2 – Preparation</td>
<td>2. To let the toddler to play with four objects for 30 seconds</td>
<td>2. To let the toddler to play with four objects for 30 seconds</td>
<td>3. A blue plastic triangle (Biku)</td>
</tr>
<tr>
<td></td>
<td>3. To ask toddler &quot;Which is your favourite?&quot; Remove the object</td>
<td>3. To ask toddler &quot;Which is your favourite?&quot; Remove the object</td>
<td>4. A grey scarf hanger (Viko)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5. Instruction sheet</td>
</tr>
<tr>
<td>Novel object</td>
<td>1. To teach the toddler the name of the three novel objects</td>
<td>1. To use whatever strategies she liked</td>
<td></td>
</tr>
<tr>
<td>Group 2 - teaching</td>
<td>2. To teach one object at a time</td>
<td>2. To teach one object at a time</td>
<td></td>
</tr>
<tr>
<td>and immediate testing</td>
<td>3. To name the object up to 5 times</td>
<td>3. To name the object up to 5 times</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4. To test the toddler's learning by saying &quot;Can you show me (name)&quot; up to 2 times</td>
<td>4. To test the toddler's learning by saying &quot;Can you show me (name)&quot; up to 2 times</td>
<td></td>
</tr>
<tr>
<td>Novel objects</td>
<td>1. To test if the toddler can delay recalling and go through each object</td>
<td>1. To ask the toddler once again &quot;Can you show me (name)?&quot;</td>
<td></td>
</tr>
<tr>
<td>Group 2 - Delay recall testing</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phase</td>
<td>Aims</td>
<td>Instructions to mother</td>
<td>Materials</td>
</tr>
<tr>
<td>-------------------</td>
<td>-------------------------------</td>
<td>-------------------------------------------------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td></td>
<td>the name of the novel</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>objects</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Task A**

Delay recall testing

1. To test if the toddler can imitate the novel head action

1. To ask the infant "Can you turn on the lamp?"

2. To not allowed to give a further demonstration, verbal or non-verbal prompt

1. Lamp

2. White box

3. Instruction sheet
Measures

Ostensive communication coding system (OCCS)

The following section outlines the development of the OCCS scale in terms how each ostensive cue and other relevant variable are measured and computed in the instrument.

Demographics Questionnaire (Appendix I)

Characteristics such as age, ethnicity, years of education and household income were collected.
**Figure 2.**
*Overview of ostensive communication coding system*

<table>
<thead>
<tr>
<th>Time</th>
<th>Level of Attention</th>
<th>IDS (Average of four components)</th>
<th>Eye contact (FC)</th>
<th>Contingent Responsivity</th>
<th>Name addressing (FC)</th>
<th>Co-occurrence</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Second)</td>
<td>(LC 0-2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Proodic Contours (LS 0-2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rhythm (LS 0-2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Linguistic structure (LS 0-2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Repetition (LS 0-2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Scaffolding (FC)**
- Question (FC)
- Command (FC)
- Exciment (FC)
- Narration (FC)
- Labelling (FC)
- Referntial Cues (FC)
- Regulating behaviours (FC)

**Contingency Response** (Average of three components)
- Contingent behaviours (LS 0-2)
- Expanding (LS 0-2)
- Joint attention (LS 0-2)

**Emotional Support** (Average of two components)
- Warmth & supportiveness (LS 0-4)
- Directive behaviour to intrusion (LS 04)

**Note:**
FC: Frequency count; LS 0-2: Likert scale rating 0-2; LS 0-4: Likert scale rating 0-4
Time

Time spent on each teaching episode is operationally defined in table 7. This is to make sure the coders are coding ostensive cues used within the same time frame. Any behaviours which happen outside this frame were not coded by the OCCS.

Table 7.
Definition of the start and the end of a teaching episode in Task A and Task B

<table>
<thead>
<tr>
<th>Task</th>
<th>Episode</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task A</td>
<td>Transition</td>
<td>Start: The mother finishes watching the video instruction for the task</td>
</tr>
<tr>
<td></td>
<td></td>
<td>End: The switch (box) is placed onto the table and ready for the mother to start the teaching</td>
</tr>
<tr>
<td></td>
<td>Demonstration</td>
<td>Start: The mother initiates the demonstration</td>
</tr>
<tr>
<td></td>
<td></td>
<td>End: Two seconds after the mother finishes the demonstration</td>
</tr>
<tr>
<td>Task B</td>
<td>Transition 1</td>
<td>Start: The mother finishes watching the instruction video</td>
</tr>
<tr>
<td></td>
<td></td>
<td>End: The mother starts asking the toddler about his/her favourite object</td>
</tr>
<tr>
<td></td>
<td>Familiar object</td>
<td>Start: The mother starts teaching</td>
</tr>
<tr>
<td></td>
<td>teaching</td>
<td>End: The mother's first trial on asking &quot;Show me the (name of the object)&quot;</td>
</tr>
<tr>
<td></td>
<td>Transition 2</td>
<td>Start: Removal of the object from the last round</td>
</tr>
<tr>
<td></td>
<td>Start</td>
<td>End</td>
</tr>
<tr>
<td>----------------------</td>
<td>-------</td>
<td>-----</td>
</tr>
<tr>
<td>Familiar object 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>teaching</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transition 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Testing object 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>teaching</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Repeat till Testing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>object 6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Attention

Attentional level is not part of the OCRS, but it is captured as part of the study. The toddler’s attentional level is measured in Likert-scale between 0 to 2. Score 0 indicates that the toddler is mostly uninterested and unable to attend or concentrate on the learning task. Score 2 suggest the toddler shows good interest in the object or tasks, and they can maintain a good level of attention throughout the process. The measures of attention allocation were adapted from procedures described in the Laboratory Temperament Assessment Battery (Coan et al., 2005), a lab-based measure of multiple dimensions of children’s temperament. Ratter needs need to take notice on the intensity of interest displayed as facial expression, duration of look given the relevant object or task and the length of an object or task manipulation.

IDS

After reviewing the pilot coding that was conducted on two videos from the pilot study, there was evidenced that these four IDS components are not always present simultaneously.

For example, a mother spoke in IDS with distinctive higher and broader range of pitch, exaggerated vowels, and used mostly simple phrases and onomatopoeia throughout the interaction, she also occasionally used full and partial reputation. Her IDS was consistent primarily with the IDS features; however, she spoke in a very fast pace and rhythm. In contrast, another mother spoke in a noticeable slower pace with longer and more frequent pauses in between isolated words and phrases throughout the task. But the IDS also came with relatively flat tone with slight variation, and there was almost no repetition of any form identified. Seemingly, these four components of IDS do not always come as a package. A demonstration of one component does not guarantee the other. Investigation of the reason and
mechanism behind this phenomenon might be beyond the scope of this study. Still, this observation indicated that these four components should be examined and coded independently to reflect the individual differences in IDS.

Besides, the pilot coding also showed that, in a given teaching episode, there was considerable variation in the prevalence of these components. For example, some mothers only demonstrated repetition once throughout the process, while another mother showed multiple. Or some mothers occasionally spoke with simplified linguistic structure while other mothers demonstrated this quality more frequently. The coding system needs to be able to capture this subtle difference. Therefore, each competent was rated in a Likert-scale between 0 to 2. Score 0 means no indication of the IDS quality. Score 1 indicates an occasional demonstration of the IDS quality. Score 2 indicates a high prevalence of the IDS quality in teaching episode.

The overall IDS performance is based on the integrated evaluation of each IDS component. Likert ratings of each component are added up to generate an overall IDS score


Contingent responsivity

1. Sub-scale of contingent behaviours

In the development of the subscale of contingent behaviours, the sub-categories (i.e. Contingent responses, joint attention and expanding) were initially measured in frequency. But pilot coding revealed that to merely measure the frequency does not capture the essence of these behaviours. Individuals can show distinctive differences in the quality of their behaviour even though they use the same number of behaviours. In regard to contingent response, for example, in the pilot coding a mother was observed to having picked up the communicative cue from the infant, but she responded in a delayed manner. Her responses were also short-lived without continuity. Another mother demonstrated a higher quality of contingent responses by having responded in a timely manner and furthermore by having back and forth exchanges with the infant. Similar, a lower quality of joint attention can be present when for a short moment mother and infant look at the object simultaneously. A higher quality of joint attention would demonstrate a slightly longer period of exchange the gaze as well as shifting in between the target and eyes, which is a clear indication that both parties are processing the information. In a higher quality Expanding, the mother broadens the cognitive awareness of the infant with multiple dimensions (e.g. colour, shape, function) by using verbal and non-verbal communication. But a lower quality expanding might only have touched on one of the aspects.

Therefore, the subscale of contingent behaviours is measured in Likert-scale 0-2. A score 0 indicate no such behaviour is identified; a score 1 means a low quality and a score 2 means high quality of contingent behaviours.
2. Subscale of emotional support

The overall score of the emotional support is computed by the average of the score of the warmth and supportiveness subscale and the reverse score of the directive and intrusive behaviours scale. These two components are rated in a Likert scale between 0 to 4. In the development of the subscale of directive and intrusive behaviours, the outcome of the pilot study suggested that a slight and appropriate degree of directive behaviours might actually be helpful to regulate the infant’s attention and interest concerning the learning episode.

Therefore, an adjustment of the reverse score is made as the following:

<table>
<thead>
<tr>
<th>Sub-scale</th>
<th>Score</th>
<th>Reversed score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Directive and Intrusive behaviours</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0</td>
</tr>
</tbody>
</table>

3. Overall score of contingent responsivity

There is no a single overall score is computed for contingent responsivity as each of its component is evaluated in different measurement.
Co-occurrence

To achieve the objective of measuring co-occurrence, the use of ostensive cues needs to be recorded at a micro-level. Therefore, each teaching episode (see the teaching of familiar object 1 (F1) below as an example) needed to be divided into several segments.

A segment is defined as the 8-seconds-period from the mother’s use the first identifiable ostensive cue. Studies have shown that within the first year of life, infants can effectively extract information from their mother's responses when these responses come within a 2-5 seconds window (Goldstein et al., 2009; Goldstein et al., 2010; Yu & Smith, 2012). With the advancement in the cognitive capacity of the infant (16-22 months old) participants in the present study, the response window should be increased. A range of duration was tested in pilot coding; an 8-seconds-window was found to be most appropriate.

Discrete ostensive cues used within the segment will be measured in frequency count. Eye contact and name addressing are straightforward; any appearance of these cues will be considered as an example of the use a discrete ostensive cue. It is more sophisticated for IDS and contingent response because there are multiple components or categories of behaviours that encompass these cues.

For a mother to be validated as using IDS in the segment, she has to demonstrate at least two unique qualities of IDS (i.e. Prosodic contours, Rhythm, Linguistic structure and repetition). The validation criteria is set to two qualities, because the mother might demonstrate one of these qualities only by chance. Therefore, it is insufficient for the speech to be distinguished from ADS.

For a similar reason, the validation criteria of contingent response is that a mother has to demonstrate a performance of at least two discrete categories of behaviours (i.e. scaffolding, contingency and emotional support). See segment three below as an example. If
the mother only uses one contingent response (e.g. taking an object from the child), it does
not give sufficient evidence to conclude that is not happening by chance. Hence, she is not
considered as using a contingent response in segment 3. In segment 4, the mother uses three
behaviours, which fall into two discrete categories (contingency and emotional support), we
can be confident to conclude she has used contingent response as a distinct ostensive cue.

The co-occurrence score of each teaching episode will be generated by adding up the
co-
Table 9.
Computation of co-occurrence score

<table>
<thead>
<tr>
<th>Teaching Episode</th>
<th>F1</th>
<th>Segments of F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>10:00-10:30</td>
<td>S1 (10:00-10:08)</td>
</tr>
</tbody>
</table>

Ostensive cues use in the segment

<table>
<thead>
<tr>
<th>Segment</th>
<th>Ostensive Cues</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>1. EC (10:00)</td>
</tr>
<tr>
<td></td>
<td>2. NA (10:01)</td>
</tr>
<tr>
<td></td>
<td>3. CR (pointing) (10:00-10:05)</td>
</tr>
<tr>
<td>S2</td>
<td>1. IDS (10:09-10:14)</td>
</tr>
<tr>
<td></td>
<td>2. CR (pointing) (10:10)</td>
</tr>
<tr>
<td></td>
<td>3. CR (Expanding) (10:12)</td>
</tr>
<tr>
<td>S3</td>
<td>1. NA (10:16)</td>
</tr>
<tr>
<td></td>
<td>2. CR (Taking the object from the child) (10:16)</td>
</tr>
<tr>
<td>S4</td>
<td>1. EC (10:26)</td>
</tr>
<tr>
<td></td>
<td>2. NA (10:26)</td>
</tr>
<tr>
<td></td>
<td>3. IDS (10:26-10:30)</td>
</tr>
<tr>
<td></td>
<td>4. CR (labelling) (10:27)</td>
</tr>
<tr>
<td></td>
<td>5. CR (turn-taking) (10:26-10:30)</td>
</tr>
<tr>
<td></td>
<td>6. CR (praising) (10:30)</td>
</tr>
</tbody>
</table>

Frequency count of ostensive cue for each segment

<table>
<thead>
<tr>
<th>Segment</th>
<th>2 discrete ostensive cues</th>
<th>2 discrete ostensive cues</th>
<th>1 discrete ostensive cue</th>
<th>4 discrete ostensive cues</th>
</tr>
</thead>
</table>

Co-occurrence score of each segment

<table>
<thead>
<tr>
<th>Segment</th>
<th>Co-occurrence Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>1 (minimal co-occurrence)</td>
</tr>
<tr>
<td>S2</td>
<td>1 (minimal co-occurrence)</td>
</tr>
<tr>
<td>S3</td>
<td>0 (no co-occurrence)</td>
</tr>
<tr>
<td>S4</td>
<td>2 (Good co-occurrence)</td>
</tr>
</tbody>
</table>

Sum of Co-occurrence score of F1

<table>
<thead>
<tr>
<th>Sum of Co-occurrence Score of F1</th>
<th>1+1+0+2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>=4</td>
</tr>
</tbody>
</table>

Note: EC: Eye contact; NA: Name addressing; CR: contingent response
The rate of using ostensive cue

Conventionally, eye contact, name addressing, and scaffolding behaviours are mostly measured in frequency. In reviewing the pilot study, it was evidenced that two of the participants used the same number of eye contact, however, they have spent different amount of time in the teaching episode. Instinctively, same number of ostensive cue might not always have the same influence on the infant’s learning if these cues are distributed in a different span of time, therefore, only measuring these ostensive cues in frequency might bring a significant limitation in the evaluation of ostensive communication.

In light of this observation, the design of OCCS aims at addressing this limitation by also capturing the performance of ostensive communicate in terms of the rate of using ostensive cue. For example, the rate of eye contact is defined as following:

\[
\text{No. of eye contact used in a given teaching episode} \quad \frac{\text{Time spent on the teaching episode}}{\text{Time spent on the teaching episode}}
\]

The rate of name addressing and scaffolding behaviours will be computed the same way. It is proposed that the rate of ostensive cue is an independent dimension of ostensive communication from frequency.
Measures used for validity analyse

The following measures were used to assess construct validity of OCCS.

1. **The Parental Reflective Functioning Questionnaire (PRFQ)** (Fonagy et al., 2016; Luyten et al., 2017)

   The PRFQ is a self-report 18-item assessment of parental reflective functioning that was primarily developed for parents with children aged 0-5 years. It focuses on a parent’s ability to be sensitive and responsive to their child’s needs (Luyten, Mayes, Nijssens, & Fonagy, 2017). The instrument consists of three sub-scales, each scale has six items that can be rated with a Likert-scale between 1 (Strongly disagree) to 7 (Strongly agree).

   The first “Pre-mentalizing (PM)” subscale is designed to capture parent’s non-mentalizing stance. It contains items such as “My child sometimes gets sick to keep me from doing what I want to do” and “Often, my child’s behaviour is too confusing to bother figuring out.” High score on PM suggests the parent tend to make malevolent attribution and an inability to tune in to the subjective world of the child (Luyten et al., 2017). The higher the score on PM, the poorer the capacity to mentalise.

   The second “Certainty of mental state (CM)” subscale is designed to assess parent’s ability to recognise the opaqueness and complexity of mental state. It consists of items such as “I always know why my child acts the way he or she does” and “I can completely read my child’s mind”. Extremely low score on CM suggests the parent’s almost complete lack of certainty about the child’s internal state, which is a feature of hypomentalising. On the contrary, extremely high score on CM reflects overly certain about the mental state, which is a feature of hypermentalizing.

   The third “Interest and curiosity in mental state (IC)” subscale captures the parent’s curiosity in the child’s internal state. It consists of items such as “I like to think about the reasons behind the way my child behaves and feels” and “I try to see situations through the
eyes of my child”. Low score on IC reflects lack of interest in the child’s internal state, yet extremely high score on IC might reflect intrusive hypermentalisation.

The normative data of PRFQ is drawn from the study conducted by (Pazzagli et al.) in 2017. In their study, parental reflective functioning questionnaire was administered to a community sample of Italian mothers and fathers of 385 children age 3-10 years (M = 6.72, SD = 2.08). The average age of mothers was 36.62 years old (SD = 5.77). Although the infant participants in the present study is much younger, the mothers in both studies are from a comparable demographic background in terms of age, ethnicity and socioeconomic status.

2. Parenting Stress Index, Fourth Edition Short Form (PSI-4-SF)

Parenting Stress Index-Short Form (PSI Abidin, 1995) is a 36-item self-report measure designed to capture the level of parenting stress of parents who have children of 3 months to 10 years of age. It consists of three subscales; Parental Distress (PD) reflects parent’s perception of child-rearing competence, conflict between spouse or partner, level of social support and stresses associate with restrictions as a parent. Difficult Child (DC) assess the parent’s view on the child’s temperament, compliance and demandingness. Parent-Child Dysfunction Interactions (P-CDI) captures the parent’s perception on whether the child meets their expectation. Each item is rated on a 5-point scale from “Strongly Disagree” = 1 to “Strongly Agree” = 5. The PSI-SF yield a total stress score and three subscale scores. Haskett et al. (2006) investigated the psychometric properties of the PSI-SF. The measure was administered to 95 mothers and fathers (Age $M = 34.40$, $SD = 8$) with children age 4-10 ($M = 7.0$). The normative data of the total stress score ($M = 79.00$, $SD = 18.30$) reported in the study was derived from parents from the non-clinical population who acted as the control group. These parents have no known history of abuse and from a comparable demographic background with the sample in the present study.
Step 5. Generating data of ostensive communication via OCCS

Coding of ostensive communication

Two coders were 3rd-year candidates of doctoral in clinical psychology. Coders were trained through pilot coding. After two coders coded independently on three videotapes of Task B, they compared and reviewed their coding on each item. They provided feedback on item content, operational definitions and examples. Feedback was used to refine the manual and the coding system. The pilot phased ended after the coders reached a less than 20% score difference on all the item.

Subsequently, the two coders will independently apply the OCCS to code the ostensive cued used in both learning paradigms.
Step 6. Establish the psychometric property of OCCS

In order to establish the psychometric property of OCCS, the present study aimed to test the following hypotheses with the proposed inferential analysis:

1. **Hypothesis 1: OCCS achieves high to excellent level of inter-rater reliability in all sub-scale.**

   The first step to establish the psychometric property is to examine whether OCCS is a reliable instrument across raters. The inter-rater reliability of each item in OCCS will be assessed using interclass correlation coefficients (ICCs; Shrout & Fleiss, 1979) base on coding of the videotape on Task B.

2. **Hypothesis 2: Co-occurrence has positive relationship with all sub-scale frequency, quality and rate of different ostensive cues**

   It is hypothesized that ostensive cues tend to co-occur within a short period of time, and the OCCS is designed to capture this aspect of ostensive communication. Intuitively, the higher frequency and rate, better the quality of ostensive cues used in the interaction, the more likely that there is co-occurrence of ostensive cues. Therefore, we will expect a significant positive association between co-occurrence and frequency, quality and rate of ostensive cues in Pearson correlation analysis.

3. **Hypothesis 3: Significant differences are found in no. of name addressing and IDS quality between the SNOC group and ROC group.**
The use of name addressing and IDS will be restricted for the participants in the ROC group. If the OCCS is sensitive to the change of performance of ostensive communication, it should be reflected on the scores. This hypothesis will be tested with Pair-sample T-test

4. **Hypothesis 4: There is no significant relationship between frequency of ostensive cue and the rate of using ostensive cue.**

The design of OCCS aims to capture the performance of ostensive communicate in terms of the rate of using ostensive cue. It is proposed that the rate of ostensive cue is an independent dimension of ostensive communication from frequency. Therefore, we will expect no significant association between frequency items and rate items in Pearson correlation analysis.

5. **Hypothesis 5: The constructs of mentalisation captured by the parental reflective functioning questionnaires (PRFQ) have significant relationships particularly with the components that constitute contingent responsivity (CR) (i.e. Scaffolding behaviours, contingent behaviours and emotional support), and significant relationships with other ostensive cues in general.**

The capacity to mentalise is considered as the main contributing factor influencing the use of ostensive cues. In the Pearson correlation analysis, we will therefore expect the OCCS scores show a significant negative relationship with the score of pre-mentalising state (PM), a significant positive relationship with the score of Interest and curiosity of mental state (IC), and a non-linear relationship with the score of Certainty of mental state (CM).
6. **Hypothesis 6**: Different dimensions of ostensive communication measured by OCCS have significant negative relationship with PSI-SF

Various researches suggested the potential link between parenting stress level, developmental outcome and ostensive communication. We will therefore expect a significant negative relationship between the OCCS scores and the PSI-SF total score.
Covid-19

Due to the global pandemic of Covid-19, the Anna Freud Centre closed in March 2020. 30 participants had been tested (14 in SNOC group, 5 in ROC group, 5 pilots and 6 data could not be used due to factors such as mother not administering the task correctly, or infant refusing to engage with the task). Considering the low sample size and insufficient power, some of the planned inferential analyses were no longer applicable for the purpose of establishing the psychometric property of OCCS. In order to best utilise the data obtained to do basic test on reliability and validity of the instrument, the research team decided not to use the data from the ROC group but only use the data from the SNOC group. It was also decided to revised some of the hypotheses and employ an alternative methodology in data analysis to test these hypotheses. The data was analysed using a mixed quasai-quantitative and qualitative method aim at investigating individual patterns of scores among the OCCS scales, and in relation to PQRS and PSI-SF. It was a quasai-quantitative method because participants involved in the analysis were selected based on a their score on OCCS, but the patterns were observed in an qualitative way.

Revised hypotheses:

1. Hypothesis 1: OCCS achieves high to excellent level of inter-rater reliability in all sub-scale. (remain unchanged)

   Although with insufficient sample size, there was no viable alternative, the inter-rater reliability of each item in OCCS were still assessed using interclass correlation coefficients (ICC; Shrout & Fleiss, 1979) base on coding of the videotape on Task B.
2. **Hypothesis 2:** There are patterns of scores between the co-occurrence score and the score of frequency, rate, and quality of ostensive cuing

Pearson correlation analysis cannot be used to test the original hypothesis (i.e. *Co-occurrence has positive relationship with all sub-scale frequency, quality and rate of different ostensive cues*). Alternatively, patterns of scores among different dimensions of ostensive cues will be observed with the score of co-occurrence.

3. **Hypothesis 3:** The performance of different scales of OCCS in Task A are considerably different from Task B.

Pair sample T-test cannot be used to test the original hypothesis (i.e. *Significant differences are found in no. of name addressing and IDS quality between the SNOC group and ROC group.*). Alternatively, to test whether the OCCS can reliably distinguish individual with high performance of ostensive communication from those with low performance, the research team decided to make the comparison within the SNOC group. Considering the very different nature of Task A and Task B, intuitively, the mother would use different ostensive cues. Observable pattern of difference between the OCCS score on Task A and Task B could give an indication that the OCCS can reliability capture the change of performance in ostensive communication.

4. **Hypothesis 4:** There is no observable pattern between the performance of frequency items and the rate items

Pearson correlation analysis cannot be used to test the original hypnosis (i.e. *There is no significant relationship between frequency of ostensive cue and the rate of using ostensive cue.*) Alternatively, participants who have the same or very similar performance on frequency items were extracted from the sample, their performance on rate items were observed.
Similarly, participants who have the same or very similar performance on rate items were extracted from the sample, their performance on frequency items were observed.

5. **Hypothesis 5:** The constructs of mentalisation show observable pattern of relationships with components of contingent responsivity and ostensive communication in general.

Pearson correlation analysis cannot be used to test the original hypnosis (i.e. *The constructs of mentalisation captured by the parental reflective functioning questionnaires (PRFQ) have significant relationships particularly with the components that constitute contingent responsivity (CR) (i.e. Scaffolding behaviours, contingent behaviours and emotional support), and significant relationships with other ostensive cues in general.*)

 Alternatively, a number of participants were selected based on their performance on OCCS. Their performance on the scale of pre-mentalising state (PM), Certainty of mental state (CM) and Interest and curiosity of mental state (IC) were compared. We would therefore expect the OCCS scores to show an observable negative pattern with the score of PM, an observable positive relationship with the score of IC, and an observable non-linear relationship pattern with the score of CM.

6. **Hypothesis 6:** Score of PSI-SF shows observable pattern of negative relationship with the performance on OCCS.

Pearson correlation analysis cannot be used to test the original hypnosis (i.e. *Different dimensions of ostensive communication measured by OCCS have significant negative relationship with PSI-SF*). Alternatively, a number of participants were selected based on their performance on OCCS. Their performance on PSI-SF were compared. We would
therefore expect to the OCCS scores to show an observable negative performance with the PSI-SF total score.
## Result

### Demographic characteristics

*Table 10. Sample Demographic information (N=14)*

<table>
<thead>
<tr>
<th>Variables</th>
<th>M (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infant’s age (months)</td>
<td>19.38 (2.14)</td>
</tr>
<tr>
<td>Infant’s gender</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>42.9% (6)</td>
</tr>
<tr>
<td>Mother’s Age (Year)</td>
<td>34.08 (2.78)</td>
</tr>
<tr>
<td>Mother’s ethnicity</td>
<td></td>
</tr>
<tr>
<td>White-British</td>
<td>28.6% (4)</td>
</tr>
<tr>
<td>White-Any other background</td>
<td>57.1% (8)</td>
</tr>
<tr>
<td>Mixed -White &amp; Asian</td>
<td>7.1% (1)</td>
</tr>
<tr>
<td>Chinese</td>
<td>7.1% (1)</td>
</tr>
<tr>
<td>Mother’s level of education</td>
<td></td>
</tr>
<tr>
<td>A-Level, vocational level (e.g. NVQ) 3 or equivalent</td>
<td>7.1% (1)</td>
</tr>
<tr>
<td>Higher education or professional/vocational equivalent</td>
<td>21.4% (3)</td>
</tr>
<tr>
<td>Postgraduate education or professional/vocational equivalent</td>
<td>71.4% (10)</td>
</tr>
<tr>
<td>Household income</td>
<td></td>
</tr>
<tr>
<td>£10,000-20,000</td>
<td>14.3% (2)</td>
</tr>
<tr>
<td>£20,000-35,000</td>
<td>14.3% (2)</td>
</tr>
<tr>
<td>£35,000-50,000</td>
<td>7.1% (1)</td>
</tr>
<tr>
<td>£50,000-75,000</td>
<td>28.6% (4)</td>
</tr>
<tr>
<td>Income Range</td>
<td>Percentage</td>
</tr>
<tr>
<td>--------------------</td>
<td>------------</td>
</tr>
<tr>
<td>£75,000-100,000</td>
<td>7.1%</td>
</tr>
<tr>
<td>£100,000+</td>
<td>21.4%</td>
</tr>
</tbody>
</table>

Use of mental health services

<table>
<thead>
<tr>
<th>Option</th>
<th>Percentage</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>28.6%</td>
<td>(4)</td>
</tr>
<tr>
<td>No</td>
<td>71.4%</td>
<td>(10)</td>
</tr>
</tbody>
</table>

14 mother-infant participants (42.9% male; infant mean age 19.38 months-old, mother mean age 34.08 years-old) took part in the experiments. With regards to ethnicity, 57.1% identified as Any-White background, 28.6% as White-British, 7.1% as Mixed White-Asian and 7.1% as Chinese. In terms of mother participants’ education level, 71.4% reported having postgraduate education or equivalent, 21.4% higher education or equivalent, and 7.1% A-level or equivalent.
### Descriptive statistics

*Table 11. Mean (SD) of Ostensive cue items in Task A and Task B (N=14)*

<table>
<thead>
<tr>
<th>Variable groups</th>
<th>Variables</th>
<th>Task A</th>
<th>Task B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$M$ (SD)</td>
<td>$M$ (SD)</td>
</tr>
<tr>
<td>Time (second)</td>
<td></td>
<td>78.5 (31.86)</td>
<td>491.07 (190.94)</td>
</tr>
<tr>
<td>Attention</td>
<td></td>
<td>1.64 (.30)</td>
<td>1.62 (.39)</td>
</tr>
<tr>
<td>Frequency items</td>
<td>No. of Eye contact</td>
<td>5.92 (2.89)</td>
<td>14.64 (8.81)</td>
</tr>
<tr>
<td></td>
<td>No. of Name addressing</td>
<td>2.42 (.37)</td>
<td>11.64 (9.56)</td>
</tr>
<tr>
<td></td>
<td>No. of Scaffolding Behaviours</td>
<td>19.29 (6.33)</td>
<td>87.64 (21.79)</td>
</tr>
<tr>
<td>Rate items</td>
<td>Rate of Eye contact</td>
<td>0.08 (.03)</td>
<td>0.03 (.02)</td>
</tr>
<tr>
<td></td>
<td>Rate of Name addressing</td>
<td>0.04 (.03)</td>
<td>0.02 (.01)</td>
</tr>
<tr>
<td></td>
<td>Rate of Scaffolding Behaviours</td>
<td>0.26 (.09)</td>
<td>0.19 (.05)</td>
</tr>
<tr>
<td>Global rating Items</td>
<td>IDS quality</td>
<td>1.18 (.37)</td>
<td>1.42 (.26)</td>
</tr>
<tr>
<td></td>
<td>Quality of Contingent Behaviours</td>
<td>0.54 (.38)</td>
<td>1.03 (.36)</td>
</tr>
<tr>
<td></td>
<td>Overall Emotional Support</td>
<td>0.38 (.51)</td>
<td>0.92 (.62)</td>
</tr>
</tbody>
</table>
Co-occurrence

items

| Co-occurrence score | 2.8 (1.11) | 4.02 (1.41) |

As shown in Table 11 above, much less time was spent on Task A is much less than on Task B. Infant participants’ average attention levels in both tasks are quite close. Frequencies of eye contact, name addressing, and scaffolding behaviours used in Task A is considerably lower than in Task B. Participants also demonstrate a better average performance on all the global rating item and co-occurrence in Task B than Task A. The rate of ostensive cueing in Task A is considerably higher than in Task B.
Hypothesis 1: OCCS achieves high to excellent level of inter-rater reliability in all subscale.

Table 12. 
*Interrater reliability of ostensive communication coding system base of Task B (N=14)*

<table>
<thead>
<tr>
<th>Item</th>
<th>ICC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attention</td>
<td>.96</td>
</tr>
<tr>
<td>Frequency items</td>
<td></td>
</tr>
<tr>
<td>No. of Eye contact</td>
<td>.97</td>
</tr>
<tr>
<td>No. of Name addressing</td>
<td>.98</td>
</tr>
<tr>
<td>No. of Scaffolding Behaviours</td>
<td>.93</td>
</tr>
<tr>
<td>Rate items</td>
<td></td>
</tr>
<tr>
<td>Rate of Eye contact</td>
<td>.96</td>
</tr>
<tr>
<td>Rate of Name addressing</td>
<td>.95</td>
</tr>
<tr>
<td>Rate of Scaffolding Behaviours</td>
<td>.87</td>
</tr>
<tr>
<td>Global rating Items</td>
<td></td>
</tr>
<tr>
<td>IDS quality</td>
<td>.82</td>
</tr>
<tr>
<td>Quality of Contingent Behaviours</td>
<td>.92</td>
</tr>
<tr>
<td>Overall Emotional Support</td>
<td>.96</td>
</tr>
<tr>
<td>Co-occurrence items</td>
<td></td>
</tr>
<tr>
<td>Co-occurrence score</td>
<td>.97</td>
</tr>
</tbody>
</table>

Note: ICC = Intraclass correlation coefficient

As reported in table 12, the inter-rater reliability of each item in ostensive communication coding system has been assessed using interclass correlation coefficients
(ICCs; Shrout & Fleiss, 1979). Fourteen videos of Task B, which ranged in length from 177 seconds to 892 seconds ($M = 491.07; SD = 190.94$), were coded independently. According to the Cicchetti (1994), all items in the ostensive communication coding system demonstrates an “excellent” agreement.
One of the research questions regarding the validity of OCCS is whether the instrument is able to distinguish different level of ostensive cues being used in different teaching paradigms. We could have investigated this question by employing the statistical analysis of Pair-Sample T-test and test whether there is significant difference between mean score of each subscale of OCCS in the SNOC group and the ROC group. However, the sample size we have does not provide sufficient statistic power to make the necessary conclusion. Detail observation and comparison of selected individual’s performance of each sub-scale of OCCS in Task A and Tasks B may be a viable alternative.

The premise to make a meaningful observation based on the given data is to have a sound system with which to select participants from the sample for more in-depth investigation. There are ten main variables in the OC. Ideally, every single variable should be reviewed, but it is not viable within the scope of the present study. As a viable alternative, three new variables (ZF, ZR and ZGR, see table 13 below) that indicate the performance of ostensive cues for each participant are generated based on the sum of Z-score of items from the same dimension (i.e. Frequency, rate and global rating). These scores can give some ideas as to how the individual’s ostensive cues performance relate to the sample mean. A positive value indicates the individual has an overall above-average performance on a given dimension. A negative value indicates the individual has an overall below-average performance on a given dimension. The higher the value of the score, the more deviation of the individual performance from the sample mean. Co-occurrence is not included in the composition because it is not measured purely in terms of frequency and rate, and therefore not entirely consistent with either dimension.
Table 13.
Composition of ostensive communication performance score

<table>
<thead>
<tr>
<th>Ostensive communication performance Score</th>
<th>Composing variables</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ZF</strong></td>
<td>Z-score of no. of eye</td>
</tr>
<tr>
<td>Sum of the Z score of contact</td>
<td></td>
</tr>
<tr>
<td>Frequency items</td>
<td></td>
</tr>
<tr>
<td><strong>ZR</strong></td>
<td>Z-score of rate of eye</td>
</tr>
<tr>
<td>Sum of Z score of rate contact</td>
<td></td>
</tr>
<tr>
<td>items</td>
<td></td>
</tr>
<tr>
<td><strong>ZGR</strong></td>
<td>Z-score of IDS</td>
</tr>
<tr>
<td>Sum of Z score of global rating items</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>P34</td>
<td>-2.11</td>
</tr>
<tr>
<td>P35</td>
<td>0.83</td>
</tr>
<tr>
<td>P39</td>
<td>-1.69</td>
</tr>
<tr>
<td>P47</td>
<td>2.28</td>
</tr>
<tr>
<td>P49</td>
<td>2.28</td>
</tr>
<tr>
<td>P51</td>
<td>1.26</td>
</tr>
<tr>
<td>P54</td>
<td>-2.51</td>
</tr>
<tr>
<td>P55</td>
<td>-0.22</td>
</tr>
<tr>
<td>P65</td>
<td>-0.50</td>
</tr>
<tr>
<td>Minimum</td>
<td>-2.51</td>
</tr>
<tr>
<td>Maximum</td>
<td>2.28</td>
</tr>
<tr>
<td>Range</td>
<td>4.79</td>
</tr>
</tbody>
</table>
Hypothesis 2: Participants with high frequency of ostensive cues tend to have high score of co-occurrence.

Table 15.
Correlation of ZF (A) and ZR (A), ZGR (A) and Co-occurrence (A)

<table>
<thead>
<tr>
<th>Spearman's rho</th>
<th>ZF (A)</th>
<th>ZR (A)</th>
<th>ZGR (A)</th>
<th>Co-Occurrence (A)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZF (A)</td>
<td></td>
<td>.055</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZR (A)</td>
<td></td>
<td>.745**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZGR (A)</td>
<td></td>
<td>.121</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Co-Occurrence (A)</td>
<td></td>
<td>-.04</td>
<td></td>
<td>.842**</td>
</tr>
</tbody>
</table>

** Correlation is significant at the 0.01 level (2-tailed).

There was insufficient sample to conduct Person correlation analysis. Spearman correlation analysis was employed to explore the relationships among ZF, ZR, ZGR and co-occurrence in task A. Results of the Spearman correlation in table 15 suggested that there was likely to be a significant positive association between ZF(A) and ZGR(A), ($r_s (13) = .75$, $p < .001$). Co-occurrence was also found to have significant positive relationship with ZF (A) ($r_s (13) = .76$, $p < .001$), and ZGR (A) ($r_s (13) = .84$, $p < .001$).

Table 16.
Comparison of OCCS dimensions in Task A

<table>
<thead>
<tr>
<th>Score Group</th>
<th>Participant no.</th>
<th>ZF(A)</th>
<th>ZR (A)</th>
<th>ZGR (A)</th>
<th>Co-occurrence (A)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lowest</td>
<td>P54</td>
<td>-2.51</td>
<td>1.65</td>
<td>-2.79</td>
<td>1.75(-1.23)</td>
</tr>
<tr>
<td></td>
<td>P24</td>
<td>-2.18</td>
<td>-1.77</td>
<td>-2.73</td>
<td>1.5(-1.47)</td>
</tr>
<tr>
<td>Middle</td>
<td>P55</td>
<td>-0.22</td>
<td>0.55</td>
<td>-0.82</td>
<td>2 (-1.00)</td>
</tr>
<tr>
<td></td>
<td>P35</td>
<td>0.83</td>
<td>0.57</td>
<td>-1.17</td>
<td>3.25 (.17)</td>
</tr>
<tr>
<td>Highest</td>
<td>P47</td>
<td>2.28</td>
<td>-1.12</td>
<td>4.41</td>
<td>4.25 (1.10)</td>
</tr>
<tr>
<td></td>
<td>P49</td>
<td>2.28</td>
<td>0.62</td>
<td>3.38</td>
<td>4.5 (1.33)</td>
</tr>
</tbody>
</table>
Participants are sequenced based on the ZF score in Task A. Two participants who score the lowest and highest, and two participants from the middle of the sequence are extracted. Six participants are extracted from the sample for comparison in table 16 and figure 3a. Figure 3b showed trend for all 14 participants.

Consistent with the spearman correlation analysis, as shown in figure 3a, when ZF score increases, a clear trend of increase is observed in ZGR and co-occurrence. There is no clear trend or pattern observed in ZR. P54 has the lowest score on ZF and ZGR, and a second lowest score on co-occurrence among all participant. P47 has the same highest score on ZF as P49. And P47 also has the second highest in ZGR and co-occurrence among all participant.
Table 17.
Correlation of ZF (B) and ZR (B), ZGR (B) and Co-occurrence (B)

<table>
<thead>
<tr>
<th></th>
<th>ZF (B)</th>
<th>ZR (B)</th>
<th>ZGR (B)</th>
<th>Co-occurrence (B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spearman's rho ZF (B)</td>
<td>.033</td>
<td>.231</td>
<td>.982**</td>
<td></td>
</tr>
<tr>
<td>ZR (B)</td>
<td></td>
<td>-.429</td>
<td>-.024</td>
<td></td>
</tr>
<tr>
<td>ZGR (B)</td>
<td></td>
<td></td>
<td>.41</td>
<td></td>
</tr>
<tr>
<td>Co-Occurrence (B)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

** Correlation is significant at the 0.01 level (2-tailed).

Spearman correlation analysis was also employed to explore the relationships among ZF, ZR, ZGR and co-occurrence in task B. Results of the Spearman correlation in table 17 suggested that there was likely to be a significant positive association between ZF(B) and co-occurrence ($r_s (13) = .98, p < .001$).
**Table 18.**
Comparison of OCCS dimensions in Task B

<table>
<thead>
<tr>
<th>Score</th>
<th>Participant</th>
<th>ZF(B)</th>
<th>ZR (B)</th>
<th>ZGR (B)</th>
<th>Co-occurrence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lowest</td>
<td>P24</td>
<td>-4.3</td>
<td>0.09</td>
<td>-4.32</td>
<td>1.67 (1.79)</td>
</tr>
<tr>
<td></td>
<td>P65</td>
<td>-2.4</td>
<td>-1.27</td>
<td>3.14</td>
<td>2.92 (-1.02)</td>
</tr>
<tr>
<td>Middle</td>
<td>P35</td>
<td>-0.46</td>
<td>-1.02</td>
<td>-0.76</td>
<td>4.25 (-.19)</td>
</tr>
<tr>
<td></td>
<td>P03</td>
<td>0.26</td>
<td>0.53</td>
<td>-1.76</td>
<td>5 (.27)</td>
</tr>
<tr>
<td>Highest</td>
<td>P55</td>
<td>3.03</td>
<td>-1.44</td>
<td>1.06</td>
<td>6.33 (1.09)</td>
</tr>
<tr>
<td></td>
<td>P20</td>
<td>3.2</td>
<td>0.24</td>
<td>2.06</td>
<td>7.33 (1.71)</td>
</tr>
</tbody>
</table>
Figure 4a and 4b
Observation of OCCS dimensions in Task A.

Participants were also sequenced based on the ZF score in Task B. A different six participants are extracted and compared in table 18, figure 4a and 4b. Consistent with the spearman analysis, as ZF (B) increases, a clear trend of increase was observed in co-occurrence. The mild increasing trend in ZGR (B) was not detectable in the spearman test but observable in table 4a. No clear pattern was observed between ZF and ZR in task B.

These results provide support to the hypothesis 2 and suggest a potentially positive relationship between ZF and co-occurrence. In addition, there is also positive relationship
among ZF, ZGR and co-occurrence. The relationship between ZR and the other three dimensions is less clear.
Hypothesis 3: The performance of different dimensions of OCCS in Task A are considerably different from Task B.

Table 19. 
Correlation between ZF, ZR, ZGR and Co-occurrence between Task A and Task B

<table>
<thead>
<tr>
<th>Spearman's rho</th>
<th>ZF (A)</th>
<th>ZR (A)</th>
<th>ZGR (A)</th>
<th>Co-Occurrence (A)</th>
<th>ZF (B)</th>
<th>ZR (B)</th>
<th>ZGR (B)</th>
<th>Co-Occurrence (B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZF (A)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.327</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZR (A)</td>
<td></td>
<td>-1.82</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZGR (A)</td>
<td></td>
<td></td>
<td>.543*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Co-Occurrence (A)</td>
<td></td>
<td></td>
<td>.168</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Correlation is significant at the 0.05 level (2-tailed).

Spearman correlation analysis was also employed to explore the relationships of ZF, ZR, ZGR and co-occurrence in Task A and Task B. Results of the Spearman correlation in table 19 suggested that there was likely to be a significant positive association between ZGR (A) and ZGR (B) ($r_s(13) = .54$, p < .05). Performance of ZF, ZGR and Co-occurrence in Task A and Task B show no relationship with each ot
Table 20.
Comparison of ZF, ZR, ZGR, Co-occurrence and Rate of Co-occurrence scores of Task A and Task

<table>
<thead>
<tr>
<th>Score group</th>
<th>Participant no.</th>
<th>ZF(A)</th>
<th>ZF(B)</th>
<th>ZR (A)</th>
<th>ZR(B)</th>
<th>ZGR (A)</th>
<th>ZGR (B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lowest</td>
<td>P54</td>
<td>-2.51</td>
<td>-2.28</td>
<td>1.65</td>
<td>-0.7</td>
<td>-2.79</td>
<td>-1.75</td>
</tr>
<tr>
<td></td>
<td>P24</td>
<td>-2.18</td>
<td>-4.3</td>
<td>-1.77</td>
<td>0.09</td>
<td>-2.73</td>
<td>-4.32</td>
</tr>
<tr>
<td>Middle</td>
<td>P55</td>
<td>-0.22</td>
<td>3.03</td>
<td>0.55</td>
<td>-1.44</td>
<td>-0.82</td>
<td>1.06</td>
</tr>
<tr>
<td></td>
<td>P35</td>
<td>0.83</td>
<td>-0.46</td>
<td>0.57</td>
<td>-1.02</td>
<td>-1.17</td>
<td>-0.76</td>
</tr>
<tr>
<td>Highest</td>
<td>P47</td>
<td>2.28</td>
<td>0.36</td>
<td>-1.12</td>
<td>-0.95</td>
<td>4.41</td>
<td>2.67</td>
</tr>
<tr>
<td></td>
<td>P49</td>
<td>2.28</td>
<td>1.12</td>
<td>0.62</td>
<td>-0.16</td>
<td>3.38</td>
<td>4.24</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Score group</th>
<th>Participant no.</th>
<th>Co-Occurrence (A) (z)</th>
<th>Co-Occurrence (B) (z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lowest</td>
<td>P54</td>
<td>1.75 (-1.23)</td>
<td>2.58 (-1.23)</td>
</tr>
<tr>
<td></td>
<td>P24</td>
<td>1.5 (-1.47)</td>
<td>1.67 (-1.79)</td>
</tr>
<tr>
<td>Middle</td>
<td>P55</td>
<td>2 (-1)</td>
<td>6.33 (1.09)</td>
</tr>
</tbody>
</table>
Using the same method, participants are sequenced based on the ZF score in task A. Two participants who score the lowest and highest, two participants from the middle of the sequence are extracted. Six participants are extracted from the sample for comparison of their performance in each OCCS dimension in task A and task B. Further observation of each dimension in figure 5a to 8b.
As can be observed in figure 5a and 5b, apart from P54 who shows a similar performance in on ZF in task A and task B, all other participants demonstrate a distinctive level of performance on ZF across two tasks. P47, for example, has a ZF score of 2.28 in Task A, which suggests a considerable above-average performance. However, remain above-average, her ZF score in Task B indicates her at a much closer level of performance with her average counterpart (0.36). Yet, participants who have relatively extreme performance in Task A tend also to have extreme performance in Task B. Mixed performance is found in the middle score group.
Despite considerable variation in the level of performance, a loose increasing trend of ZF in task B can be observed along the increase in ZF in task A. This loose pattern seems to suggest that ZF might be a relatively stable dimension across tasks, which means an individual who demonstrates above-average frequency of ostensive cues in Task A appears more likely to demonstrate the same in Task B.
Distinctive differences are also found among the individual scores on ZR (A) and ZR (B). There is no pattern or trend that can be observed, which suggest the rate of ostensive cue usage might vary not only from individual to individual, but also from task to task.
Although no positive relationship was detected in the spearman correlation analysis, a clear general pattern of increase in ZGR (A) can be observed alongside an increase in ZGR (B). The comparison of ZGR in Task A and Task B shows a similar result as ZF, where individuals display a considerable difference in the level of performance on Task A and Task B, but those who have above-average performance (or below-average) in one task are likely to have a similar performance in the other.
Consistent with the spearman correlation analysis, as can be observed in figure 8a and 8b, individual performance relatively well in co-occurrence in task A is likely to have a similar level of good performance in task B. It is the same for those who has relative poorer performance. Yet, there is also observable difference between the co-occurrence score in task A and in task B, which indicates the co-occurrence scale is also task sensitive.

These results seem to support hypothesis 3 that there are observable differences in individual performance in ZF, ZR ZGR and co-occurrence in different tasks, which suggest these dimensions are task sensitive. And there seem to be potential positive relationships of ZF, ZGR and co-occurrence in task A and task B, which suggests performance in these
dimension is stable across tasks. (i.e. individual who performs well in ZF in task A, is like to also perform well in task B).
Hypothesis 4: There is no observable pattern between the performance of frequency items and the rate items

The initial exploration in the section above has suggested ZF and ZR are measuring two unique dimensions of ostensive communication. Intuitively, even when two mothers have used the same amount of eye contact, name addressing or scaffolding cues, it is likely to lead to a different effect if the intensity of the cues is different. This is investigated by reviewing participants’ raw data.

Table 2.1. Frequency, rate and time

<table>
<thead>
<tr>
<th>Participant no.</th>
<th>No. of eye contact (z)</th>
<th>Rate of eye contact (z)</th>
<th>Time (second) (z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>P65</td>
<td>5 (-1.09)</td>
<td>0.014 (-1.05)</td>
<td>346 (-.76)</td>
</tr>
<tr>
<td>P24</td>
<td>5 (-1.09)</td>
<td>0.028 (-.14)</td>
<td>177 (-1.64)</td>
</tr>
<tr>
<td>P06</td>
<td>12 (-.30)</td>
<td>0.043 (.81)</td>
<td>282 (-1.09)</td>
</tr>
<tr>
<td>P03</td>
<td>22 (.84)</td>
<td>0.047 (1.07)</td>
<td>473 (-.09)</td>
</tr>
</tbody>
</table>

Participants are sequenced based on their number of eye contacts in Task B. A few participants are extracted from the sample.

P65 and P24 use the same number of eye contact, but these eye contacts are distributed over a considerable different time range. The rate of eye contact of P24 is double of P6. Meanwhile, although P06 has the same rate of eye contact as P03, the latter mother has considerable above-average performance in the number of eye contacts when compared to the former one. That is because they spent a noticeable different amount of time on teaching the same task.
Hypothesis 5: The constructs of mentalisation show observable pattern of relationships with components of contingent responsivity and ostensive communication in general.

Due to a small sample size, this hypothesis cannot be tested with parametric correlation analysis. Nevertheless, we may be able to speculate the relationship between the two by looking at individual patterns of scores. Participants are extracted from the sample base on their performances on each subscale of PRFQ in correspondent to the conceptual proposition. Their performance in PRFQ and contingent response components in both tasks are compared. The z-scores in PRFQ scales are generated by comparison between individual score against normative mean and standard deviation drawn from Pazzagli, et al. (2017)’s study. It can provide a better idea as to how different dimensions of OCCS might behave in relation to the population. Patterns and trends are reported.

Table 22.

<table>
<thead>
<tr>
<th>Measures</th>
<th>Sample</th>
<th>Normative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parental Reflective</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Functioning</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-Mentalizing Modes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(PM)</td>
<td>1.61 (.57)</td>
<td>1.84 (.85)</td>
</tr>
<tr>
<td>Certainty about Mental</td>
<td></td>
<td></td>
</tr>
<tr>
<td>States (CM)</td>
<td>4.07 (.84)</td>
<td>4.01 (1.24)</td>
</tr>
<tr>
<td>Interest and Curiosity in</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mental States (IC)</td>
<td>6.08 (.65)</td>
<td>6.08 (.77)</td>
</tr>
</tbody>
</table>

As shown in table 22, the sample mean of all the subscales of parental reflective functioning are close to the normative means reported in Pazzagli et al.’s study, which suggests the participants participated in present study has similar capacity of mentalization in comparison with the wider population.
1. Pre-mentalising state (PM)

Conceptually, a high score on PM generally reflects a non-mentalizing stance, malevolent attribution and an inability to enter the subjective world of the child. Therefore, it is hypothesised that high scores on PM might associate with poorer performance in contingent responsivity. Three participants who score the lowest and three participants who score the highest are extracted from the sample, their score contingent responsivity components are compared (Luyten et al., 2017).

Table 23.
Correlation between Pre-Mentalizing modes (PM) with Scaffolding, Quality of contingent behaviours and emotional support in Task A & Task B

<table>
<thead>
<tr>
<th>Spearman's rho</th>
<th>Pre-Mentalizing Modes</th>
<th>Scaffolding (A)</th>
<th>Contingent behaviours quality (A)</th>
<th>Emotional sup (A)</th>
<th>Scaffolding (B)</th>
<th>Contingent behaviours quality (B)</th>
<th>Emotional support (B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-.056</td>
<td>-.318</td>
<td>-.421</td>
<td>.111</td>
<td>.281</td>
<td>-.251</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Spearman correlation analysis was also employed to explore the relationships between Pre-mentalising modes (PM) and the three sub-dimensions of contingent responsivity (scaffolding, quality of contingent behaviours and emotional support). Results of the Spearman correlation in table 22 suggested that there is no relationship among these constructs.
Table 24.  
Pre-mentalising state (PM) and Contingent responsivity (CR)

<table>
<thead>
<tr>
<th>Score group</th>
<th>Participant no.</th>
<th>PM (z)</th>
<th>Scaffolding (z)</th>
<th>Contingent Behaviours (z)</th>
<th>Emotional Support (z)</th>
<th>Scaffolding (z)</th>
<th>Contingent Behaviours (z)</th>
<th>Emotional Support (z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lowest</td>
<td>P35</td>
<td>1 (-.99)</td>
<td>21 (.27)</td>
<td>0.25 (-.75)</td>
<td>0 (-.76)</td>
<td>96 (.37)</td>
<td>1.11 (.20)</td>
<td>0.67 (-.41)</td>
</tr>
<tr>
<td></td>
<td>P39</td>
<td>1 (-.99)</td>
<td>10 (-1.47)</td>
<td>0.50 (-.09)</td>
<td>0 (-.76)</td>
<td>86 (-.08)</td>
<td>0.47 (-1.16)</td>
<td>0.25 (-1.07)</td>
</tr>
<tr>
<td></td>
<td>P47</td>
<td>1 (-.99)</td>
<td>28 (1.38)</td>
<td>0.92 (1.00)</td>
<td>1.25 (1.71)</td>
<td>77 (-.49)</td>
<td>1.28 (.66)</td>
<td>1.71 (1.27)</td>
</tr>
<tr>
<td></td>
<td>P49</td>
<td>1 (-.99)</td>
<td>22 (.43)</td>
<td>1.25 (1.88)</td>
<td>0.63 (.48)</td>
<td>92 (.19)</td>
<td>1.33 (.81)</td>
<td>2.29 (2.2)</td>
</tr>
<tr>
<td>Highest</td>
<td>P65</td>
<td>2 (.19)</td>
<td>23 (.59)</td>
<td>1.25 (1.88)</td>
<td>0.63 (.48)</td>
<td>79 (-.40)</td>
<td>1.5 (1.27)</td>
<td>1.17 (.40)</td>
</tr>
<tr>
<td></td>
<td>P20</td>
<td>2.17 (.39)</td>
<td>26 (1.06)</td>
<td>0.25 (.75)</td>
<td>0 (-.76)</td>
<td>116 (1.28)</td>
<td>1.53 (1.35)</td>
<td>1.46 (.86)</td>
</tr>
<tr>
<td></td>
<td>P54</td>
<td>3 (1.36)</td>
<td>7 (-1.94)</td>
<td>0.33 (-.53)</td>
<td>0 (-.76)</td>
<td>89 (.06)</td>
<td>1.17 (.36)</td>
<td>0.46 (-.74)</td>
</tr>
</tbody>
</table>
To further observe potential relationship between PM and contingent responsivity, four participants who have the lowest score on PM and three participants who have the highest score on PM are extracted from the sample and listed on table 23. Performance of each component of contingent responsivity is then observed from figure 9a to 11b.
No observable pattern of relationship can be observed between PM and scaffolding in both tasks. Individual with the lowest and highest score of PM seem to have similar level of performance in scaffolding.
There seem to be a mild positive relationship between PM and the quality of contingent behaviours in task B but not task A.
There is an observable negative relationship between PM and emotional support in both tasks when all participants are included.

Individual participants’ performance (see table 23) is further reviewed in conjunction with video data. For example, P47 and P49 performed exceptionally well on most of the CR components. For example, P47 had five out of six scores higher than the average.
performance than other mothers, four among these scores exceeded at least 1 SD. Her performance in scaffolding in Task A was the highest among all participant. In reviewing the videotape, it was evidenced that she used way more hand gestures, exaggerated facial expression, labelling of object and body part (head), questions and narrations than her counterparts in Task A. She also smiled and often laughed, which created a very warm atmosphere. There was a strong sense of connection between her and her toddler.

P49 had all six above-average scores. She had the highest score on emotional support in Task B among all participants, and a second highest score in quality of contingent behaviours in Task A. In the video data, this mother demonstrated a lot more smile and laugh than average, what makes her stood out was that she also used physical touch and gestures (e.g. leaning forward) more than average, which conveyed very warm feeling.

P35 showed mixed performance, which was quite different from P47 and P49. She used a similar level of scaffolding behaviours. She had mixed performance in contingent behaviours and considerably less emotionally supportive than average mothers. She was quite serious in both tasks. There was less sense of joy, and seemingly she was eager to have her toddler to accomplish the tasks.

Although P65 and P20 showed a relatively higher scores on PM in comparison to other participants, their PM scores were also very close to mean, therefore not distinctively high. Their performance on the overall contingent responsivity were similar to P49 and P47 as most of their scores were above mean. P65 showed outstanding performance in contingent behaviours in Task A, while P20 showed particularly good performance in scaffolding behaviours in both tasks.

P54 was the only participant who showed a distinctive high score on PM. She demonstrated low emotional support on both tasks, and a major decrease in the use of scaffolding behaviours in task A. Four out of six scores are below mean suggested her capacity to contingent response was consistently lower than the average mothers. P54 spent
only 34 seconds ($z = -1.40$) in Task A, which was the second lowest among all participants, and 320 seconds ($z = -.90$) on Task B. She swiftly through task A in particular without giving much room for her child to react to the demonstration with communication cue or emotional expression.

In sum, the result of this comparison provided a loose support to the hypothesis that relatively low to medium score on PM seems to connect to better performance in overall contingent responsivity. In contrast, distinctive high score on PM showed a considerable reduction in ability to respond contingently.
2. Certainty of Mental State (CM)

Conceptually, extreme high score of CM indicates a tendency of mothers to be overly certain about the mental state of her child, which reflects the inability to recognise the opacity of mental states and results in hypermentalising; while an extreme low score of CM indicates an almost completely lack of certainty about the child’s mental state and results in hypomentalising (Luyten et al., 2017). It is therefore hypothesised that participants who have a relatively higher and lower score on CM will have relatively poorer performance on components of CR.

Table 25.
Correlations between Certainty of Mental state (CM) with Scaffolding, Quality of contingent behaviours and emotional support in Task A & Task B

<table>
<thead>
<tr>
<th></th>
<th>Scaffolding (A)</th>
<th>Contingent behaviours quality (A)</th>
<th>Emotional sup (A)</th>
<th>Scaffolding (B)</th>
<th>Contingent behaviours quality (B)</th>
<th>Emotional support (B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spearman's rho</td>
<td>-.06</td>
<td>.125</td>
<td>.146</td>
<td>.116</td>
<td>-.251</td>
<td>.364</td>
</tr>
</tbody>
</table>

Spearman correlation analysis was also employed to explore the relationships between certainty about mental states (CM) and the three sub-dimensions of contingent responsivity (scaffolding, quality of contingent behaviours and emotional support). Results of the Spearman correlation in table 24 suggested that there is no relationship among these constructs.
Table 26.

Certainty of Mental State (CM) and Contingent responsiveness (CR)

<table>
<thead>
<tr>
<th>Score group</th>
<th>Participant</th>
<th>CM (z)</th>
<th>Task A</th>
<th>Scaffolding (z)</th>
<th>Contingent Behaviours (z)</th>
<th>Emotional Support (z)</th>
<th>Task B</th>
<th>Scaffolding (z)</th>
<th>Contingent Behaviours (z)</th>
<th>Emotional Support (z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lowest</td>
<td>P24</td>
<td>3 (-.81)</td>
<td>14 (-.84)</td>
<td>0.17 (-.97)</td>
<td>0 (-.76)</td>
<td>35 (-2.4)</td>
<td>0.47 (-1.56)</td>
<td>0 (-.148)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>P51</td>
<td>3 (-.81)</td>
<td>27 (1.22)</td>
<td>0.42 (-.31)</td>
<td>0 (-.76)</td>
<td>100 (0.56)</td>
<td>1.33 (.81)</td>
<td>0.42 (-.81)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle</td>
<td>P55</td>
<td>3.83 (-.15)</td>
<td>16(-.52)</td>
<td>0.17(-.97)</td>
<td>0.63 (.48)</td>
<td>129(1.88)</td>
<td>0.92(-.33)</td>
<td>1.13 (.33)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>P15</td>
<td>4 (-.01)</td>
<td>17 (-.36)</td>
<td>0.58 (.13)</td>
<td>0.88 (.97)</td>
<td>89 (.06)</td>
<td>1.25 (.59)</td>
<td>0.92 (0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Highest</td>
<td>P03</td>
<td>5.33 (1.06)</td>
<td>23 (.59)</td>
<td>0.83 (.78)</td>
<td>1.38 (1.96)</td>
<td>79 (-.40)</td>
<td>0.67 (-1.02)</td>
<td>1.21 (.46)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>P34</td>
<td>5.67 (1.34)</td>
<td>15 (-.68)</td>
<td>0.25 (-.75)</td>
<td>0 (-.76)</td>
<td>95 (.33)</td>
<td>0.75 (-.79)</td>
<td>0.75 (-.27)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Based on the conceptual framework, participants with the lowest, middle and the highest participants of CM are extracted from the sample in table 25. Their performance in scaffolding, CB quality and emotional support are further observed and compared in figure 12a to 14.
A very loose pattern of anti-V shape is observed. Participants scored the lowest and highest in CM tend to have slightly poorer performance in scaffolding than those who have middle (optimal) score of CM.
Similarly, a very loose pattern of anti-V shape is observed. Participants scored the lowest and highest in CM tend to have slightly poorer performance in CB quality than those who have middle (optimal) score of CM.
Figure 14a and 14b
Observation of the relationship between CM and emotional support in Task A and Task B.

Similarly, a very loose pattern of anti-V shape is observed. Participants scored the lowest and highest in CM tend to have slightly poorer performance in emotional than those who have middle (optimal) score of CM.
Participants’ individual performance (see table 25) is further reviewed in conjunction with video data. For example, P24, one of the participants who scored lowest on CM, showed a below-average performance on all CR components in both tasks. She scores particular low on scaffolding behaviours (34, $z = -2.4$) and the quality of contingent behaviours (0.47, $z = -1.56$) in task B; both are the lowest among all participants. The way she taught in both tasks was very mechanical with little variation in expression, there was very little sense of warmth. On a few occasions, she was intrusive and tried to removed objects from her toddler forcefully.

P51, although she demonstrated a relatively higher level of scaffolding behaviour in task A and somewhat good level in task B comparing with her counterparts, her performance on half of the CR components are below-average.

P34 scores the highest on CM, her scores on five out of six CR components are below the average mothers, which suggests a consistent relatively poor performance. This mother rarely used contingent responses like expanding or joint attention. She occasionally responded to her toddler’s communication cues, but the exchanges were short-lived.

As the scores converge to the middle, participants showed more above-average performance on CR components. For example, P15 showed above-average performance on five out of six CR components, reflects a consistent. P35 also has above-average performance on half of the six CR components.

In sum, the result of this comparison provided a loose support to the hypothesis.
3. Interest and Curiosity in mental state (IC)

Table 27. Correlations between Interest and Curiosity in mental state (IC) with Scaffolding, Quality of contingent behaviours and emotional support in Task A & Task B

<table>
<thead>
<tr>
<th>Spearman's rho</th>
<th>PRFQ Interest and Curiosity in Mental States</th>
<th>Scaffolding (A)</th>
<th>Contingent behaviours quality (A)</th>
<th>Emotional sup (A)</th>
<th>Scaffolding (B)</th>
<th>Contingent behaviours quality (B)</th>
<th>Emotional support (B)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>.307</td>
<td>.404</td>
<td>.058</td>
<td>.35</td>
<td>.465</td>
<td>.236</td>
</tr>
</tbody>
</table>

Spearman correlation analysis was also employed to explore the relationships between Interest and curiosity in mental state (IC) and the three sub-dimensions of contingent responsivity (scaffolding, quality of contingent behaviours and emotional support). Results of the Spearman correlation in table 26 suggested that there is no relationship among these constructs.
Table 28.

Interest and Curiosity in mental state and Contingent responsivity

<table>
<thead>
<tr>
<th>Score group</th>
<th>Participant</th>
<th>IC (z)</th>
<th>Task A</th>
<th>Quality of Contingent Behaviours (z)</th>
<th>Emotional Support (z)</th>
<th>Task B</th>
<th>Quality of Contingent Behaviours (z)</th>
<th>Emotional Support (z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lowest</td>
<td>P24</td>
<td>4.83 (-1.62)</td>
<td>14 (-.84)</td>
<td>0.17 (-.97)</td>
<td>0 (-.76)</td>
<td>35 (-2.4)</td>
<td>0.47 (-.156)</td>
<td>0 (-.148)</td>
</tr>
<tr>
<td></td>
<td>P54</td>
<td>5.17 (-1.18)</td>
<td>7 (-1.94)</td>
<td>0.33 (-.53)</td>
<td>0 (-.76)</td>
<td>89 (.06)</td>
<td>1.17 (.36)</td>
<td>0.46 (-.74)</td>
</tr>
<tr>
<td></td>
<td>P55</td>
<td>5.5 (-.75)</td>
<td>16 (-.52)</td>
<td>0.17 (-.97)</td>
<td>0.63 (.48)</td>
<td>129 (1.88)</td>
<td>0.92 (-.33)</td>
<td>1.13 (0.33)</td>
</tr>
<tr>
<td>Highest</td>
<td>P39</td>
<td>6.67 (.77)</td>
<td>10 (-1.47)</td>
<td>0.5 (-.09)</td>
<td>0 (-.76)</td>
<td>86 (-.08)</td>
<td>0.47 (-1.56)</td>
<td>0.25 (-1.07)</td>
</tr>
<tr>
<td></td>
<td>P20</td>
<td>6.83 (.97)</td>
<td>26 (1.06)</td>
<td>0.25 (-.75)</td>
<td>0 (-.76)</td>
<td>116 (1.28)</td>
<td>1.53 (1.35)</td>
<td>1.46 (.86)</td>
</tr>
<tr>
<td></td>
<td>P15</td>
<td>7 (1.19)</td>
<td>17 (-.36)</td>
<td>0.58 (.13)</td>
<td>0.88 (.97)</td>
<td>89 (.06)</td>
<td>1.25 (.59)</td>
<td>0.92 (0)</td>
</tr>
</tbody>
</table>

Three participants with the lowest and highest score of IC are extracted from the sample for comparison of their performance of contingent responsivity in table 27. Their performance in scaffolding, CB quality and emotional support are further observed and compared in figure 15a to 17b.
Figure 15a and 15b
Observation of the relationship between IC and Scaffolding in Task A and Task B.

A mild positive relationship can be observed between IC and scaffolding in both tasks.
Figure 16a and 16b
Observation of the relationship between IC and CB quality in Task A and Task B.

A mild positive relationship can be observed between IC and CB quality in both tasks.
Figure 17a and 17b
Observation of the relationship between IC and Emotional support in Task A and Task B.

A mild positive relationship can be observed between IC and emotional support in task B but not in task A.

Low level of IC might reflect a lack of interest and curiosity in the child’s mental state, which significantly contributes to the state of hypomentalising. While a relatively higher score on IC usually reflects a good capacity of mentalisation. But extremely high score might reflect intrusive hypermentalising.

Participants’ individual performance (see table 28) is further reviewed in conjunction with video data. For example, P24, who reported the lowest CM score, also showed the lowest IC score among all participants. As reported earlier, her performance on all the CR

186
component in both tasks is lower than the average mother. P54, who has the highest score on PM, is the second lowest on IC. Her performances on scaffolding and her quality of contingent behaviours in Task B are slightly above, the rest is below-average. In fact, she has the lowest number of scaffolding behaviours in Task A among all participants.

With a slightly higher score on IC, P55 shows a general improvement in her scores on the CR components. Her emotional support in both tasks are above average. She also used highest number of scaffolding among all participants.

Although P39 has a relatively high score on IC, none of her scores on CR components is above mean. Her performance of scaffolding in Task A, and quality of contingent behaviours in Task B are one of the lowest among the participants. In the video, P39 seemed to be Impatient in both tasks and there was minimal sense of warmth. The way she taught her toddler was lacked in variation. She did not demonstrate much attempt to mentalise, she does not come across as someone who has high interest and curiosity of her toddler’s mental state.

P20 and P15 have the highest scores on IC. There are observable improvements in general among the CR components in both tasks. Although she has below-average performance in the quality of contingent behaviours and emotional support, P20 demonstrates an outstanding performance in scaffolding in both tasks. She also has the highest score on the quality of contingent behaviours among all participant in task B. P20 gently taught her toddler, she took her time and gave sufficient space for him to respond with cues. She demonstrated outstanding effort in trying to meet where the toddler was cognitively and emotionally. As for P15, she has the extremely high score on IC, yet, five out of six of her score on CR components are above mean, which indicate a consistency of higher than average performance in CR. In the video, she does not come across as a mother with hypermentalisation.
Although the performance of P24, P54 and P20 provide some support to the hypothesis, P39 and P15 did not perform as expected. It is inconclusive whether the hypothesis is loosely accepted or rejected.
Hypothesis 5 continue, and hypothesis 6: Score of PSI-SF shows observable pattern of negative relationship with the different dimensions of OCCS

Similarly, correlation analysis among other dimensions of OCCS, PQRS and PSI-SF cannot be implemented due to insufficient sample size. Therefore, to speculate the relationships among them, we employed the same methodology to look at individual pattern of scores among these constructs.

Table 29.
Mean (SD) of Parenting Stress Index – short form (PSI-SF)

<table>
<thead>
<tr>
<th>Measures</th>
<th>Sample</th>
<th>Normative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parenting Stress Index (PSI-SF)</td>
<td>67.07(23.47)</td>
<td>79.00</td>
</tr>
</tbody>
</table>

As shown in table 23, the sample mean of PSI-SF in is noticeably lower than the normative mean reported in Mary et al.'s study. It suggests that the sample of this study might not sufficiently represent the wider population.

Table 30.
Summary of observable patterns of scores among OCCS, PRFQ and PSI-SF

<table>
<thead>
<tr>
<th>Ostensive cue performance score</th>
<th>PRFQ</th>
<th>PSI-SF</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZF (A)</td>
<td>PM</td>
<td>CM</td>
</tr>
<tr>
<td>-VE</td>
<td>+VE</td>
<td>+VE</td>
</tr>
</tbody>
</table>
Table 29 summarises the observable relationships among ostensive cue performance scores, the three sub-scales of PRFQ, as well as the PSI.

The capacity to mentalise had a more noticeable overall influence on the mother’s use of ostensive cues in Task A. It’s influence in task B was less clear. Mother’s interest and curiosity to her child’s mental state (IC) seemed to have the most noticeable impact on different dimensions of ostensive communication. The Pre-mentalising scale seemed able to loosely predict some variance of the quality of ostensive communication, where hyper-mentalisation caused by the inability to enter the child’s subjective world appeared to reduce the quality of ostensive communication.

The parenting stress level (PSI) also demonstrated some ability to predict the performance of ostensive communication. The table suggested that high parenting stress might be less favourable for ostensive communication.

Detail explore of the relationships among these constructs are presented below in table 30 to table 38. Observation of trends and patterns can be found in figure 18a to 32b below.
Table 31. 
Correlations between Interest and Curiosity in mental state (IC) with Scaffolding, Quality of contingent behaviours and emotional support in Task A & Task B

<table>
<thead>
<tr>
<th>Spearman's rho</th>
<th>Pre-Mentalizing Modes (PM)</th>
<th>Certainty about Mental States (CM)</th>
<th>Interest and Curiosity in Mental States (IC)</th>
<th>Parenting Stress Index – short form (PSI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ZF (A)</td>
<td>ZR (A)</td>
<td>ZGR (A)</td>
<td>ZF (B)</td>
</tr>
<tr>
<td>Pre-Mentalizing Modes (PM)</td>
<td>-.424</td>
<td>.261</td>
<td>-.524</td>
<td>-.536*</td>
</tr>
<tr>
<td>Certainty about Mental States (CM)</td>
<td>.157</td>
<td>-.428</td>
<td>.086</td>
<td>.285</td>
</tr>
<tr>
<td>Interest and Curiosity in Mental States (IC)</td>
<td>.46</td>
<td>.073</td>
<td>.287</td>
<td>.466</td>
</tr>
<tr>
<td>Parenting Stress Index – short form (PSI)</td>
<td>-.436</td>
<td>.396</td>
<td>-.22</td>
<td>-.469</td>
</tr>
<tr>
<td>Co-Occurrence (A)</td>
<td>-.569*</td>
<td>.348</td>
<td>.610*</td>
<td>-.392</td>
</tr>
<tr>
<td>Co-Occurrence (B)</td>
<td>.668*</td>
<td>.356</td>
<td>.418</td>
<td>-.268</td>
</tr>
</tbody>
</table>

* Correlation is significant at the 0.05 level (2-tailed).

Spearman correlation analysis was also employed to explore the relationships between PRFQ, PSI and different dimensions of OCCS. Results of the Spearman correlation in table 28 suggested that there is likely to be a significant negative association between PM and Co-occurrence (A) ($r_s (13) = -.56, p < .05$). A significant positive relationship is found between CM and ZF (B) ($r_s (13) = .66, p < .05$). CM was also found to be positively associated with Co-Occurrence (B) ($r_s (13) = .61, p < .05$).
1. Comparison of ostensive cue performance scores with Parenting reflective function (PRFQ) and Parenting stress index (PSI).

The following tables sequences participants base on their score on ZF, ZR, ZGR, and co-occurrence with the same methodology, two participants who score the lowest, two participants who score the highest, and two participants who’s score in the middle of the sequence are extracted from the sample. Their performance on the Pre-mentalising state scale (PM), Certainty of mental state (CM), Interest and curiosity of mental state scale (IC) and Parenting stress index (PSI) are observed and compared. Trends and patterns are explored in figures 18a to 32b.

ZF (A)

Table 32.
The relationship between ZF (A), Parenting reflect functioning and Parenting Stress Index

<table>
<thead>
<tr>
<th>Score group</th>
<th>Participant</th>
<th>ZF(A)</th>
<th>PM (z)</th>
<th>CM (z)</th>
<th>IC (z)</th>
<th>PSI (z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lowest</td>
<td>P54</td>
<td>-2.51</td>
<td>3 (1.36)</td>
<td>3.5 (-.41)</td>
<td>5.17 (-1.88)</td>
<td>133 (2.95)</td>
</tr>
<tr>
<td></td>
<td>P24</td>
<td>-2.18</td>
<td>1.83 (-.01)</td>
<td>3 (-.81)</td>
<td>4.83 (-1.62)</td>
<td>67 (-.66)</td>
</tr>
<tr>
<td>Middle</td>
<td>P55</td>
<td>-0.22</td>
<td>1.5 (-.40)</td>
<td>3.83 (-.15)</td>
<td>5.5 (-.75)</td>
<td>96 (.93)</td>
</tr>
<tr>
<td></td>
<td>P35</td>
<td>0.83</td>
<td>1 (-.99)</td>
<td>3.83 (-.15)</td>
<td>6.33 (.32)</td>
<td>46 (-1.80)</td>
</tr>
<tr>
<td>Highest</td>
<td>P47</td>
<td>2.28</td>
<td>1 (-.99)</td>
<td>3.83 (-.15)</td>
<td>5.67 (-.53)</td>
<td>44 (-1.91)</td>
</tr>
<tr>
<td></td>
<td>P49</td>
<td>2.28</td>
<td>1 (-.99)</td>
<td>4.83 (.66)</td>
<td>6.5 (.55)</td>
<td>43 (-1.97)</td>
</tr>
</tbody>
</table>
Figure 18a and 18b
Observation of the relationship between PRFQ and ZF(A).
As the ZF (A) score increases, the score on the PM scale PSI show a decreasing trend. For example, the PM score decreases from 3.00 ($z = 1.36$) and 1.84 ($z = -0.01$) in the low ZF (A) score group, to 1.50 ($z = -0.400$) and 1.00 ($z = -0.99$) in the middle ZF (A) score group and is maintained at the 1.00 ($z = -0.99$) in the high ZF (A) score group. The decreasing trend of PSI is less clear but observable. Participants in the lowest ZF (A) score group show a distinctive difference in PSI score from those in the highest ZF score group.

Although there is mixed performance in the middle group and the highest group, a less robust but observable trend of increase is identified in the IC scale. There is a distinctive difference in IC score between the lowest ZF(A) score group (e.g. P54, 5.17 ($z = -1.18$)) and the highest ZF (A) score group (e.g. P49 (6.50, $z = 0.55$).
Meanwhile, as the ZF(A) score increases, the CM scale shows a very mild increasing trend. The CM score first shows a tendency to increase from 3.5\((z = -0.41)\) and 3 \((z = -0.81)\) in the lowest ZF (A) score group to 3.83 \((z = -0.15)\) in the middle group and then further increased to 4.83 \((z = -0.66)\) in the highest ZF score group.

**ZR (A)**

*Table 33.*
*The relationship between ZR (A), Parenting reflect functioning and Parenting Stress Index*

<table>
<thead>
<tr>
<th>Score group</th>
<th>Participant no.</th>
<th>ZR (A)</th>
<th>PM ((z))</th>
<th>CM ((z))</th>
<th>IC ((z))</th>
<th>PSI ((z))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lowest</td>
<td>P24</td>
<td>-1.77</td>
<td>1.93 (-.01)</td>
<td>3.00 (-.81)</td>
<td>4.83 (-1.62)</td>
<td>67 (-.66)</td>
</tr>
<tr>
<td></td>
<td>P47</td>
<td>-1.12</td>
<td>1.00 (-.99)</td>
<td>3.83 (-.15)</td>
<td>5.67 (-.53)</td>
<td>44(-1.91)</td>
</tr>
<tr>
<td>Middle</td>
<td>P03</td>
<td>-1.07</td>
<td>1.33 (-.60)</td>
<td>5.33 (1.06)</td>
<td>5.83 (-.32)</td>
<td>69(-.55)</td>
</tr>
<tr>
<td></td>
<td>P39</td>
<td>0.12</td>
<td>1.00 (-.99)</td>
<td>4.67 (.53)</td>
<td>6.67 (.77)</td>
<td>61(-.98)</td>
</tr>
<tr>
<td>Highest</td>
<td>P55</td>
<td>0.55</td>
<td>1.50 (-.40)</td>
<td>3.83 (-.15)</td>
<td>5.50 (-.75)</td>
<td>96(.93)</td>
</tr>
<tr>
<td></td>
<td>P54</td>
<td>1.65</td>
<td>3.00 (1.36)</td>
<td>3.50 (-.41)</td>
<td>5.17(-1.18)</td>
<td>133(2.95)</td>
</tr>
</tbody>
</table>
Figure 20a and 20b

Observation of the relationship between PRFQ and ZR(A).

As the ZR (A) score increases, scores on the PM, CM and IC scale show a mixed performance among the three groups. There is no clear pattern identified.
The scores on PSI scale shows a clear increasing trend from 55.00 ($z = -1.31$) and 58.00 ($z = -1.15$) in the lowest ZR(A) group, to 66.00 ($z = -0.71$) and 61.00 ($z = -0.98$) in the middle ZR (A) score group, and further increase to 72.00 ($z = -0.38$) in the highest ZR (A)score group.
ZGR (A)

Table 34.
The relationship between ZGR (A), Parenting reflect functioning and Parenting Stress Index

<table>
<thead>
<tr>
<th>Score group</th>
<th>no.</th>
<th>ZGR(A)</th>
<th>PM (z)</th>
<th>CM (z)</th>
<th>IC (z)</th>
<th>PSI (z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lowest</td>
<td>P54</td>
<td>-2.79</td>
<td>3 (1.36)</td>
<td>3.5 (-.41)</td>
<td>5.17 (-1.18)</td>
<td>133 (2.95)</td>
</tr>
<tr>
<td></td>
<td>P24</td>
<td>-2.73</td>
<td>1.83 (-.01)</td>
<td>3 (-.81)</td>
<td>4.83 (-1.62)</td>
<td>67 9.66</td>
</tr>
<tr>
<td>Middle</td>
<td>P39</td>
<td>-0.84</td>
<td>1 (-.99)</td>
<td>4.67 (0.53)</td>
<td>6.67 (0.77)</td>
<td>61 (-.98)</td>
</tr>
<tr>
<td></td>
<td>P55</td>
<td>-0.82</td>
<td>1.5 (-.40)</td>
<td>3.83 (-.15)</td>
<td>5.5 (-.75)</td>
<td>96 (.93)</td>
</tr>
<tr>
<td>Highest</td>
<td>P03</td>
<td>3.93</td>
<td>1.33 (-.60)</td>
<td>5.33 (1.06)</td>
<td>5.83 (-.32)</td>
<td>69 (-.55)</td>
</tr>
<tr>
<td></td>
<td>P47</td>
<td>4.41</td>
<td>1 (-.99)</td>
<td>3.83 (-0.15)</td>
<td>5.67 (-.53)</td>
<td>44 (-1.91)</td>
</tr>
</tbody>
</table>
As the ZGR (A) score increases, the scores on the PM scale showed a loosely decreasing trend from 3.00 ($z = 1.36$) and 1.83 ($z = -0.01$) in the lowest ZGR (A) score group, to 1.5 ($z = -0.4$) and 1.00 ($z = -0.99$) in the middle ZGR score group, and maintain in a similar level of 1.33 ($z = -0.60$) and 1.00 ($z = -0.99$) in the highest ZGR score group.

The score of the IC scale shows a loosely increasing trend. The IC score increase from 4.83 ($z = -1.62$) and 5.17 ($z = -1.18$) in the lowest ZGR score group, to 5.50 ($z = -0.75$) in the middle ZGR score group, and further raise to 5.67 ($z = -0.53$) and 5.83 ($z = -0.32$) in the highest ZGR score group. Although P39 shows a score of 6.67 ($z = 0.77$) in the middle ZGR score group, which is relatively higher than her counterpart, the score is consistent with the increasing trend.
There are mixed performance in the CM scale among the three groups, no clear trend is identified.

Figure 23a and 23b

Observation of the relationship between PSI and ZGR(A).

![Graph showing relationship between PSI and ZGR(A)]

Meanwhile, although P55 shows a score of 96.00 ($z = 0.93$) in the middle ZGR (A) score group that deviates from the pattern, the score of the PSI scale also shows a loosely decreasing trend from the 133.00 ($z = 2.95$), which is the highest among all participants, in the low ZGR (A) score group, to 61.00 ($z = -0.98$) in the middle ZGR score group. And it further decreases to 69.00 ($z = -0.55$) and 44.00 ($z = -1.91$), which is the lowest among all participates, in the highest ZGR (A) score group.
Table 35.
The relationship between Co-occurrence (A), Parenting reflect functioning and Parenting Stress Index

<table>
<thead>
<tr>
<th>Score group</th>
<th>Participant</th>
<th>Co-Occurrence</th>
<th>PM (z)</th>
<th>CM (z)</th>
<th>IC (z)</th>
<th>PSI (z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lowest</td>
<td>P24</td>
<td>1.5 (-1.47)</td>
<td>1.83 (-.01)</td>
<td>3 (-.81)</td>
<td>4.83 (-1.62)</td>
<td>67 (-.66)</td>
</tr>
<tr>
<td></td>
<td>P34</td>
<td>1.75 (-1.23)</td>
<td>1.67 (-.20)</td>
<td>5.67 (1.34)</td>
<td>6.17 (.12)</td>
<td>55 (-1.31)</td>
</tr>
<tr>
<td>Middle</td>
<td>P06</td>
<td>3 (-.07)</td>
<td>1.67 (-.20)</td>
<td>4.17 (.13)</td>
<td>5.67 (-.53)</td>
<td>57 (-1.20)</td>
</tr>
<tr>
<td></td>
<td>P51</td>
<td>3 (-.07)</td>
<td>2 (.19)</td>
<td>3 (-.81)</td>
<td>6.5 (.55)</td>
<td>72 (-.38)</td>
</tr>
<tr>
<td>Highest</td>
<td>P49</td>
<td>4.5 (1.33)</td>
<td>1 (-.99)</td>
<td>4.83 (.66)</td>
<td>6.5 (.55)</td>
<td>43 (-1.97)</td>
</tr>
<tr>
<td></td>
<td>P03</td>
<td>4.75 (1.57)</td>
<td>1.33 (-.60)</td>
<td>5.33 (1.06)</td>
<td>5.83 (-.32)</td>
<td>69 (-.55)</td>
</tr>
</tbody>
</table>

Co-occurrence (A)
As the co-occurrence (A) score increases, although there was a participant in the middle score group slightly deviated from the trend, the score on the PM scale showed a loosely decreasing trend from 1.83 (-.01) and 1.67 (-.20) in the lowest co-occurrence score group, to 1.00 (-.99) and 1.33 (-.60) in the highest co-occurrence score group.

Meanwhile, mild increasing trends can be observed both in CM and IC as Co-occurrence (A) score increases.
Figure 25a and 25b
Observation of the relationship between PSI and Co-occurrence (A).

A mild decreasing trend can be observed in PSI scores as co-occurrence (A) decreases.
### ZF (B)

**Table 36. The relationship between ZF (B), Parenting reflect functioning and Parenting Stress Index**

<table>
<thead>
<tr>
<th>Score group</th>
<th>Participant no.</th>
<th>ZF(B)</th>
<th>PM (z)</th>
<th>CM (z)</th>
<th>IC (z)</th>
<th>PSI (z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lowest</td>
<td>P24</td>
<td>-4.30</td>
<td>1.83 (-.01)</td>
<td>3 (-.81)</td>
<td>4.83 (-1.62)</td>
<td>67 (-.66)</td>
</tr>
<tr>
<td></td>
<td>P65</td>
<td>-2.40</td>
<td>2 (.19)</td>
<td>3 (-.81)</td>
<td>6.5 (.55)</td>
<td>72 (.38)</td>
</tr>
<tr>
<td>Middle</td>
<td>P35</td>
<td>-0.46</td>
<td>1 (-.99)</td>
<td>3.83 (-.15)</td>
<td>6.33 (.32)</td>
<td>46 (-1.80)</td>
</tr>
<tr>
<td></td>
<td>P03</td>
<td>0.26</td>
<td>1.33 (-.60)</td>
<td>5.33 (1.06)</td>
<td>5.83 (-.32)</td>
<td>69 (-.55)</td>
</tr>
<tr>
<td>Highest</td>
<td>P55</td>
<td>3.03</td>
<td>1.5 (-.40)</td>
<td>3.83 (-.15)</td>
<td>5.5 (-.75)</td>
<td>96 (.93)</td>
</tr>
<tr>
<td></td>
<td>P20</td>
<td>3.20</td>
<td>2.17 (.39)</td>
<td>4.33 (.26)</td>
<td>6.83 (.97)</td>
<td>66 (-.71)</td>
</tr>
</tbody>
</table>
A very mild anti-V shaped trend can be observed in CM as the ZF(B) score increases, while there is no observable trends found in PM and IC.
No noticeable trend is identified in PSI as the ZF(B) scores increases.
Table 37. The relationship between ZR (B), Parenting reflect functioning and Parenting Stress Index

<table>
<thead>
<tr>
<th>Score group</th>
<th>Participant no.</th>
<th>ZR(B)</th>
<th>PM (z)</th>
<th>CM (z)</th>
<th>IC (z)</th>
<th>PSI (z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lowest</td>
<td>P55</td>
<td>-1.44</td>
<td>1.5 (-.40)</td>
<td>3.83 (-.15)</td>
<td>5.5 (-.75)</td>
<td>96 (.93)</td>
</tr>
<tr>
<td></td>
<td>P65</td>
<td>-1.27</td>
<td>2 (.19)</td>
<td>3 (-.81)</td>
<td>6.5 (.55)</td>
<td>72 (-.38)</td>
</tr>
<tr>
<td>Middle</td>
<td>P49</td>
<td>-0.16</td>
<td>1 (-.99)</td>
<td>4.83 (.66)</td>
<td>6.5 (.55)</td>
<td>43 (-1.97)</td>
</tr>
<tr>
<td></td>
<td>P24</td>
<td>0.09</td>
<td>1.83 (-.01)</td>
<td>3 (-.81)</td>
<td>4.83 (-1.62)</td>
<td>67 (-.66)</td>
</tr>
<tr>
<td>Highest</td>
<td>P06</td>
<td>0.95</td>
<td>1.67 (-.20)</td>
<td>4.17 (.13)</td>
<td>5.67 (-.53)</td>
<td>57 (-1.20)</td>
</tr>
<tr>
<td></td>
<td>P39</td>
<td>3.19</td>
<td>1 (-.99)</td>
<td>4.67 (.53)</td>
<td>6.67 (.77)</td>
<td>61 (-.98)</td>
</tr>
</tbody>
</table>
Figure 28a and 28b
Observation of the relationship between PRFQ and ZR(B).
Figure 29a and 29b

Observation of the relationship between PSI and ZR(B).

As the ZR score increases, all the scale shows a mixed performance, no observable trend is identified.
**ZGR (B)**

*Table 38.*

*The relationship between ZGR (B), Parenting reflect functioning and Parenting Stress Index*

<table>
<thead>
<tr>
<th>Score group</th>
<th>no.</th>
<th>ZGR(B)</th>
<th>PM (z)</th>
<th>CM (z)</th>
<th>IC (z)</th>
<th>PSI (z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lowest</td>
<td>P24</td>
<td>-4.32</td>
<td>1.83 (-.01)</td>
<td>3 (-.81)</td>
<td>4.83 (-1.62)</td>
<td>67 (-.66)</td>
</tr>
<tr>
<td></td>
<td>P06</td>
<td>-3.76</td>
<td>1.00(-.99)</td>
<td>4.67(.53)</td>
<td>6.67(.77)</td>
<td>61.00 (-.98))</td>
</tr>
<tr>
<td>Middle</td>
<td>P35</td>
<td>-0.76</td>
<td>1.00(-.99)</td>
<td>3.83 (-.15)</td>
<td>6.33 (.32)</td>
<td>46 (-1.80)</td>
</tr>
<tr>
<td></td>
<td>P15</td>
<td>0.59</td>
<td>1.33(-.60)</td>
<td>4 (-.01)</td>
<td>7 (1.19)</td>
<td>58 (-1.15)</td>
</tr>
<tr>
<td>Highest</td>
<td>P65</td>
<td>3.14</td>
<td>2.00(.19)</td>
<td>3 (-.81)</td>
<td>6.5 (.55)</td>
<td>72 (-.38)</td>
</tr>
<tr>
<td></td>
<td>P49</td>
<td>4.24</td>
<td>1.00(-.99)</td>
<td>4.83 (.66)</td>
<td>6.5 (.55)</td>
<td>43 (-1.97)</td>
</tr>
</tbody>
</table>
As ZGR (B) score increases, no trends can be observed in PM and CM.

The IC scores show a mild but observable increasing trend from 4.83 ($z = -1.62$) in the lowest ZGR score group to 6.50 ($z = 0.55$) in the highest ZGR score group.
Although P72 in the highest score group show a relatively higher score (72, $z = -0.38$) that is deviated from the trend, the PSI scale shows a very mild decreasing trend from 67.00 ($z = -0.66$) in the lowest ZGR score group to 43.00 ($z = -1.97$) in the highest ZGR score group.
**Co-occurrence (B)**

*Table 39.*

*The relationship between Co-occurrence (A), Parenting reflect functioning and Parenting Stress Index*

<table>
<thead>
<tr>
<th>Score group</th>
<th>Participant no.</th>
<th>Co-Occurrence (B) (z)</th>
<th>PM (z)</th>
<th>CM (z)</th>
<th>IC (z)</th>
<th>PSI (z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lowest</td>
<td>P24</td>
<td>1.67 (-1.79)</td>
<td>1.83 (-.01)</td>
<td>3.00 (.81)</td>
<td>4.83 (-1.62)</td>
<td>67 (-.66)</td>
</tr>
<tr>
<td></td>
<td>P54</td>
<td>2.58 (-1.23)</td>
<td>3.00 (1.36)</td>
<td>3.50 (-.41)</td>
<td>5.17 (-1.18)</td>
<td>133 (2.95)</td>
</tr>
<tr>
<td>Middle</td>
<td>P15</td>
<td>4.25 (-.19)</td>
<td>1.33 (-.60)</td>
<td>4.00 (-.01)</td>
<td>7.00 (1.19)</td>
<td>58 (-1.15)</td>
</tr>
<tr>
<td></td>
<td>P03</td>
<td>5.00 (.27)</td>
<td>1.33 (-.60)</td>
<td>5.33 (1.06)</td>
<td>5.83 (.32)</td>
<td>69 (-.55)</td>
</tr>
<tr>
<td>Highest</td>
<td>P55</td>
<td>6.33 (1.09)</td>
<td>1.50 (-.40)</td>
<td>3.83 (-.15)</td>
<td>5.50 (-.75)</td>
<td>96 (.93)</td>
</tr>
<tr>
<td></td>
<td>P20</td>
<td>7.33 (1.71)</td>
<td>2.17 (.39)</td>
<td>4.33 (.26)</td>
<td>6.83 (.97)</td>
<td>66 (-.71)</td>
</tr>
</tbody>
</table>
As the co-occurrence (B) score increases, the PM scores show a mixed performance and no clear trend is identified.

The CM scale shows a mild increasing trend as co-occurrence (B) score increases.

Although mixed performance in the middle and highest score group, the IC scale shows a loose increasing trend from the score of 4.83 ($z = -1.62$) and 5.17 ($z = -1.18$) in the lowest score group to the score of 5.50 ($z = -0.75$) and 6.83 ($z = 0.97$) in the highest score group.
Mixed performance of the PSI among the three groups. No noticeable trend can be identified.
Further observation and comparison of two individual participants were conducted to derive an in-depth understanding of the qualitative difference of ostensive communication measured by OCCS. Their scores on PRFQ and PSI-SF were also compared. The detail of the comparison can be found in Appendix J.
Discussion

In light of the limitations that have been exposed regarding the supportive evidence presented in the body of literature about the theory of natural pedagogy and ostensive communication, this paper has called for addressing these limitations by re-examining the phenomenon in a naturalistic setting. In order to provide a reliable and validated tool with which achieve this objective, the present study aimed at developing the ostensive communication coding system (OCCS) and establish its psychometric properties. Findings from the study provided initial support for the reliability and validity of the OCCS.

Inter-rater reliability and its implication

The initial examination suggested that all scales of OCCS are likely to demonstrate an exceptionally high level of inter-rater reliability. The result could be caused by a couple of different reasons. Implications for the future application of the OCCS are discussed in this section.

The frequency items performed exceptionally well (e.g. ICC of No. of eye contact = .97; No. of Name addressing = .98; No. of scaffolding behaviours = .87; Co-occurrence = .98). High reliabilities are achieved probably because these scales are “physically based” and employed a micro-coding method. Physically based codes require less human judgement, there is not much room for ambiguity and therefore a disagreement between coders is less likely (Ekman & Rosenberg, 2005). Micro-coding methods assess observed behaviours in discrete, overt units based on clear and explicitly behavioural operational definitions (Bell & Bell, 1989; Weiss & Dehle, 1994), therefore it is easier for coders to establish good level of inter-observer agreement (Bakeman & Gottman, 1997; L'Abate & Bagarozzi, 1993)
Although the scale of no. of scaffolding behaviours involves a lot of sub-categories of discrete behaviours (e.g. excitements, narration, regulating behaviours), it still achieved a very satisfying ICC, probably because of good amount of concrete and paradigm-specific examples (See table 3, operational definition of contingent responsivity) regarding each of the behaviours which were developed in the pilot coding. Well-developed examples tend to sensitise the coder to various behaviours and enhance the accuracy of the coding (Chorney et al., 2015).

Another factor of the coding method might also have contributed to the high ICC: in order to capture the co-occurrence, this system has employed a method that resembles the “Timed-event sequential continuous coding”. This type of coding provides information on frequency, order and also of the timing of the behaviours (Chorney et al., 2015). The coding system requires the coders to mark the exact time of the occurrence of a behaviour (see appendix K for detail). The coders also needed to actively monitor for occurrence of other ostensive cues in the next 5 to 8 seconds after an ostensive cue appeared for the first time. That prompted the coders to attend to the detail of the interaction with absolute focus. There was therefore less error or uncounted behaviour. The drawback of this methodology is that it is the most time- and resources-intensive approach (Adamson et al., 2014). In the present study, it took almost four hours to code every ten minutes of video data.

A relatively lower level of agreement can be observed in items (e.g. ICC of IDS quality = .82) which used the macro-coding method. This method requires the coder to observe an entire interaction and make one global rating. (Weiss & Heyman, 1990). That involves a more subjective judgement (Kubicek et al., 2013). Although there was some level of consensus about the prosodic contours and rhythm components of IDS, there was no objective data regarding the participant’s performance of pitch and pace based on which coders could make objective evaluation. Coders had to evaluate the performance of these
components based on their subjective judgement. There was therefore a higher variation in the IDS quality score.

An ICC of 0.82 is still considered as a very satisfying performance. That is probably the case because even though each IDS components coding has used macro-coding, they were also evaluated independently from each other. The clear breakdown of IDS and a detailed definition of these components allowed the coders to make more precise evaluation, which contributed to the high level of agreement between coders.

Another major reason that might have contributed to the high ICC is probably that the coders had high degree of familiarity of the OCCS. The primary coder was also the primary developer of the OCCS; he therefore was extremely familiar with the coding system. The second coder was also involved in the process of development, where she contributed to the refinement of the behavioural operational definitions and examples. The coders further developed their common understanding of the system through the pilot coding. They discussed each disagreement and achieved an over 70%-of items-agreement before commencing the coding process. The video data used for pilot coding was recoded with the latest manual. Therefore, naïve coders who have not been involved in such process might need to take extensive training to achieve a similar level of inter-rater reliability.

In addition, the fact that the coding process was not “double-blind” might have contributed to the high ICC. Ideally, coders should not know about the study hypotheses or be involved in administering the intervention (Chorney et al., 2015). Due to limited human resources though, both coders had been involved in administering the experiment. Their direct observation of the participants’ behaviours within and outside the coding time frame might have biased their judgement. Besides, due to the small sample size, both coders coded the same fourteen participants. These factors are likely to have boosted the inter-rater reliability.
In sum, the OCCS in the present study is likely to have achieved high ICC in all scales for a combination of reasons. Further investigation of the reliability should definitely employ a larger sample size. The researchers should be mindful of the potential bias. Therefore, should such human resources be available, the OCCS should be follow a “triple-blind approach”: three or more coders should be involved. That means the coders should not be involved with the administration of the experiments or have any interaction with the participants. Also the coders should get their video data through a random allocation; additionally the coders should be kept blind to study hypotheses.

**The dimension of rate**

In the present study, the initial exploration of the raw data suggested significant variations in the time spent on teaching, not only across tasks, but also between individuals. The time spent on Task B, for example, ranged from the shortest of 177 seconds to the longest of 892 seconds. The mean was 491.07 seconds with a relatively wide standard deviation of 190.94 seconds. This aspect of the ostensive communication has so far been rarely captured in experimental studies because the tested ostensive cues were mostly demonstrated by experimenters in a controlled manner, and therefore standardised in both frequency and rate. The effect of ostensive cues was then compared between groups in which ostensive cueing was present and those were it was absent (e.g. Egyed et al., 2013; Foursha-Stevenson et al., 2017; Nielsen, 2006; Sage & Baldwin, 2011). Clearly, the results in this section suggested that ostensive communication in the mother-infant dyad in naturalistic setting is far more complicated and dynamic, which supported the proposition of insufficient ecological validity of experimental studies.

As shown in the results, not only is there an individual variation in frequency of ostensive cueing, but there is also variation in intensity (i.e. the rate of ostensive cues in a
given period of time). Mothers who use the same or similar number of ostensive cues might distribute them very differently. Likewise, a mother who uses a smaller number of ostensive cues could present these cues in a shorter span, which results in a similar intensity in comparison with the mother who used higher number of ostensive cues.

Different rates of using ostensive cues could convey a very different quality of ostensive communication. For example, intuitively, a moderate or appropriate rate of name addressing might convey a gentle intention to communicate. Yet, an exceptionally high intensity of name addressing might convey a sense of eagerness to catch the infant’s attention and therefore intrusive. It is not uncommon to observe mothers addressing their infant by their name with a high rate when their infant is not responding or distracted. In fact, in reviewing the raw data, it is evidenced that the two participants (P55 and P39, see table below) who used the highest number of names addressing in Task B, have also shown the highest rate. At the same time their toddlers showed the least attention among all participants. At the same time, the two participants (P15 and P65) who used the lowest numbers of name addressing have also shown the lowest rate and highest attention among all participants.

Table 40.
Comparison of participants with lowest and highest no. of name addressing

<table>
<thead>
<tr>
<th>Participant no.</th>
<th>Attention (z)</th>
<th>No. of Name addressing (z)</th>
<th>Rate of Name addressing (z)</th>
<th>Time (z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>P15</td>
<td>2 (.96)</td>
<td>2 (-1.00)</td>
<td>0.005 (-1.31)</td>
<td>420 (-.37)</td>
</tr>
<tr>
<td>P65</td>
<td>2 (.96)</td>
<td>3 (-0.90)</td>
<td>0.009 (-1.0)</td>
<td>346 (-.76)</td>
</tr>
<tr>
<td>P39</td>
<td>0.75 (-2.3)</td>
<td>26 (1.50)</td>
<td>0.051 (2.36)</td>
<td>505 (.07)</td>
</tr>
<tr>
<td>P55</td>
<td>1 (-1.60)</td>
<td>31 (2.03)</td>
<td>0.035 (1.05)</td>
<td>892 (2.1)</td>
</tr>
</tbody>
</table>
Of course, this result might reflect patterns which only surface in extreme cases, and the relationship between the frequency and rate of ostensive cueing remains unclear. But this result provides support to the notion that the rate of ostensive cueing is an independent dimension of ostensive communication. The cause and effect of different rate of ostensive cueing should be investigated in future studies.

For the purpose of standardisation and computation of the rate items, time spent on a given teaching episode was precisely defined (See table 7 in the method section). Any behaviour that happened outside this specific timeframe was not captured by OCCS, but this behaviour might well be contributing to the overall ostensive communication. In the target object teaching episode of Task B, the coding of ostensive cues stopped after the mother’s first trial to test what the toddler has learned. Some mothers continued using ostensive cues in the testing phase, especially when the toddler lost interest in the task, became exhausted or distracted. P06, for example, after one of her teaching episodes ended at 21 minutes 20 seconds, she started testing her toddler by asking her to show the target object. Therefore, any behaviour after that was not coded. Her toddler was not cooperating with the testing because she was upset and wanted to see her father who was in the waiting area. The mother addressed her toddler by name, used eye contact and other cues trying to reassure her, which lasted for quite a few minutes. Although these cues might not be directly relevant to the teaching, they reflected the mother’s capacity to regulate the toddler’s emotions. A stable emotional state would have positive influence on her recalling of newly learned information later (Rose et al., 1999). If this period had been defined as a part of the teaching episode, the time spent on this episode would have been extended, which could change the overall result regarding the rate of ostensive cueing. Therefore, the definition of a teaching episode should be further considered in future studies when investigating the rate of ostensive cueing.
Co-occurrence and the relationships among the three dimensions

In the examination of hypothesis 2, the result acquired from detail observation of individual patterns of scores provided evidence that supported the hypothesis. Participants who used higher frequency of ostensive cues tended to have higher co-occurrence score. This is consistent with the literature that mothers tend to use these cues together within the window of five to eight seconds (Goldstein & Schwade, 2008; Goldstein et al., 2010; Yu & Smith, 2012). In addition, the result also suggested that, there could be strong positive relationships among frequency items, global items and co-occurrence. This means participants who used higher frequency of ostensive cues also tended to demonstrate better quality in IDS, contingent behaviours and emotional support, and again these ostensive cues tended to present together in a coordinated manner. This result provides support to the proposition that ostensive cues come as a package rather than in isolation in naturalistic setting. There were no observable patterns of individual scores among the rate items and the rest of the dimensions. This is a bit surprising. One might think mothers who show higher rate in ostensive cueing would likely to have high co-occurrence. This result again called for future investigation on rate of ostensive cueing. Nevertheless, the result suggested that the subscale of co-occurrence can reliably capture the variation in degree of coordination among ostensive cues.

Differential validity

The present study aimed to design the OCCS as an instrument that can be used for evaluation of ostensive communication universally in any natural pedagogy situation and experimental paradigm, although the original proposal of validating this aspect of OCCS by comparing the score difference between the SNOC group and the ROC group could not be
implemented. The detail observation of score difference between a participant’s performance in Task A and Task B provided supportive evidence. In the comparison of the ostensive communication performance scores (i.e. ZF, ZR, ZGR and co-occurrence), we could observe that there is a likely to be a strong positive association between the performance of ostensive communication in terms of frequency (ZF), quality (ZGR) and co-occurrence in both Tasks. Which means a mother who has an above-average performance of ostensive communication in Task A tends to also have an above-average performance in Task B. It seemed to suggest that the capacity to use ostensive cues in these aspects is a relatively stable quality for an individual. This is consistent with the proposal that there is likely to be a stable underlying psychological construct(s) that is governing the tendency to use ostensive cues. Despite this stable tendency, there are also distinctive differences found in the score of these dimensions in Task A and Task B for the participants. That means a mother might tend to have an above-average performance of ostensive cueing in general, but in some tasks, she is significantly better than her counterparts, in other tasks she is only mildly better. Similar patterns are observed in below-average performance. This seems to suggest that a considerable amount of the variance of ostensive cueing is influenced by the type of teaching and the type of knowledge a mother is trying to pass on.

This result indicated that not only the OCCS can reliably distinguish individual’s level of capacity of ostensive communication, it further confirms that the OCCS can also reliably detect the subtle changes of ostensive cueing from one paradigm to the other. This result gave some level of confidence that the OCCS can be used to evaluated ostensive communication in a range of pedagogical situations. In addition, it might also suggest that it is not sufficient to give a full account on an individual’s capacity of ostensive communication if the person is only evaluated in one type of teaching. The individual needs to be observed in a range of natural pedagogical situations.
Construct validity

Reflective functioning and Contingent responsivity (CR)

The present study examined the patterns of scores among the constructs of mentalisation and components of CR (i.e. scaffolding behaviours, contingent behaviours and emotional support). Findings were loosely consistent with hypotheses; results were mostly in the expected directions.

1. Pre-mentalisation state (PM)

Although no significant association was found in the spearman correlation analysis, the patterns of scores found between Pre-mentalisation state (PM) and the performance of CR showed a loose support to the hypothesis. For example, the one participant (P54) who has a significantly higher score of PM (3, $z = 1.36$) showed considerable poorer performance in most of the components of CR. In contrast, participants who have a PM score lower than or similar to the normative mean showed more above-average performances. In prior study, PM, in particularly relevant to a tendency to make malevolent attribution to child’s thought and behaviours, was found highly negatively associated with secure attachment and all dimensions of emotional availability (Fonagy et al., 2014; Fonagy & Luyten, 2016; Luyten et al., 2017). The component of emotional support in CR tapped not only into the warmth and supportiveness, but also the how empathic the mother is. It is therefore unsurprisingly, in the present study, mothers with higher score of PM was found lower in emotional support in both tasks (see figure 11b). The result was only loosely consistent with the hypothesis probably because most of the participants in this sample have a generally lower score on PM than the norm. The sample mean is 1.61 ($SD = .57$) is considerably lower than the normative mean is (1.84 $SD = .85$). The sample might have insufficient representation of the higher PM group in the population.
2. **Certainty of Mental state (CM)**

The patterns of score found between Certainty of Mental state (CM) and the performance of CR showed some support to the hypothesis. Relative to mother with securely attached infants, mothers with higher level of CM tended to have infants with anxious-resistant attachment pattern, whereas mothers with relative lower level of CM tended to have anxious-avoidant infants (Luyten et al., 2017). Excessive (hyper-) and insufficient (hypo-) levels of parental contingency was found to be positively associated with nonoptimal child development, but moderate level of parental contingency found positively associated with secure attachment relationship (Bornstein & Manian, 2013; Hane et al., 2003; Isabella & Belsky, 1991; Jaffe et al., 2001; Jay et al., 1984). This is consistent with the result in present study that participants whose CM scores fell on the middle showed more above-average performance on CR, whereas the participants with an extreme score of CM in the sample tended to show below-average performance in all components of CR. Optimal level of scaffolding and good quality of contingent behaviours involve certainly level of “knowing” the infant’s mind relatively well but at the same time recognizing the opacity of infants’ mental state. (Luyten et al., 2017; Smaling et al., 2016). Individual with relatively low and high CM might have undermined their reflective function, hence, reduce their capacity to produce optimal scaffolding and contingent behaviours. It is also obvious that participants with optimal level of CM tended to have higher score in emotional support (see table 14a &14b).

3. **Curiosity and Interest of Mental state (IC)**

The patterns of scores found between Curiosity and Interest of Mental state (IC) and the performance of CR were only partly consistent with the hypothesis. The Participant with
the lowest IC scores showed consistent below-average performance across all component of CR in both tasks. There were some observable improvements of the performance of CR in Task B as the IC score increased. This is consistent with the previous findings that mothers with good level of IC tended to have infants with secure attachment and social cognitive development (Luyten et al., 2017). Higher IC usually means the mother has better capacity in recognizing the opacity of the child’s internal subjective world and showing a genuine interest and curiosity, in turn, the mother tend to use more scaffolding and contingent behaviours to support the infant’s learning but using less directive or intrusive behaviours.

P39’s performance was inconsistent with the hypothesis. She had relatively higher score on IC, yet all her performance of CR was below-average. In fact, her performance of scaffolding in Task A, and quality of contingent behaviours in Task B was one of the lowest among the participants. In reviewing video data, P39 seemed to be inpatient in both tasks, there was minimal sense of warmth. There was very minimal variation in how she taught. She did not demonstrate much attempt to mentalise, she did not come across as someone who has high interest and curiosity of her toddler’s mental state. P39 might be considered an outlier as she has low score on PM and optimal level of CM, yet her performance in most CR components has not been consistent with other participants who have similar portfolio in PRFQ.

In contrast, P20, who had similar score of IC as P39, showed result that’s consistent with the hypothesis. She had very good performance in scaffolding behaviours across both tasks, and significantly better performance on all components of CR in Task B than P39.

P15 had the maximum score of IC (7, z = 1.19). It is proposed that extreme high score in IC might reflect intrusive hypermentalising, which might result in poor performance of CR. Inconsistent with this proposal, P15 showed above-average performance on almost all CR components across both tasks. Given the small sample size, there was only one
participant with extreme score of IC, it might not be representative to the population with intrusive hypermentalisation.

In sum, despite a small portion of results that were not fully consistent with the hypotheses, the overall patterns of scores between reflective functioning and contingent responsive seemed to provide initial support to the construct validity of the CR scale, where the individual who has better reflective functioning tend to have better performance on contingent responsivity.

**Further exploration of reflective functioning and ostensive communication in general**

Further exploration on the patterns of scores between reflective functioning (RF) and different dimensions of ostensive communication was conducted.

As can be seen in Table 24 (the summary of observable patterns of scores among OCCS dimensions, PRFQ and PSI-SF), there are different observable patterns of scores between OCCS and PRFQ in Task A and Task B. For example, RF seemed to play an important role in frequency items (ZF) in Task A, but not as much in Task B. On the contrary, RF seemed to have a stronger influence on co-occurrence in Task B than in Task A. The level of PM and IC showed a similar pattern of scores with global rating items (i.e. quality of IDS, quality of contingent behaviours and emotional support) in both tasks. This seems to suggest that although the RF plays a role in the mother’s use of ostensive cues, the role shifts depending on situation.

In Task A, for example, as there was less opportunity for trial and error, the mother might need to quickly figure out whether the toddler has learned the novel action within the small windows in between demonstrations. High in RF might allow the mother to actively mentalise the toddler’s cognitive state in facing the challenge of the learning task and
therefore use more ostensive cue to facilitate the learning. While in Task B, there was sufficient room for the mother to come to a similar level of using ostensive cues through trial and error as long as the mother has a reasonable level of RF, but the higher in RF made a difference in how the mother was coordinating these cues to make them co-occur.

Pre-mentalising state (PM) seems to have consistent negative relationship with ZGR items in both tasks, at the same time, Curiosity and Interest of mental state (IC) seems to have consistent positive relationship. Considering the quality of contingent behaviours and emotional support are the main components contributing to the ZGR score, this is consistent with the result in construct validity analysis of the contingent responsivity scale. The other component of ZGR is IDS quality. With the current level of analysis, it is unsure whether PM and IC also impact on the mother’s use of IDS, or which component of IDS (i.e. Prosodic contour, Rhythm, linguistic structure and Repetition). This should be further investigated.

Among all the dimensions in both tasks, certainty of mental state (CM) only showed potential positive relationship with frequency items in Task A. This might be because there was no extreme high score of CM in the sample to demonstrate potential non-linear relationship between CM and different dimensions of OCCS.

In sum, although with the current level of analysis these results, the relationship among RF and eye contact, name addressing and IDS remain unclear, these results did seem to suggest that RF plays a role in overall quality in ostensive communication. Further investigation is needed to explore this notion.

**Parenting stress (PSI-SF) and ostensive communication**

The investigation of patterns of scores between PSI-SF and OCCS showed evidence that partly support the hypothesis and contribute to the validation of the construct validity of OCCS.
Low in total parenting stress seemed to consistently relate to higher in quality of IDS, contingent behaviours and emotional support, which provided some support to the construct validity. This is also consistent with the literature that low parenting stress, the use of IDS, contingent responses, emotional support and children’s developmental outcome and language development are related to each other. (Anthony et al., 2005; Campbell et al., 1991; Crnic et al., 2005; Crnic & Greenberg, 1990; Ma et al., 2011; Magill-Evans & Harrison, 2001; Moss et al., 1998; Pianta & Egeland, 1990; Spinelli et al., 2017). It seemed to suggest that ostensive communication takes a role as a mediator in this process, which is to be verified in future study.

Interestingly enough, parenting stress showed observable patterns of scores with OCCS mainly in Task A, but only with the ZGR dimension in Task B. Similar to the finding in the section above that, it seems a considerable amount of variance of ostensive cueing is influenced by the type of teaching. The effect of psychological factors on ostensive cueing seems to change in different situations.

Moreover, with the current level of analysis, it is unclear which component (i.e. Parental stress, Difficult child, Parent-Child dysfunction interaction) of parenting stress plays what role in facilitating or prohibiting the use of ostensive cues. Further detailed exploration is recommended in future study.

Applying OCCS in two learning paradigms and its implications

Task A and Task B were very different in many ways and likely to have brought difference in the mother’s use of ostensive cues. For example, there was in general less interaction between the mother and the toddler in Task A than in Task B simply because the mother was only allowed to have three demonstrations in Task A, there was not much window to get “feedback” from the toddler. The mother did not know if the toddler has
learned or not until the very end of the experiment. While there were all together nine
teaching episodes (nine objects to teach) in Task B. Some mothers seemed to show a gradual
change in ostensive cueing as each teaching and testing was an opportunity of trial and error.
There was more opportunity for the mother to figure out what works and what doesn’t.

On the other hand, participants spent significantly longer time in Task B than Task A.
Considering the very limited cognitive resources for toddlers at this developmental stage,
Task B was a more demanding task than Task A. No toddler showed exhaustion in Task A,
but quite a few of them found it challenging to continue halfway through Task B. However,
these toddlers also showed similar level of attention in both tasks. This seem to be consistent
with the observation that mother needed to use more ostensive cues in Task B to maintain the
toddler’s state of mind needed for the learning to take place. Therefore, naturally, participants
were shown to have more eye contact, name addressing and scaffolding behaviours. They
were also shown to have better quality in IDS, contingent behaviours. The emotional support
score in Task B was more than double of Task A, this might be because the mother needed to
use more supportive behaviours, which conveyed a stronger sense of warmth, to help the
toddler to regulate their feelings when they were tired, distracted and grumpy. The
considerable higher score of co-occurrence in Task B seems to reflect that there was an
increasing need to coordinate these cues. Interesting enough, there was higher density (rate)
of these cues in Task A despite lower in frequency. The rate of eye contact of Task A is
almost triple of Task B, the rate of name addressing is double.

As discussed earlier, although the performance of ostensive communication seems to
be a fairly stable quality that that is governing by underlying psychological functioning (e.g.
RF parenting stress), there also seems to have significant variation of the use of ostensive
cues subject to the infant’s immediate state (e.g. emotional state, physical state) in relation to
pedagogical situation, or the task itself (Laosa, 1980). The dynamic relationships among
these variables should be further investigated in future study once the OCCS is fully validated.

**Clinical implication**

In previous studies, assessment of maternal contingent responsivity often naturally linked to direct or indirect intervention programs for infants and their mothers (e.g. Responsivity Education/ Prelinguistic Milieu Teaching (RE/PMT) (Yoder and Warren 1998; 2001;2002); Playing and Learning strategies (PALS) program (Landry et al., 2006); and Mediational Intervention for Sensitising Caregivers (MISC), Klein et al., 1987)) because the development of the tool usually involved investigation of the underlying mechanism of the behaviours concerned, as well as the establishment of very detail operational definition of these behaviours. The operational definition not only provided helpful guidance as to what behaviours to look out for in the assessment, but also allowed clinician to know precisely where to intervene.

The development of OCCS in present study shares the same feature. Once the OCCS is fully validated in future study, it’s application can potentially be extended into other clinical and research settings. For example, it can be used as an assessment tool in mother-infant mental health clinic for screening individuals who might have difficulties in using ostensive communication due to other mental health issues (e.g. postnatal depression). On a service level, the OCCS is helpful in building up evidence base for the case of setting-up particular services for certain clinical population. On a clinical level, the OCCS provides means to assess each component of ostensive communication and therefore allows clinician to formulate and operationalise the interventions that’s customised to each individual. After the intervention is completed, OCCS can also be used as in post-intervention assessment and evaluate the effectiveness of the intervention. Of course, OCCS can also be used to evaluate other existing interventions and might provide new insights as to how to refine these
programs. In a more macro perspective, future researches can also look into the relationship between individual performance in OCCS and demographic characteristic (e.g. level of household income, working hours, SES), outcomes of these studies can potentially inform the development of family friendly office policies and social benefit policies.
The Case study

By investigating individual cases in an in-depth and qualitative approach, the case study aimed to verify if the OCCS scores are correspondent to the quality of ostensive communication in the experiment.

The detail description of the performance of P24 and P49 (see Appendix J) thoroughly reflected their distinctive difference in ostensive communication between the two mothers. P49 not only used more ostensive cues, she used them flexibly with plenty of variation. She showed an advanced capacity to tune-in with her toddler and adjusted her use of cues subject to her daughter’s responses. There was a clear demonstration of high reflective functioning and contingent responsivity. Meanwhile, P24 used a smaller number and lower quality of ostensive cues in general, she also used them rigidly and mechanically. There was little evidence of her attempting to tune-in with her daughter and adjust her strategies accordingly. P24 also did not seem to enjoy the task with her daughter as much as P49. This seemed to be consistent with how she approached the task as something to be accomplished rather than something fun to do with her daughter. Their performances are accurately reflected on their score in OCCS, which gives support to the validity of the scale.

P24 and P49’s performance in PRFQ and PSI-SF were compared. As indicated by the result, P24 showed poorer reflective functioning in all three dimensions of PRFQ than P49. These results are also consistent with their performance in the experiment where P49 showed way better reflective function than P24. Unsurprisingly, P24 also showed higher level of parenting stress than P49. The cases study demonstrated the potential relationship among the constructs of PRFQ, PSI-SF and ostensive communication, which gives further support to the construct validity of OCCS, although further statistical investigation with a larger sample size is needed to verify these findings.
These results seem to suggest that the OCCS can reliably reflect the differences in ostensive communication in different teaching and learning situations.

**Limitation and further development of OCCS**

**Measuring eye contact and name addressing**

Although the scale of no. of eye contact and name addressing are likely have achieved a good level of ICC, the accuracy of capturing these behaviours can still be enhanced. Collecting data of simple physically based behaviours as such could for example be computer automated (Chorney et al., 2015) which could at the same time reduce the demand on the coder and avoid unnecessary human error.

For example, Ye et al. (2012) developed a system for detecting eye contact between adult and child through having the adult wear eye-tracking glasses. They used the gaze track technology to determine the adult’s point of gaze, then combined it with a computer vision analysis of the child’s face to determine her/ his gaze direction, thereby identifying an exchange of looks by the dyad. With the help of this technology, the system can potentially have more accurate accounts of joint attention as well. Name addressing can also easily captured by computer-based speech recognition program like such as Python (Manaswi, 2018).

Unlike the scale of IDS or emotional support, which is captured and measured in marco-coding with attributions such as “good performance” or “poor performance”, eye contact and name addressing are measured in frequency in the current version of OCCS. One of the limitations of micro-coding is that it might not capture the “meaning” of the behaviours. Take the example of rate of name addressing as an example. Capturing only the frequency of name addressing might not be able to reflect whether this ostensive cue was
used in a helpful or intrusive manner. Excessive eye contact could also be unhelpful or convey an intimidating impression (Csibra, 2010).

Therefore, eye contact and name address should also be measured under a macro-coding approach to capture the “appropriateness” of the behaviour. Combination of macro-coding and micro-coding could potential capitalise the strength

**Measuring and validating IDS**

Similarly, although the instrument has achieved satisfying inter-rater reliability with the IDS sub-scale, to objectively evaluate the quality of IDS, future research could consider employing the electrolaryngographic measure (e.g. Grieser & Kuhl, 1988) that produces actual data of the prosodic contours and rhythm for the researcher.

Besides, in the present study, the construct validity of the subscale of IDS quality has not been tested independently from other subscales yet. Future studies should address this limitation. Various studies have reported that depressed mothers tend to produce IDS with significantly lower fundamental pitch (Kaplan et al., 2001) and less effect on infant’s learning (Kaplan et al., 1999). Kaplan et al. (2011) confirmed this prediction and reported evidence that a mother’s depression score as measured by Beck Depression Inventory (BDI-II) (Beck et al., 1996) is positively associated to IDS with lower foundation pitch. Therefore, it is recommended that future research can use BDI-II as a measure for validity testing the IDS quality subscale in OCCS.

**Measuring contingent responsivity**

One of the essential criteria to decide whether a behaviour can be regarded as a contingent response is how timely the mother is responding to the infant’s signal. In the
present version of OCCS, there is no concrete assessment of this aspect. Although Bormstein and Manian (2013) measured contingent responsivity with a relatively narrow behavioural operational definition, the distinctive merit of their study is that they used micro-code to capture the maternal response based on the appearance of the infant’s initial signal. These signals included (1) looking at mother, (2) looking at object, (3) non-distress vocalisation, and (4) smile. In order to thoroughly evaluate contingent responsivity, the next version of OCCS should consider employing this approach. It is recommended that there should be another scale to capture the exact timing of the infant’s signal. When enough human resources are available, this scale should be administered by an independent coder. The timing of the maternal responses should be compared against the timing of the infant’s signal to determine the contingent responsivity.

Bormstin and Manian (2013) reported another insightful finding. They compared the maternal sensitivity (measured by Emotional Availability Scale (EA Scales; Biringen, Robinson, & Emde, 1998) with the frequency of contingent responses and reported that under- and over-contingency are both judged as less sensitive. In other words: There is an optimal level or range of contingent responsivity. No formal evaluation of optimality in the current version of OCCS is a major limitation to be addressed in future research. Last but not least, Bornstin and Manina (2013) has also indicated that the Emotional Availability scale can be a useful measure for further validation of the subscale of contingent responsivity.

Conclusion

The present study aimed to develop and establish the psychometric property the ostensive communication coding system (OCCS). However, due to a disruption in data collection, the study did not have sufficient statistical power to undertake inferential examination. The study therefore employed a quasi-qualitative approach to make a detailed
observation of the patterns of scores among different dimensions of OCCS, PRFQ and PSI-SF. Overall, the initial examination of the instrument provided some support to the reliability and validity of the instrument. The findings should be further verified with a larger sample size in future research.

Besides, the findings of the study also gave inspiration for further investigation. For example, a potentially dynamic relationship among psychological functioning (e.g. RF, parental stress, maternal depression), ostensive communication and infant’s immediate state and learning paradigm is yet to be explored. The cause and effect of a different rate of ostensive cueing is a novel area that has not been explored in previous research. And refinement of the subscales of OCCS could be achieved by employing new tools (e.g. Python, wearable eye-tracking glassless) and additional methodology.
References


Adams, K. L. (2006). Parental stress, parenting behavior and observed parent-child interaction. Pace University,


Reflective Functioning Questionnaire. *PLoS ONE, 11*(7), e0158678. doi:10.1371/journal.pone.0158678


Part 3: Critical Appraisal
Introduction

I will start by outlining the influence that my prior experience and theoretical orientation had on my approach to this research. I will then reflect on the challenges encountered during the course of the study. It will be followed by a summary of the experiences and lessons learned. Learning how we managed these challenges hopefully will provide helpful guidance to researchers who are interested in studying this field. The last part of the reflection will focus on the impact of global pandemic and Covid 19 to this particular area of research and conclude with proposed ideas of adaptations for fellow researchers to explore.

Influence of prior experience and theoretical orientation on my approach

My theoretical orientation is heavily influenced by psychodynamic and psychoanalytic schools. I started my postgraduate academic and clinical training in the Anna Freud Centre around 10 years ago. I completed the MSc in Psychoanalytic Developmental Psychology, which contributed to establishing the frame of reference through which I understand the world and my personal experiences; from macro phenomena, like human culture, civilization and politics, to the micro level, as it manifests itself in every day human relationships, interactions and also the clinical practice. The robust evidence in the literature has convinced me that early experiences between mother and infant are absolutely crucial for individual development. From then on, I have been very interested in the investigation of mother-infant interaction.

Although my clinical practical work has been mainly with adults, those experiences have only deepened my interest in these schools of thought as they have given me plenty of opportunity to see the link between a person’s early development and their current psychological state. What I have also learned from the consultation room is that human interactions are extremely complex, and it can be very challenging to observe and study all
the subtleties of these interactions without a sound theoretical framework and suitable instruments. Years of academic training have led me to appreciate the value of scientific research and statistical analysis. Yet, it is not uncommon for quantitative researches to fall short when it comes to explaining a complex phenomenon like mother-infant interaction because of the difficulty in capturing the data necessary for the analysis.

I joined a research team during the MSc. In exchange of having access to the data for my dissertation at the time, I was involved in the study of Object Relation and Inhibitory Control (ORIC) in Early and Late Adolescence. For the first time I took part in transforming qualitative data captured by the Thematic Apperception Test (TAT) into quantitative data through coding, following the manual of the Social Cognition and Object Relations Scale. I was fascinated how this methodology succeeded in capturing the richness of human experience and turns it into quantitative data that makes statistical investigation possible. The choice to take part in the present study was motivated by my interest in mother-infant dyads and my interest in bridging the gap between quantitative research and psychological constructs.

**Major challenge – recruitment**

Our team encountered major challenges during this study in the phase of recruitment and testing, which were confronted creatively and thoroughly. These experiences are worth reflecting upon, for the future benefit of my own development as well as for this project.

The targeted population of the present study were mothers with toddler aged 16-22 months. Most of the mothers at this stage had finished their maternity leave and returned to work either full time or part time. They therefore have limited time to take part in community activities, such as playgroups, under 5’s class and the Baby Bounce in the local library. This
characteristic of this population made them difficult to reach. This was evidenced in the pilot study.

According to the researcher who conducted the pilot study, the research team has only been able to recruit 15 pairs of mother and toddler over the course of a few months, despite intensive efforts. We encountered the same difficulties. In the first two months, the research team visited around ten libraries where the Baby Bounce and under 5’s class take place. We also visited around ten activity groups. There were roughly ten to fifteen mothers who took part in each of these activities. All together we met around two to three hundred mother and infant pairs. There were usually a few mothers who showed interest, but only one or two agreed to leave contact details. The other mothers were either unavailable for the experiment, or her infant was out of the age range. The response rate was low. We managed to recruit less than 20 pairs in the first two months. The data of demographic characteristics indicated that over 50% of our sample have a yearly household income over £50 000, suggesting that most of these mother are from middle class background. This is probably one reason why a financial incentive of £20 pounds did not hold much appeal. We struggled to find a better way to advertise the study and boost the respond rate.

I did a very small scale survey within my social network trying to identify what incentive might tap into these mothers’ interest. Apparently, we had not been mentalising our target audience enough and overlooked what is most obvious – a mother’s primary interest is her infant and the experience of spending time together. This helpful feedback led me to consider different marketing strategies. I realised that it is important for the mother to feel she is doing something fun and meaningful with her infant, and that it is also crucial to have something tangible from that experience, helping them to retain this memory. Another helpful aspect would be that mother and infant can make daily use of this object. While it could make them feel proud about having taken part in the study, the object at the same time could draw
interest of other potential participants. Giving out a T-shirt with an interesting print might meet all these objectives. A custom printed T-shirt would have cost over £10 each. Unfortunately, the project did not have a big enough budget to implement this idea. A custom printed baby bib turned out to be a more suitable idea.

I then designed a baby bib (Appendix L) that attempted to meet the listed objectives. The central theme was to make the mother feel proud about her infant’s involvement in the study. I therefore chose the slogan “A born Baby Psychologist”. To recognise their contribution to the scientific community, I added the logo of the Anna Freud Centre and UCL. The cartoons of two playing and thinking infants would allow the mother to identify with her infant. On the poster, we stressed on the fact that taking part in the study involves some fun activities and is therefore a relaxed and interesting experience to have.

A wordy poster was not helpful in catching mother’s attention, as probably they are very preoccupied with their infants and had no extra cognitive recourses to entertain lengthy information. A picture of the baby bib is a better eye catcher than a poster that is full of words. The initial design received very positive feedback from my social circle and was sent for mass production.

In the meantime, the research team tried to find better ways to reach the target mothers. The conventional way of recruitment through visits to mother-infant community events had proven to be ineffective. Since during the 3rd year of the doctorate in clinical psychology training, each trainee has only got roughly two research days a week, so logistically, we could visit one or a maximum two events per day. Therefore this approach was not cost-effective. In order to maximise the number of potential participants, we used social media. We spread the information through facebook and twitter within our network and spoke to our contacts individually to encourage them to further spread the information by re-posting the poster on their social media, while asking people to do the same in their
network. We also approached some key opinion leaders on Instagram, such as weanin15_@instagram. This account is operated by two parents who have over 465k followers on Instagram. After we explained our research, they kindly agreed to post our poster and information in their newsfeed.

These promotion strategies have proven to be a great success. The respond rate was significantly boosted two- or threefold. We successfully recruited over 65 participants over the course of two to three months. On reflection, recruitment is a very important step for any research project - more so, when the targeted population is difficult to reach. But it might generally not receive the investment in time and creativity it deserves. Therefore I believe sharing this experience could be helpful to fellow researchers who are going to conduct studies within a similar population.

**Major challenge - testing**

As we successfully accomplished the task of recruitment, we moved onto the next stage of the study: the testing. It was meant to be more straightforward, yet, like many other research projects, limited resources on multiple fronts made it more challenging than expected. To give the reader a bit of context: The experiments had to take place in the Anna Freud Centre; in order to capture the behaviour data in the mother-infant interaction, the laboratory needed to be set up in a precise way. There were four cameras attached on each side of the room on the walls and one laptop to control the cameras. Only one room in the Anna Freud Centre had such equipment installed; this room was frequently used by clinicians therefore its availability was very limited in the first place. On the other hand, as mentioned above, the population we researched is characterised by its very limited availability; not only because the mothers have usually returned to work, but also because there are only a couple
of hours a day in which the infants are active and in the right state to take part in the experiments. Together with the fact that the researchers in our team could only work three out of the five working days, the likelihood of having all parties (i.e. participants, researchers and facilities) available during the same time slot was small. The lack of certainty about the availability of the participants and the testing laboratory created major difficulties in coordination. For example, often we were not able to get confirmation from the mother to schedule an experiment. The team had to send a mass email to all participants with all available slots for them to consider. Yet often, when one participant finally found a slot that would work for her and the infant and replied by email, the slot had already been reserved for another participant. The back-and-forthing to find a workable time slot for the experiment through email was very time consuming and inefficient.

The team tried to confront this challenge on multiple fronts. Firstly, we tried to increase the capacities for testing by exploring the possibility of conducting the experiment at non-office hours or on Saturdays, in cooperation with the property management team. Secondly, the team tried to work with the finance and IT department to install the necessary equipment in an additional room. Unfortunately, these proposed solutions could not be realised for financial reasons. Subsequently, in order find a more efficient scheduling system for the experiment, the team consulted the IT department to design a dynamic booking system which would allow all participants to see all the available slot. The participant can log into the system with her participant number and choose a slot which works for her. To avoid double booking, once a slot is taken, it will be removed from the list and is no longer available to the other participants. The system operates very much like the system we use to select a seat on a flight. It might look like a simple step, but it can save tremendous amount of time for the researchers. This simple system is not only applicable for this study, but for all projects which require participants and researchers to come together for an appointment.
Although this project had to be paused as the global pandemic lead to a suspension of all experiment, it is a system that is worth investing in for the benefit of many research projects in the institute.

Taking part in this project has been an invaluable learning experience. For the first time I have been fully involved in every stage of a project; there was a lot of room to make autonomous decisions and shape the everyday operation of the study. I have particularly appreciated the opportunity to solve problems by using creativity and technology (e.g. social media, design software). On reflection, while the team has put a lot of thought and effort in the research proposal, especially to the scientific element including literature review, study design and statistical analysis, we may not have given sufficient consideration to the everyday logistics and operative aspect of the study. We could have developed a more comprehensive plan from the start by consulting researchers who have conducted studies with a similar population and setting, thereby learning from their experiences. It might be helpful to recognise that such valuable experiences often get lost in the process. We as researchers tend to, quite understandably, put our primary focus on the content and the outcome of the study, not its logistical foundation. The critical appraisal has provided the space that was necessary to reflect and pass on these experiences.

Adaption to COVID 19

The global pandemic of COVID 19 has had a profound impact on almost all industries; the field of academic study being no exception. It was most unfortunate that after resolving the challenges that came along in each stage, the scheduled experiments had to be called off. While making the necessary adjustment to maximise the utility of the data we collected, the big challenge ahead is the continuous influence of COVID 19. Due to the
reduction of confirmed infections, the lockdown of the country is being lifted progressively as I am writing this critical appraisal. There is the hope that life comes back to normal and various research projects can pick up their momentum. However, in the meantime, second waves of major outbreaks of COVID-19 have been reported in multiple countries and regions, including the United States, Australia, Hong Kong and Japan. Many leaders have called for a long-term management plan to prepare for maintaining the operation of their respective industries under the continuous threat of the virus. In the academic field, we might have to acknowledge the fact that, for the foreseeable near future, the impossibility of conducting experiments in the laboratory will become the “new normal”. In light of this circumstance, I would like to take this opportunity to discuss and explore the potential adaption to this new situation in the context of this particular project.

As reported in the empirical paper, the ultimate objective of this research project is to study the phenomenon of ostensive communication between mother and infant in real life. The present study has tested the feasibility of observing the ostensive communication through having the mother-infant dyad engage in a learning task in a laboratory setting. The interaction was video recorded and coded under the newly developed OCSS and eventually transformed into quantitative data for statistical analysis. One of the limitations of the study was that the pedagogical process might not fully correspond to a natural interaction, because the experiment did not take place in a natural environment. Additionally, the presence of the experimenters might have inducted the mother’s performance anxiety. Not only this limitation could potentially be addressed if the experiment took place in the participant’s home; a home-based experimental paradigm could be a necessary adaptation for the project, enabling it to be able to be continued in the context of the current pandemic. Mahoney et al. (1986) videotape-recorded mother and infant playing together in their living room and successfully produced data for analysis. But it might not be possible for experimenters to do
home visits due to the guidelines of social distancing. Therefore, one possible step forward is to design a learning task that can be administered and manipulated remotely.

The first step is to recruit participants who are interested to take part, and screen for those who have the necessary equipment (e.g. a laptop, tablet, webcam, wireless earphone, stable internet) and an appropriate environment (e.g. a spacious enough living room). The paradigm could be a partial replication of the current design but replace word-object learning by word-picture learning. Or it could be a computer-based learning task that requires the mother’s involvement. Experiment materials could be posted or emailed to the participants. The mother could watch the training or instruction video or read a written manual to familiarise herself with the procedures. The researcher could do some practice on the paradigm with the mother through video conference to make sure she fully understands how to administer the learning task. In the actual experiment, the interaction in the teaching and learning could be live-streamed through video conference. The researchers could record the video from their end, while the participant records the process from another angle with their mobile phone. During the testing, the experimenter could give instant instructions or assistance through the wireless earphone.

With these adaptions, the mother and infant do not need to make a visit to the laboratory in person and therefore would have more flexibility to choose when to take part in the study. This is likely to boost the response rate as limited availability was the main obstacle for recruitment. Of course, there might still be potential challenges which require the researcher’s further consideration. For example, data protection is an issue, and how to make sure the video data is sent through a secured network. When participants will be videoing in their home environment, confidentiality can be a concern. Researchers will also have much less control over the potential confounding variables brought by the unstandardized setting of
the home environment. More detailed planning and piloting can be helpful to address these issues.

Conclusion

In conclusion, taking part in this project has been an invaluable learning experience as well as a process of integration. I have come to appreciate that running a research project is indeed like making a journey. Endless expected and unexpected challenges arise along the way. It does not only take an intellectual mind to resolve different obstacles; one also needs to be thinking out of the box, to be creatively flexible and adaptive. The outcome of the research is not only a product of my academic endeavour, but an integration of other skills and capacities (e.g. use of technology, marketing strategies, designing, working with professionals who have expertise from another field). In this appraisal, I have reflected upon the experiences in managing challenges that came with recruitment and testing and have added an initial exploration of making necessary adaptations due to the COVID 19 pandemic. I hope sharing these experiences and thoughts can be helpful for fellow researchers who are interested in continuing their exploration in the area of mother and infant ostensive communication.
Reference


Adams, K. L. (2006). Parental stress, parenting behavior and observed parent-child interaction. Pace University,


doi:10.1371/journal.pone.0158678


doi:10.1017/S0305000900002464


doi:10.1111/j.1467-9280.2009.02295.x

doi:10.1006/jecp.1998.2438


Phillips, R. S. C. (1994). *Infant-directed speech in African-American mothers.* University of Illinois at Urbana-Champaign,


Appendices

Appendix A: Contributions to the Research Project

Both Jeffy Chun Yu Ho and Shruti Jamnadass completed the ethics application and designed the project. The development of the ostensive cues coding tool was carried out by Jeffy and Shruti supported this process. Participant recruitment and data collection was carried out jointly. The coding of the study data into the tool was entered individually and Jeffy then used this information to establish inter-rater reliability for the tool. Analysis of results, and the write-up of the theses were carried out individually.
Appendix B: Ethical Approval

26th September 2019

Dr Tobias Nolte
Faculty of Brain Sciences
UCL

Dear Dr Nolte,

Notification of Ethics Approval with Provisos

Project ID/Title: 16673/001: ‘Ostensive Communication, Mentalising and Learning in Mother-Infant Interactions.’

I am pleased to confirm in my capacity as Joint Chair of the UCL Research Ethics Committee (REC) that your study has been ethically approved by the UCL REC until 26th September 2020.

Approval is granted on condition that you:

- provide us with the data protection registration number when issued.
- amend your participant info sheet:
  - please note that there is new legislation on data protection - the Data Protection Act 2018.
  - Please include a Data Privacy Notice –see our template point 14.
  - please include information who the participants can complain to, including ethics@ucl.ac.uk
- amend your consent document:
  - We would expect that details of the researchers and data protection officer (Ms Alex Potts data-pro) will be included on the consent form.
  - As you are collecting personal data, you need to inform the participant about the lawful basis for processing their personal data –public task –please see point 3 of our template for your reference

Ethical approval is subject to the following conditions:

Notification of Amendments to the Research

You must seek Chair’s approval for proposed amendments (to include extensions to the duration of the project) to the research for which this approval has been given. Each research project is reviewed separately and if there are significant changes to the research protocol you should seek confirmation of continued ethical approval by completing an ‘Amendment Approval Request Form’ http://ethics.grad.ucl.ac.uk/responsibilities.php

Adverse Event Reporting – Serious and Non-Serious

It is your responsibility to report to the Committee any unanticipated problems or adverse events involving risks to participants or others. The Ethics Committee should be notified of all serious adverse events via the Ethics Committee Administrator (ethics@ucl.ac.uk) immediately the incident occurs. Where the adverse incident is unexpected and serious, the Joint Chairs will decide whether the study should be terminated.
Appendix C: Participant Information Sheet

A study on mother-infant communication and infant learning

UCL Research Ethics Committee Approval ID Number: 16673/001

We would like to invite you and your child to participate in this study. Before you decide whether you would like to take part, it is important for you to read the following information carefully. Please ask us if there is anything that is not clear or if you would like more information.

**What is the purpose of this study?**

This research is being carried out to help us understand more about how babies learn and what aspects of the mother-child communication promote their learning. We are inviting mothers with babies between 16 and 22 months of age to take part in this study.

**Do I have to take part?**

It is entirely up to you to decide whether you take part in this study. If you do agree to participate you will still be free to change your mind at any time without having to give a reason. A decision to pull out of the study at any time will not affect the standard of care you receive and you may withdraw your data from the study at any point up until it is transcribed for use in the final report.

**What will I have to do if I take part?**
If you decide to take part in this study, a researcher will arrange to see you and your child together at the Anna Freud National Centre for Children and Families (AFNCCF) where you will be asked to do the following:

- **Complete some questionnaires** about you and your child and a **demographics form**. These should take about 45 minutes to complete.

- **Spend time with your child** for a little while as you would normally do at home. This should take about 10 minutes and will be video-recorded.

- **Teach your child to perform some simple tasks** which will be explained to you by the research team and then let your child carry out the task on his/her own. This should take about 45 minutes and will be video-recorded.

- **Complete a short screening measure** on the day of the study. This will be carried out after the study has taken place and will take approximately 20 minutes.

**Will my taking part in the study be confidential?**

All information collected will be kept strictly confidential and the research team will be the only ones with access to it. We make sure of this by keeping the questionnaires and video-recordings locked away, and by only writing your assigned identity number not your name or any other personal details on these. The consent form that you sign will be stored separately to the video footage. We will not disclose any personal information you give us, unless in the very rare circumstances that there are concerns about the safety of you and/or your child. In these cases we would talk with you first before discussing these concerns with anybody else.

All data collected will only be used for the purposes of this study and will be stored in accordance with the Data Protection Act (1998). All data will be destroyed 1 year after the study ends.

**What are the possible disadvantages and risks of taking part?**
We do not anticipate any particular risks to taking part in this study. If you are unhappy about any aspect(s) of it, please contact a member of the research team, who will do their best to help resolve any difficulties.

What will happen to the results of the study?

We may publish out results in a scientific journal and present findings at scientific conferences. Your data will be personally unidentifiable at all times.

Who is organising and funding the project and who has granted ethical approval?

The project is funded by the AFNCCF.
The study has been approved by the Ethics Committee of University College London, reference number: 16673/001

Can I speak to someone to find out more information?

Yes, please do get in touch if you would like to know more. You can contact:

- Tobias Nolte: Tobias.NolteMD@annafreud.org
- Shruti Jamnadass: Shruti.Jamnadass@annafreud.org
- Jeffy C. Y. Ho: Jeffy.Ho@annafreud.org

If you have any complaints or concerns about the study, please contact Tobias Nolte at Tobias.NolteMD@annafreud.org and the UCL Research Ethics Committee at ethics@ucl.ac.uk

Thank you for taking the time to read through this information sheet!

Local Data Protection Privacy Notice
The controller for this project will be University College London (UCL). The UCL Data Protection Officer provides oversight of UCL activities involving the processing of personal data, and can be contacted at data-protection@ucl.ac.uk

This ‘local’ privacy notice sets out the information that applies to this particular study. Further information on how UCL uses participant information can be found in our ‘general’ privacy notice: https://www.ucl.ac.uk/legal-services/privacy/ucl-general-research-participant-privacy-notice

The categories of personal data used will be as follows:
Your contact details such as email address and phone number. Other demographic information such as yours and your child’s age, ethnic background, marital status, housing, education and employment. The study will be video recorded for the purpose of data collection and later analysis. This video material will be stored for the purposes of the study and will be discarded once the purpose of the study has been met.

The lawful basis that would be used to process your personal data will be to carry out a study in the public interest.

The lawful basis used to process special category personal data will be for scientific and historical research or statistical purposes.

Your personal data will be processed so long as it is required for the research project. If we are able to anonymise or pseudonymise the personal data you provide we will undertake this, and will endeavour to minimise the processing of personal data where possible.
If you are concerned about how your personal data is being processed, or if you would like to contact us about your rights, please contact UCL in the first instance at data-protection@ucl.ac.uk.
Appendix D: Participant Consent Form

CONSENT FORM

Please complete this form after you have read the Information Sheet and/or listened to an explanation about the programme evaluation. Thank you for your interest in taking part in this study. If you have any questions arising from the Information Sheet or explanation already given to you, please ask the researcher before you decide if you would like to take part. You will be given a copy of this Consent Form to keep and refer to at any time.

Initials

1. I confirm that I have read and understood the information sheet for this study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.

2. I understand that my participation is voluntary and that I am free to withdraw from this study at any time, without having to give a reason and without my legal rights being affected.

3. I understand that my personal information (contact details, age and other demographic information such as ethnicity) will be used for the purposes explained to me. I understand that according to data protection legislation, ‘public task’ will be the lawful basis for processing.

4. I agree to be filmed for this study and understand the purpose of the filming and that all footage will only be used for purposes relating to this study.

5. I understand that it may be possible to identify me and my child from the video footage that will only be used for research purposes.

6. I confirm I have parental responsibility for the named child.
7. I agree to be interviewed for this study and understand that my responses will remain confidential and will only be used for research purposes.

8. I agree to have my interview audio-recorded. I understand that this information will remain confidential and anonymous.

9. I agree to take part in the above study.

<table>
<thead>
<tr>
<th>Participant name</th>
<th>Signature</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Researcher name</th>
<th>Signature</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Names and Contact Details of Researchers:

Tobias Nolte: Tobias.NolteMD@annafreud.org

Shruti Jamnadass: shruti.jamnadass.17@ucl.ac.uk

Jeffy C. Y. Ho: chun.ho.10@ucl.ac.uk

Principal Researcher:

Profossor Peter Fonagy: p.fonagy@ucl.ac.uk

Data Protection Officer: Ms Alex Potts
Appendix E: Participant Debrief Form

A study on mother-infant communication and infant learning

UCL Research Ethics Committee Approval ID Number: 16673/001

Thank you for taking part in our study.

Our study aims to explore whether mothers’ use of cues when teaching infants, aged 16-22 months, impacts whether they can learn a task. Some examples of cues we are looking at include eye contact and calling the baby’s name.

We are also curious to see whether babies imitate their mother, such as copying the head action to switch on the lightbulb, or whether they can determine that the goal is to switch the light on and thus use their hands instead.

We measured other constructs with the questionnaires we sent you and so will also be exploring the interaction of those constructs with child learning.

As we said before, there is no expectation that that an infant of your baby’s age will learn all the tasks, and this study will add to research in this area so that we are able to learn more about development in this age range.
If you have any further questions, feel free to contact us before June 2020:

- Shruti Jamnadass: shruti.jamnadass.17@ucl.ac.uk
- Jeffy C. Y. Ho: chun.ho.10@ucl.ac.uk

Thank you again and well done to you and your baby!
Appendix F: Recruitment Poster

Mums and Toddlers needed!

Is your child between the age of 16 and 22 months?

The Anna Freud National Centre for Children and Families is conducting a study about how babies learn, and we need your help!

What will taking part involve?

A one hour visit to the Centre of:

- structured play activities for your baby;
- a couple of brief questionnaires;

The Centre is at 4-8 Rodney Street, it’s 10 minutes’ walk from Angel tube Station (Northern Lines).

What are the benefits of taking part?

It’s a visit with Good Fun and you will be making an important contribution with your baby to research on children’s development.

You are compensated £17 for your time and to cover travel expenses to the Centre. Your child will also receive a cool UCL & Anna Freud Centre limited edition Baby Bib.

Interested in participating?
If you are interested in taking part or would like more information, please contact Shruti or Jeffy via Email:

Shruti: Shruti.Jamnadass@annafreud.org

Jeffy: Jeffy.Ho@annafreud.org
Appendix G: Appointment Letter to Participants

The Anna Freud National Centre for Children and Families Anna Freud Centre

4-8 Rodney Street,
London,
N1 9JH

Dear

As discussed, your appointment slot for our Infant Research Study is booked for:

Date: [ ]
Time: [ ]

At: The Anna Freud National Centre for Children and Families Anna Freud Centre
4-8 Rodney Street
London N1 9JH

If you are travelling via underground with a pram, the nearest stations are King’s Cross St Pancras station (Lift & Escalator Access) and Angel station (Escalator Access Only).

Once arrive, please report to the reception and we will meet you there.

For the purpose of the study, please note that we will initially try to have minimal interaction with your baby. For example, we might come across as a little bit cold or less enthusiastic, but please be reassured that it is part of the study.

Please see enclosed some questionnaires. We would be grateful if you could complete these prior to your visit and bring them with you.
If you have any questions or need to re-book your appointment, please either contact Jeffy at Jeffy.Ho@annafreud.org or Shruti at Shruti.Jamnadass@annafreud.org.

We look forward to meeting you and your baby!

Best wishes,

Jeffy C. Y. Ho and Shruti Jamnadass

Clinical Psychologists in Training
Research Department of Clinical, Educational and Health Psychology
University College London
1-19 Torrington Place
London WC1E 7
Appendix H: Instruction Sheet for Experiment

Instruction Sheet for TASK A (first part)

PURPOSE: Help your baby learn to switch on the light with their head

1. Demonstrate touching your head to the box

2. Demonstrate this 3 times

3. Not allowed to physically move your baby’s head to the box

4. After you have shown this 3 times, **DO NOT** let your baby try on immediately and return the box to the experimenter

Instruction Sheet for TASK B

PURPOSE: Teach your baby names of objects

1. Show your baby all 4 objects in the tray by showing and naming them, one at a time, using whatever strategy you like

2. Then, name one object for your baby to select. Do this again.
3. After your baby has successfully shown you 2 objects, we will move on

**Task B continued**

1. There are now 3 objects with labels on them indicating what they are called

2. Show your baby all 3 objects by showing and naming them, one at a time

3. After you have shown and named all objects, repeat this two more times

4. After this, your baby can play with the items for a few seconds

5. Place all 4 items on the tray

6. Ask your baby which one they prefer

7. Say a name of an object that you would like your baby to select
8. Do this for all 3 objects

9. Don’t worry if your baby doesn’t get it right!

10. We will now give you another set of 3 objects – please do the same as above
    (showing, naming, asking baby which they prefer and then naming each object one at a time
    for your baby to select)

   **Instruction Sheet for TASK A (2nd part)**

1. Place the box in front of your baby and ask them to switch on the light

2. Do not demonstrate the task anymore and do not tell them how to switch on the light

   😊

Thank you for taking part in our study
Appendix I: Demographics Questionnaire

Demographic Questionnaire

PPT ID: _______________ Date:___________________

1. What is the age of your baby? ______ (in months)

2. What is your age? ______ (in years)

3. How many other children do you have? ______

4. How old are each of your other children? Please list their ages (in years):
   ______ ______ ______ ______

5. What language(s) do you speak at home?
   __________________________________________________________

6. Place of birth: ________________________________ (City/Country)

7. Have you had a diagnosis of learning disability or learning difficulty? YES / NO
   If Yes, please specify: ________________________________

Ethnicity (please circle):

1. White – British
2. White – Irish
3. White – any other white background
4. Black/Black British – Caribbean
5. Black/Black British – African
6. Black/Black British – any other Black background
7. Mixed – White and Black Caribbean
8. Mixed – White and Black African
9. Mixed – White and Asian
10. Any other mixed background
11. Asian/British Asian – Indian
12. Asian/British Asian – Pakistani
13. Asian/British Asian – Bangladeshi
14. Asian/British Asian – any other Asian background
15. Chinese
16. Any other background not stated

**Mother’s ethnicity (please circle):**

1. White – British
2. White – Irish
3. White – any other white background
4. Black/Black British – Caribbean
5. Black/Black British – African
6. Black/Black British – any other Black background
7. Mixed – White and Black Caribbean
8. Mixed – White and Black African
9. Mixed – White and Asian
10. Any other mixed background
11. Asian/British Asian – Indian
12. Asian/British Asian – Pakistani
13. Asian/British Asian – Bangladeshi
14. Asian/British Asian – any other Asian background
15. Chinese
16. Any other background not stated

**Father’s ethnicity (please circle):**

1. White – British
2. White – Irish
3. White – any other white background
4. Black/Black British – Caribbean
5. Black/Black British – African
6. Black/Black British – any other Black background
7. Mixed – White and Black Caribbean
8. Mixed – White and Black African
9. Mixed – White and Asian
10. Any other mixed background
11. Asian/British Asian – Indian
12. Asian/British Asian – Pakistani
13. Asian/British Asian – Bangladeshi
14. Asian/British Asian – any other Asian background
15. Chinese
16. Any other background not stated (please state):
17. Not stated

**Employment status:**

1. Employed - full time
2. Employed - part time
3. Employed – casual work
4. Self employed
5. Internship/apprenticeship
6. Student
7. Retired
8. Carer
9. Unemployed
Occupation (please state): __________________________________________

Household Income (please circle):

1. Less than £10,000
2. £10,000-20,000
3. £20,000-35,000
4. £35,000-50,000
5. £50,000-75,000
6. £75,000-100,000
7. £100,000 +

Mother’s level of education (please circle):

1. No qualifications
2. Other qualification not listed (e.g. certificate)
3. Vocational level (e.g. NVQ) 1, GCSE (<5 A*-C) or equivalent
4. GCSE (5 or more grades A*-C), vocational level (e.g. NVQ) 2 or equivalent
5. A Level, vocational level (e.g. NVQ) 3 or equivalent
6. Higher education or professional/vocational equivalent
7. Post graduate education or professional/vocational equivalent (e.g. Masters, PhD, MD)

Years in education (total): ______________years

Father’s level of education (please circle):

1. No qualifications
2. Other qualification not listed (e.g. certificate)
3. Vocational level (e.g. NVQ) 1, GCSE (<5 A*-C) or equivalent
4. GCSE (5 or more grades A*-C), vocational level (e.g. NVQ) 2 or equivalent
5. A Level, vocational level (e.g. NVQ) 3 or equivalent
6. Higher education or professional/vocational equivalent

7. Post graduate education or professional/vocational equivalent (e.g. Masters, PhD, MD)

Years in education (total): ______________ years

Have you experienced any significant losses during your childhood (i.e. parents, siblings, close family member)?

<table>
<thead>
<tr>
<th>Who</th>
<th>How old were you at the time of loss? (years)</th>
<th>Due to what circumstances did the loss occur?</th>
</tr>
</thead>
</table>

Have you experienced any significant separations during your childhood (i.e. parents, siblings, close family member)?

<table>
<thead>
<tr>
<th>Who?</th>
<th>How old were you at the time of separation? (years)</th>
<th>Due to what circumstances did the separation occur?</th>
<th>How long for? (months)</th>
</tr>
</thead>
</table>

Are you currently OR have been seen by mental health services?

Yes      No
If you are currently receiving treatment, at what stage of the treatment are you?:

1. Preliminary assessment
2. Waiting list for treatment
3. Psycho-educational group started
4. Psycho-educational group finished waiting for individual therapy
5. Psycho-educational group finished waiting for further group sessions
6. Individual treatment started
7. Other. Please specify: __________________________

When were you first seen by mental health services? ____________________________
(month/year)

If you are no longer being seen, when did you stop working with them? ______
(month/year)

If you are still being seen, when did you start working with your current mental health
team? ____________________________ (month/year)

Are/have you had any psychological therapies as part of your treatment?

| Therapy Type | Service Provider (i.e. Trust, private) | Number of sessions | Average contact time (hrs) | Dates (from/to) |
Are you currently on any medication prescribed by your mental health service
| Name of Drug | Dosage (mg) | Frequency | How long taken for? |
Appendix J: Case Study

Table 41.
ZF, ZR and ZGR total of P24 & P49

<table>
<thead>
<tr>
<th>Participant no.</th>
<th>ZF Total</th>
<th>ZR Total</th>
<th>ZGR Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>P24</td>
<td>-6.48</td>
<td>-1.68</td>
<td>-7.05</td>
</tr>
<tr>
<td>P49</td>
<td>3.40</td>
<td>0.46</td>
<td>7.62</td>
</tr>
</tbody>
</table>

Two participants are selected from the sample for more in-depth observation based on their performance of total ZF, ZR and ZGR in table 33.

P24 has the most consistent below-average performance in all dimensions. She has the lowest total ZF and total ZGR score among all participants. Her ZR total score is 3rd lowest.

P49 has the most consistent above-average performance in all dimensions. She has the highest total ZGR score, the second highest total ZF score. Her total ZR score sits around the middle compared to her counterparts.
Table 42.  
Comparison of ostensive communication performance between P24 and P49

<table>
<thead>
<tr>
<th>Task</th>
<th>Participant no.</th>
<th>Time (z)</th>
<th>Attention (z)</th>
<th>No. of eye contact (z)</th>
<th>Rate of eye contact (z)</th>
<th>No. of Name addressing (z)</th>
<th>Rate of Name addressing (z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task A</td>
<td>P24</td>
<td>64 (-.46)</td>
<td>2 (1.23)</td>
<td>3 (-1.01)</td>
<td>0.05 (-1.07)</td>
<td>2 (-.33)</td>
<td>0.03 (-.21)</td>
</tr>
<tr>
<td></td>
<td>P49</td>
<td>94 (.49)</td>
<td>2 (1.23)</td>
<td>10 (1.41)</td>
<td>0.11 (1.12)</td>
<td>3 (.45)</td>
<td>0.03 (-.19)</td>
</tr>
<tr>
<td>Task B</td>
<td>P24</td>
<td>177 (-1.64)</td>
<td>1.42 (-.53)</td>
<td>5 (-1.09)</td>
<td>0.03 (-1.4)</td>
<td>4 (-.8)</td>
<td>0.02 (.09)</td>
</tr>
<tr>
<td></td>
<td>P49</td>
<td>593 (.53)</td>
<td>1.92 (.74)</td>
<td>16 (.15)</td>
<td>0.03 (-2.3)</td>
<td>19 (.77)</td>
<td>0.03 (.84)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Task</th>
<th>Participant no.</th>
<th>IDS quality (z)</th>
<th>Scaffolding (z)</th>
<th>Rate of Scaffolding (z)</th>
<th>Quality of Contingent Behaviours (z)</th>
<th>Emotional Support (z)</th>
<th>Co-occurrence (z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task A</td>
<td>P24</td>
<td>0.81 (-1)</td>
<td>14 (-.84)</td>
<td>0.22 (-.48)</td>
<td>0.17 (-.97)</td>
<td>0 (-.76)</td>
<td>1.5 (-1.19)</td>
</tr>
<tr>
<td></td>
<td>P49</td>
<td>1.56 (1.02)</td>
<td>22 (.43)</td>
<td>0.23 (-.31)</td>
<td>1.25 (1.88)</td>
<td>0.63 (.48)</td>
<td>4.25 (1.29)</td>
</tr>
<tr>
<td></td>
<td>P24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>Task B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P49</td>
<td>1.75</td>
<td>92</td>
<td>0.16</td>
<td>1.33</td>
<td>2.29</td>
<td>5.17</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.1</td>
<td>35</td>
<td>0.2</td>
<td>0.47</td>
<td>0</td>
<td>1.67</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.25)</td>
<td>(-2.42)</td>
<td>(.15)</td>
<td>(-1.56)</td>
<td>(-1.48)</td>
<td>(-1.68)</td>
<td></td>
</tr>
</tbody>
</table>
1. Task A

P24 spent considerably less amount of time than P49 was Task A. Not only P49 did use a lot more eye contact, the rate of eye contact is also more than double of P24. This has resulted in more frequent and better quality of the back and forth joint attention between P49 and her toddler than P24, which might have partly contributed to the considerable difference in the quality of contingent behaviours between them. In fact, there was no evidence of P24 forming any joint attention with her toddler in the video and the raw data. P24 was pretty much preoccupied with her own agenda of demonstrating the novel action and did not show much concern about where her toddler was cognitively and emotionally at. This is reflected on her relatively low score on quality of contingent behaviours. This was evidenced in the video that there were only a few seconds between each demonstration, it did not seem like she had any intention or attempt to make space for mentalising. In contrast, P49 has formed excellent quality of joint attention. She looked at the switch (box) back and forth for a few times. She also had timely responses to her toddler’s expression in a turn-taking manner before making the next demonstration. She took her time in between demonstration to give her toddler sufficient space to process the new information presented.

P24 also used a lot less amount of scaffolding behaviours to prepare her toddler to engage in the learning task. Her demonstration was rather mechanical without much variations. She tended to use a monotonous mode in her teaching. She used a minimal range of cues, such as commands like “sit” and “look”, and an occasional exaggerated noise, but nothing beyond that. In contrast, P49 used not only considerably more scaffolding behaviours, but also a much more variation. On top of commands and facial expression, P49 used a lot of hand gestures, labelling and facial expression.
In terms of IDS, although P24 seemed to speak in a higher pitch, her pitch tended to stay at a similar level; there was little variation. As she was rushing through the task, she also spoke in a relatively fast pace with little pauses. She used simple language and tended to have full repetition, but the lack of variation in speech content gave a mechanical impression. P49 spoke in a softer pitch with much more variety. She also spoke considerably slower with a lot more pauses. Her used simple language like “shall we look at this”, “let’s looks at this”. She used mostly partial repetition, which created more variations.

P24’s monotonous style might have also reflected on the major difference of co-occurrence when compared with P49. P49 used a much broader combination of ostensive cues than P24.

Overall speaking, P49 showed an impression that she was naturally engaging with her toddler through the task. She enjoyed the interaction. She was attentive to subtle cues while given sufficient space for the toddler to process and react. There was no sense that she was intrusive in any way. P24 brought along a very different impression. She was eager to have her toddler to look at her demonstration without giving much concerns of where her daughter was at. P24 was doing her own thing rather than “having a dance between the mother-toddler dyad”.

2. Task B

P29 used way less amount of time in Task B than P49. P29’s toddler was also much more distracted and less engaged in the learning process. The difference in use of ostensive cues between the two participants becomes very noticeable. P49 used 3 to 4 times more of eye contact, name addressing and scaffolding behaviours than P24.

In terms of contingent behaviours, P24 engaged in more joint attention with her toddler. Yet, these joint attentions were usually short-lived. There was little sense that this
mother-toddler dyad was thinking about the object together. P24 used very limited expanding to broaden her toddler’s cognitive awareness of the object. In contrast, P49 showed a much better quality of joint attention with her toddler. P49 used her gaze to guide her toddler’s attention and looking back and forth between the object and her toddler. It was clear to the observer that they were on a similar page of processing the information. P49 also often used expanding. She described the objects in different dimensions (e.g. colour, shape, sounds and function), which often enlightened her toddler’s curiosity. This is reflected on their distinctive difference in their score on the quality of contingent behaviours.

Similar to Task A, P24 showed a rather monotonous style with minimal variation in the combination of ostensive cues, which is again reflected on the extremely low co-occurrence score ($1.67, z = -1.68$). P24 showed little flexibility in terms of what ostensive cue to use. She tended to use a very similar set of ostensive cues (e.g. repeated state of “this is a ball” with the same hand gesture) despite the fact that it has proven unsuccessful to catch her toddler’s interest or attention. Her use of ostensive cues did not see to have a reciprocal relationship with her toddler’s response, where she made little adjustment according to her toddler’s reaction or communication. In contrast, P49 was way more flexible. She made an adjustment of her use of ostensive cues subject to the responses of her toddler.

Their performance on IDS was similar to that in Task A. The rhythm of P24 remained consistently fast while P49 was gentle, soft and slow. P24 scored low on repetition because at times she said almost nothing apart from repeating the name of the object.

P24 was quite serious and harsh. There was little joy in the interaction. It was like she was taking a serious test with her toddler rather than having a bit of fun alongside the learning. In a few occasions, P24 was a bit upset or impatient, she removed objects from her toddler forcefully. P49 showed a lot more laughter and smiled through the task. She enjoyed the task and tuned into her toddler’s curiosity of exploring unusual objects and names.
### Table 43.
Comparison of PRFQ & PSI between P24 & P49

<table>
<thead>
<tr>
<th>Participant no.</th>
<th>PM (Z)</th>
<th>CM (Z)</th>
<th>IC (Z)</th>
<th>PSI (Z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>P24</td>
<td>1.83 (-.01)</td>
<td>3 (-.81)</td>
<td>4.83 (-1.62)</td>
<td>67 (-.66)</td>
</tr>
<tr>
<td>P49</td>
<td>1 (-.99)</td>
<td>4.83 (.66)</td>
<td>6.5 (.55)</td>
<td>43 (-1.97)</td>
</tr>
</tbody>
</table>
Appendix K: Coding system

<table>
<thead>
<tr>
<th>PP Number</th>
<th>49</th>
</tr>
</thead>
<tbody>
<tr>
<td>Starting</td>
<td>End</td>
</tr>
<tr>
<td>Demonstration 2</td>
<td>0:10:34</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

320
Appendix L: Baby Bib