Prunty, Jonathan Edward (2019) ***Developing Methodological Tools for Ecologically-v***
Infant Eye-Tracking Research. **Doctor of Philosophy (PhD) thesis, University of Kent,.**

Developing Methodological Tools for Ecologically-valid Infant Eye-Tracking Research

Jonathan E. Prunty

A thesis submitted for the degree of Ph.D. in the Faculty of Social Sciences at the University of Kent

September 2019

Abstract

Conducting robust and relevant research is uniquely challenging when investigating infant participants given their limited linguistic and behavioural repertoires. However, the location of an infant's gaze is a reliable index of their attention, and this was exploited in key methodological innovations during the 60s which allowed researchers to answer foundational questions about infants' perceptual abilities within an experimental laboratory setting. Since then, technological improvements such as remote eye-tracking have enabled infants' eye movements to be recorded with greater accuracy and precision. Yet fundamental questions remain concerning the validity of recording the duration of infant looking toward highly artificial stimuli, particularly when investigating complex, higher-order social and cognitive abilities which naturally occur within the context of dynamic interactions. Nevertheless, using naturalistic (e.g. dynamic, interactive) stimuli presents substantial data processing challenges. This thesis introduces the 'Gaze Contingent Social Interaction paradigm', as a flexible methodological template for conducting lab-based eye-tracking experiments with greater ecological validity. This paradigm is applied within several empirical 'worked examples'; research topics of current interest to infant socio-cognitive research that predominantly use static stimulus presentations (e.g. facial race and affect processing). Within these experiments, several methodological tools (both novel and adapted for infant research) are also demonstrated, (e.g. 'Dynamic AOIs', heatmap and temporal analyses) that will hopefully aid developmental researchers to reliably analyse the rich data generated within this paradigm. Overall, the empirical chapters of this thesis highlight key differences between infants' eye movements for naturalistic stimuli compared to static images, which encourages us to consider if the methods currently used in lab-based studies reliably generalise to the 'real-world'.

Acknowledgements

I'd first like to thank Dr David Kelly for his excellent supervision over the last three years. His support has been unwavering, and unfailingly positive despite setbacks. He has been incredibly generous with his time and resources, and has provided an academic example which I myself hope to emulate. I'd also like to thank Jolie Keemink, my co-PhD in the Kent Infant Lab for her help and companionship, as well as all the students that I have worked with in the lab. I am also grateful to the University of Kent for providing the GTA scholarship that has allowed me to pursue this research, and to all the support teams that have facilitated it.

This thesis was completed despite numerous interruptions from my fellow PhD students, my friends; Martina De Lillo, Matt Fysh, Maria Zajaczkowska, Giulia Mangiaracina, Courtney Allen and others, but without whom my time here at Kent would have been immeasurably poorer. This thesis would also not have been possible without my constant companion over these last three years… MATLAB. We began as strangers, but are now firm friends (though these days I seem to be spending more time with my new friend *R*). Finally, I would like to thank my family; my parents, grandparents and siblings, whose support and guidance over the years has made me the person that I am today.

Forthcoming publications from this thesis

Prunty, J. E., Keemink, J. R. & Kelly, D. J. (accepted proposal). Developing ecologically valid eye-tracking methods to investigate how infants use social information: Facial expressions, a worked example. *Perspectives on Psychological Science.* **(Chapter's 5, 6 and 8)**

Prunty, J. E., Norman, M., Keemink, J. R., Keshavarzi-Pour, M. J. & Kelly, D. J. (under review). Infant race preferences are removed following simulated social interactions. *Journal of Cognition and Development.* **(Chapter 3)**

Prunty, J. E., Keemink, J. R., & Kelly, D. J. (in prep). Infants' pupillometry responses to the 'basic' emotions reveal sensitivity to happy and fear faces only. **(Chapter 7)**

Declaration

I declare that this is my own work carried out under the normal terms of supervision.

Jonathan E. Prunty

Table of contents

ABSTRACT .... **I**

ACKNOWLEDGEMENTS .... **II**

CHAPTER 1 .... **1**

METHODOLOGICAL CHALLENGES IN INFANT RESEARCH **1**

A BRIEF HISTORY OF BEHAVIOURAL METHODS IN INFANT PSYCHOLOGY 2
  From the home to the lab *3*
  The science of looking *5*
  From perceiving to thinking *8*
  The ecological approach *10*
  Tracking infant eye movements *13*
CURRENT METHODOLOGICAL CHALLENGES IN INFANT RESEARCH 21

CHAPTER 2 .... **25**

SIMULATING SOCIAL INTERACTIONS USING GAZE-CONTINGENT EYE-TRACKING **25**

DEVELOPMENT WITHIN SOCIAL INTERACTIONS 26
SIMULATING A SOCIAL INTERACTION 29
THE GAZE-CONTINGENT PARADIGM 32
  Procedure *33*
  Participants *34*
  Eye tracking *36*
  Data processing *37*
  Potential modifications of the GC Social Interaction paradigm *38*
  Analysis *40*
  Theoretical contribution *41*
EMPIRICAL 'WORKED EXAMPLES' 42

CHAPTER 3 .... **43**

PREFERENCES FOR RACE AND SOCIAL BEHAVIOUR WITHIN SIMULATED SOCIAL INTERACTIONS ... **43**

ABSTRACT 43
INTRODUCTION 43
  Hypotheses *47*
EXPERIMENT 1 48
METHOD 48
RESULTS 48
DISCUSSION 50
EXPERIMENT 2 50
METHOD 50
  Participants *50*
  Stimuli *51*
  Procedure *52*
  Data processing and visualisation *54*
RESULTS: RACE PREFERENCES 55
  Familiarisation phase: Eye region onset *55*
  Familiarisation phase: Duration *56*
  Preference phase *57*
RESULTS: SOCIAL INTEREST AND SCANNING BEHAVIOUR 58
  AOIs: Familiarisation phase *59*
  AOIs: Preference phase *60*

Discussion 64
Experiment 3 65
Method 65
Participants *65*
Stimuli and procedure *65*
Results: Race preferences 66
Familiarisation phase *66*
Preference phase *67*
Results: Developmental differences in scanning 69
Discussion 72
General Discussion 73
Conclusion *77*

CHAPTER 4 79

DO INFANTS SHOW A PREFERENCE FOR SOCIALLY-ENGAGING BEHAVIOURS? 79

Abstract 79
Introduction 79
Experiment 1 81
Method 81
Results 82
Familiarisation phase *82*
Preference phase *83*
Discussion 83
Experiment 2 84
Method 84
Results 85
Familiarisation phase *85*
Preference phase *85*
Discussion 86
Experiment 3 87
Method 87
Results 87
Familiarisation phase *87*
Preference phase *88*
Combined analyses: Developmental differences *88*
General Discussion 89

CHAPTER 5 92

DO INFANTS SCAN STATIC AND DYNAMIC FACIAL EXPRESSIONS DIFFERENTLY? 92

Abstract 92
Introduction 92
Hypotheses *95*
Methods 95
Participants *95*
Stimuli *96*
Procedure *98*
Analytical tools *100*
Analysis plan *105*
Results 105
Looking durations toward the stimulus *105*
Looking durations toward AOI regions *106*
Summary *108*

Happiness and anger: Heatmap analysis *109*
Happiness and anger: Temporal dynamics............ *110*
Summary............ *114*
Further differences between conditions *115*
DISCUSSION 118

CHAPTER 6 122

THE DEVELOPMENT OF DYNAMIC FACIAL EXPRESSION SCANNING 122

ABSTRACT 122
INTRODUCTION............ 122
METHODS 124
Analysis plan *124*
RESULTS............ 126
AOI analysis *126*
Summary............ *128*
Temporal dynamics *128*
Temporal dynamics: Face *130*
Temporal dynamics: Lower features *132*
Temporal dynamics: Lower face *135*
Summary............ *137*
Heatmaps *137*
Combining temporal and spatial analyses: Surprise............ *143*
DISCUSSION 146
Expression interest............ *146*
Expression scanning............ *148*
Conclusion............ *152*

CHAPTER 7 153

USING PUPILLOMETRY AS A PHYSIOLOGICAL MARKER FOR EMOTIONAL PROCESSING IN INFANTS............ 153

ABSTRACT 153
INTRODUCTION............ 153
Facial expressions *155*
Methodological considerations *157*
METHOD............ 159
Experiment 1: Expressive Eyes *159*
Experiment 2: Expressive Faces *163*
Participants............ *164*
Data processing *165*
Analysis plan............ *166*
RESULTS 1: EXPRESSIVE EYES 166
Mean pupil size............ *166*
Functional data analysis............ *167*
RESULTS 2: EXPRESSIVE FACES 168
Mean pupil size............ *168*
Functional data analysis............ *170*
Luminance-adjusted analysis............ *172*
DISCUSSION 174

CHAPTER 8 180

GENERAL DISCUSSION OF METHODS AND FINDINGS 180

Infants' preferences for behaviours within social interactions *181*
Limitations of spontaneous preference paradigms *186*
Infants' scanning of dynamic faces *188*
Convergent methods: Pupillometry and eye movements *193*
METHODOLOGICAL TOOLS 197
Temporal analysis *200*
Heatmaps *203*
IMPLICATIONS FOR FUTURE RESEARCH 206
CONCLUSIONS 207

REFERENCES 209

Chapter 1

Methodological challenges in infant research

The transformation from newborn to child is remarkable. Many of the complex social, cognitive, perceptual and motor abilities that are fundamental to human adult functioning have their foundations in infancy, and therefore the research that investigates this period of life is invaluable. To consider an analogy, imagine a cook has prepared a dish for you to taste for the first time. After a lifetime of experiencing different flavours, you will likely be able to isolate many of the ingredients. If you have experience cooking, you might be able to intuit several of the methods the cook used. However, it is not until you are allowed into the kitchen, to observe when the cook introduces each of the ingredients (even the secret ones), and how they are combined, that you can have a full appreciation of the nature of the dish. Similarly, our understanding of how humans perceive, think, feel and behave is immeasurably enriched by our understanding of how these important facets of human life develop across ontogeny.

Yet despite its scientific merit, conducting infant research presents unique and substantial methodological challenges. During the first year of life, infants' linguistic and behavioural repertoires are profoundly limited. How do we determine, for instance, at which age infants are able to perceive colour? We cannot ask verbal questions to infants, and even if they could understand what response could they give? What about more complex abilities such as logical reasoning? How could we determine if infants' represent the beliefs and intentions of others? The current thesis will consider the methods currently used in infant research, and will investigate the effectiveness of the 'Gaze-Contingent

Social Interaction' as a methodological solution to some of the challenges faced in the study of infant social development (see Chapter 2) using empirical examples from experiments with six-, nine- and twelve-month-old infants (Chapters 3 – 7). In the present chapter however, the historical development of infant research methods will be considered, and some of the current methodological challenges facing the field will be identified.

A brief history of behavioural methods in infant psychology

The earliest scientific descriptions of infant development consist mainly of observations in the home environment. Notable examples include Jean Héroard's seventeenth-century medical journal of King Louis XIII of France and the eighteenth-century account by Dietrich Tiedmann, which is hailed as the first 'scientific baby diary' (see Wallace, Franklin, & Keegan, 1994 for a review). In the nineteenth century, Charles Darwin (1877) in is his 'biographical sketch of an infant' also preserves his personal observations and comments concerning the development of his infant son, William (or 'Doddy'), using diary entries. Several of his observations presage important findings in developmental psychology, yet this account also contains many errors and ultimately Darwin's observations are little more than unempirical, anecdotal accounts of a father and his son. Nevertheless, by the early twentieth century the systematic observation of one's own children was an established and prominent methodology in the emerging discipline of developmental psychology (Bühler, 1930; W. Stern, 1924; Werner, 1926). Jean Piaget also drew extensively upon observations of his own three children when formulating his stage theory of development (Piaget, 1929, 1964), but crucially, Piaget also supplemented his naturalistic observations with loosely-structured experiments to provide empirical support for his ideas (i.e. the 'clinical method', Figure 1a).

From the home to the lab

Home observations in the form of baby diaries were useful during the initial stages of developmental psychology as they provided rich, naturalistic data from which to form ideas and build theories. However the conclusions we can draw from such methods are necessarily limited in scope due to their subjective nature, small sample and lack of experimental control. Twentieth century American psychologists such as John B. Watson (1913, 1928) and B. F. Skinner (1935) realised this, however instead of increasing the rigour of existing methods, the behaviourists instigated a methodological revolution in which only highly-controlled empirical investigations of particular behaviours were deemed permissible. In the pursuit of objectivity and reproducibility, and in a frustrated kick-back against the ‘medieval’ study of ‘intangible’ and ‘vitalistic’ entities such as consciousness (i.e. the introspective and psychoanalytical methods), the complexities of mental and social development were reduced to simplistic stimulus-response associations (Watson & McDougall, 1929). According to their perspective, there was only behaviour, and behaviours were learned, and thus the development of an infant could be entirely ‘controlled’ by behavioural conditioning; an assertion which Watson tested in the controversial ‘Little Albert’ experiments (Watson & Rayner, 1920, Figure 1b). In hindsight, the behaviourist movement was undoubtedly overzealous and oversimplistic, and we now see a valued role for biological, cognitive and social factors in our psychological development (see Vygotsky, 1978). However, their emphasis on experimental control in psychology has had an enduring effect on the field, culminating in the study of infants moving away from the family home and into the university lab.

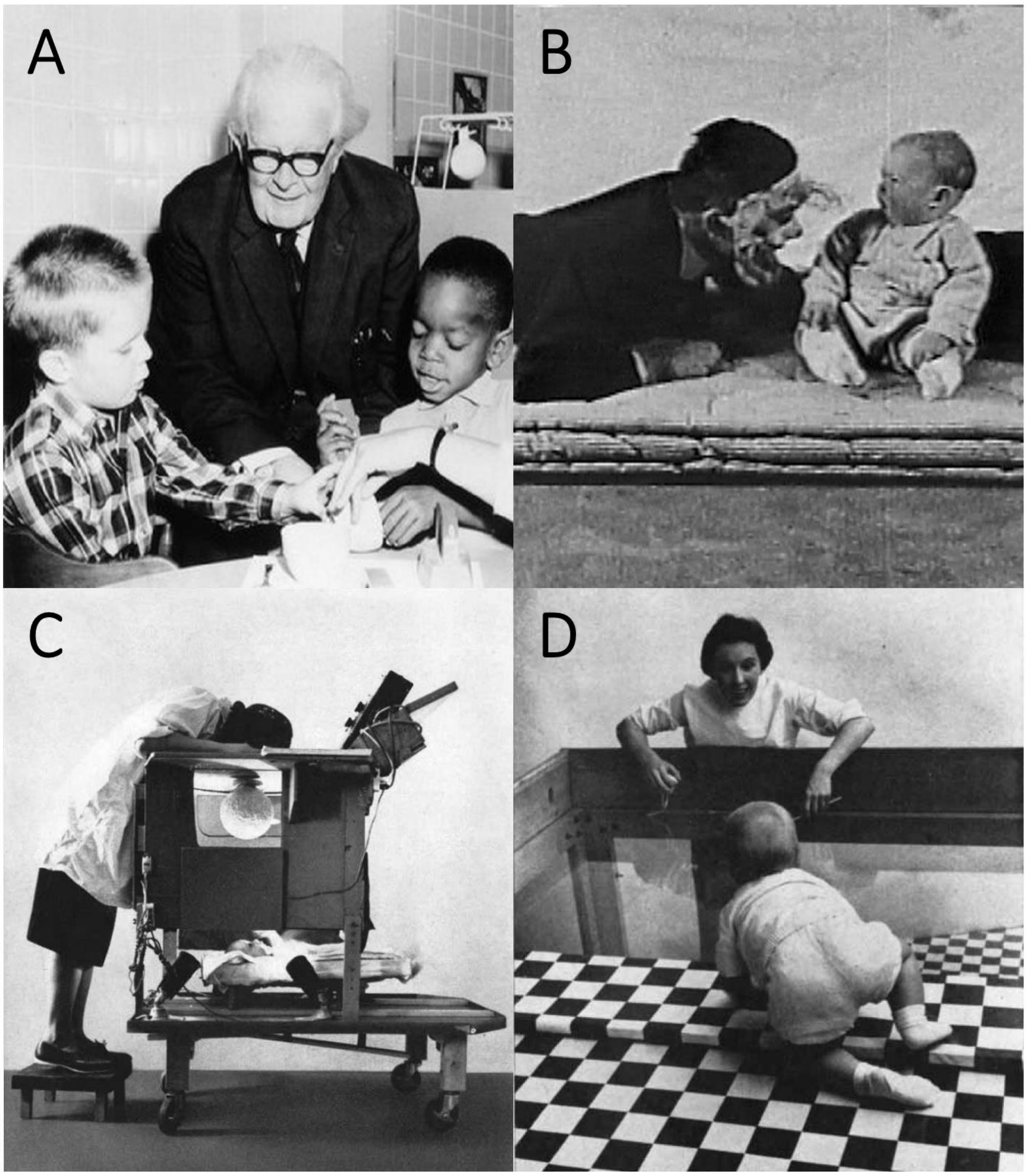

Figure 1. Images depicting changing infant behavioural methods from the 1920s to the 1960s.

Piaget's (1929) 'clinical method' (A), which consisted of a mixture of observation and experimentation in the home environment. Watson & Raynor's (1920) 'Little Albert' experiments (B) which symbolise a revolutionary shift in psychology toward highly-controlled measurements of behaviour in a lab setting. Fantz's (1961, 1963) 'preferential looking chamber' (C) and Gibson & Walk's (1960) 'visual cliff' (D) are examples of methodological breakthroughs in the 1960s which lead to a rapid increase in our understanding of infant perceptual abilities via the experimental research conducted throughout the 1970s and 1980s.

The science of looking

Until the mid-twentieth century, our understanding of infant perception, attention and cognition can be summarised by William James's (1890) assertion that the world for an infant is a 'blooming, buzzing, confusion'; that although infants can sense light, sound and motion, they cannot perceive their form. This conception of the infant all changed after a critical methodological breakthrough was made by Robert L. Fantz (1956, 1961, 1963). Fantz realised that if an infant consistently turned their gaze toward one visual pattern more than another, they must be able to perceptually discriminate between these patterns. From this key insight, he developed a method which would become known as a 'preferential looking task' (or 'forced-choice preferential looking procedure'; see also Dobson & Teller, 1978; Teller, 1979). After first testing the method with chicks and chimpanzees, Fantz (1961, 1963) measured relative looking durations to various pattern-pairings in newborn infants. To do this he used an experimental 'looking chamber' in which an observer could peep through a hole to view reflections of the illuminated stimuli depicted on the infants' cornea, and thus record which stimulus they were attending at any given time point (Figure 1c). Using this technique, Fantz was able to determine that the ability to perceive certain visual forms was innate. Newborn infants preferentially attended to complex patterns over simplistic ones, regardless of the side they were presented, and showed a particularly pronounced preference for face-like patterns.

Yet as the preferential looking task relies on infants possessing an *a priori* or 'spontaneous' preference for one stimulus over another, Fantz (1964) enhanced this method using what became known as a 'habituation-recovery' (or 'visual paired comparison'; VPC) procedure (see also Berkson & Fitz-Gerald, 1963; Friedman, 1972; Sokolov, 1963; see Colombo & Mitchell, 2009 for a review) which enabled researchers to *induce* visual preferences that might not already exist. This was achieved by preceding the paired comparison with a period in which one stimulus was repeatedly presented to the infant until they began to show disinterest (‘habituation’), and thus, if infants were able to perceptually discriminate between the stimuli, they would show greater interest in the relative novelty of the unhabituated stimulus (‘recovery’). Through the careful choice of stimuli within this and similar paradigms, researchers were then able to empirically investigate the inner worlds of infants, and were at last able to accurately determine their cognitive and perceptual capabilities (e.g. can infants perceive depth? Can they discriminate between facial expressions? What about recognising individuals from another race?) The prevailing view at the time that infants inhabited a formless and chaotic world until they *learned* how to create stable and orderly representations, was now no longer tenable, and the ideological battle between nature and nurture was renewed.

Fantz’s key insight was that the location of an infant’s gaze within their visual environment is not random and could be reliably used to indicate the subject of their interest. From the early 70s methods founded on this principle (e.g. Spelke, 1976; Teller, 1979) were used extensively to establish a foundational understanding of infant perceptual functioning on which we have built subsequent knowledge (see Aslin & Smith, 1988; Braddick & Atkinson, 2011; Teller & Movshon, 1986 for reviews). Alternative paradigms were also developed that successfully measured other infant behaviours; both spontaneous (e.g. willingness to crawl, see Figure 1d; E. J. Gibson & Walk, 1960; or infant sucking rates; Kaye, 1967) and conditioned (e.g. conditioned head-turning; Bower, 1966), as well as implicit physiological measurements (e.g. changes in pupil size; Fitzgerald, 1965; or heart rate; Lewis, Kagan, Campbell, & Kalafat, 1966), to index infants’ perceptual abilities. However, the measurement of looking behaviours proved the most valuable, with major developments in our understanding of visual acuity and contrast sensitivity (e.g. Atkinson, Braddick, & Braddick, 1974), colour vision (e.g. Bornstein, 1975), binocularity (e.g. Birch, Gwiazda, & Held, 1982), perceptual constancy (e,g, Slater & Morison, 1985), face and emotion perception (e.g. Field, Cohen, Garcia, & Greenberg, 1984; Goren, Sarty, & Wu, 1975; Nelson, Morse, & Leavitt, 1979), object perception (e.g. Kellman & Spelke, 1983), motion perception (e.g. Fox & McDaniel, 1982) and orienting (e.g. Aslin & Salapatek, 1975) to name but a few. To illustrate this point, compare the difference in visual acuity graphs from 1959 to 1986 (Figure 2). Hypothetical curves projecting from ‘zero’ at birth to adult functioning at 5 or 6 years of age are replaced by empirically-based trajectories describing rapid development during infancy. In his review of infant looking paradigms, Richard N. Aslin commented: “It is no exaggeration to say that without looking time measures we would know very little about nearly any aspect of infant development” (Aslin, 2007, p. 48). Therefore, as a consequence of these methodological innovations, the 70s and early 80s were a burgeoning period for infant perceptual research. Once this perceptual foundation was laid, researchers then began to apply looking behaviour methods to other questions, such as those relating to higher-order cognitive development. If infants can *perceive* in a similar way to adults, they reasoned, could they also *think* similarly?

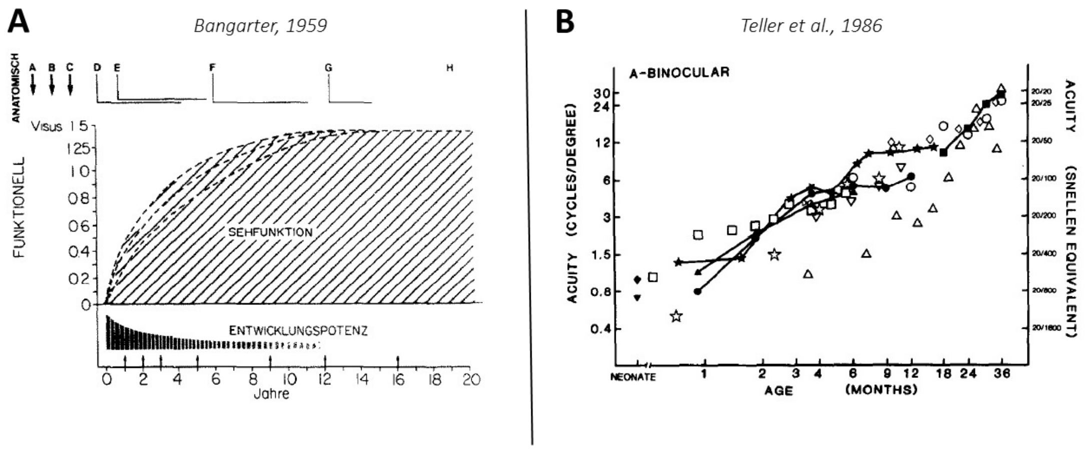

Figure 2. Developmental trajectories of visual function from (A) 1959 and (B) 1986.

The earlier graph depicts hypothetical curves inferred from anatomical landmarks (‘Anatomisch’, a-h), with visual ‘function’ (‘Funktionell’) rising from zero at birth to an adult-level at 5-6 years (‘Jahre’). The later graph demonstrates the advances in perceptual knowledge brought about by new methodological techniques during the 60s and 70s. It uses empirical data from a VPC paradigm to plot the developmental trajectory of visual acuity which rapidly develops to approximately adult levels during infancy. Graphs reproduced from Teller & Movshon (1986).

From perceiving to thinking

Thought is intrinsically connected to language. Can preverbal infants think? And if so, is it possible to assess what they know without being able to provide verbal instruction or receive verbal responses? Are looking behaviours a valid tool to investigate infant cognitive abilities? These questions were part of an important methodological debate that occurred during the late 90s and early 2000s (see discussions in Aslin, 2000, 2007; Baillargeon, 1999; Bogartz, Shinskey, & Speaker, 1997; Cohen, 2004; Haith, 1998; Meltzoff & Moore, 1998; Munakata, 2000; Munakata, McClelland, Johnson, & Siegler, 1997; Smith, 1999; Spelke, 1998), following seminal papers which seemed to demonstrate remarkably precocious cognitive abilities in infants. For instance, these studies claimed that young infants possess sophisticated knowledge about the world; knowledge about objects and their physical properties (Baillargeon, 1986; Baillargeon, Spelke, & Wasserman, 1985; Johnson & Aslin, 1995; Spelke, Kestenbaum, Simons, & Wein, 1995), knowledge about numerical and logical relationships (Leslie, 1982, 1984; Wynn, 1992) and knowledge about the beliefs and goals of others (Gergely, Knadasdy, Csibra, & Biro, 1995; Onishi & Baillargeon, 2005; Woodward, 1998). The critics argued that many of these studies were guilty of an *overly rich* interpretation of looking behaviours; that researchers were too readily attributing high-level cognitive constructs to explain behaviours where a low-level perceptual interpretation would suffice (e.g. Haith, 1998; Smith, 1999).

The Violation of Expectation (VoE) paradigm attracted the most criticism. This method, still popular today (e.g. Cesana-Arlotti et al., 2018), is a modified version of the widely-used habituation-recovery procedure. Infants are familiarised with an event sequence before being presented with either a ‘possible’ or ‘impossible’ outcome; the logic being that the *unexpected*, impossible event would be *surprising*, and thus consistently attract longer looking from infants. In a well-known example, Baillargeon and colleagues (Baillargeon, 1987; Baillargeon et al., 1985) familiarised infants with an animation of a flat screen (or ‘drawbridge’) which rotated 180° toward or away from them. The test events included a cube placed behind the drawbridge which either caused it to stop rotating once making contact (possible event) or did not perturb the drawbridge which continued its rotation ‘through’ the cube (impossible event). Infants looked longer at the impossible event, despite it being more ‘familiar’. In contrast to the scientific consensus at the time, this finding suggested that infants already *know* that occluded objects continue to exist (object permanence; c.f. Piaget, 1954), and that one solid object cannot move through another. Yet subsequent work showed that infants who aren’t sufficiently habituated will look longer at the impossible event due to a perceptual preference for familiarity, not because their cognitive expectations were violated (Cashon & Cohen, 2000; Rivera, Wakeley, & Langer, 1999; Schilling, 2000; see Cohen, 2004), casting doubt on the nativist interpretation of these findings.

This methodological discussion about the validity of VoE paradigms was only a prominent skirmish symptomatic of larger divides in infant psychology. Researchers questioned whether humans possess innate knowledge (see Spelke, 1998; Spelke, Breinlinger, Macomber, & Jacobson, 1992), whether infant cognition can and should be investigated via looking behaviours (see Haith, 1998), and what, exactly, does infant looking really mean anyway? For instance, an infant's increased looking to an impossible event may reflect their *perception* that something is 'odd', without their *knowing* precisely why it is odd. Even if we grant that infants possess these cognitive constructs to some degree, their comprehension and representation of the world is likely to be substantially different from an adult's. And although we can easily conceive of a graded development in visual acuity (many of us experience less than perfect vision), it is less straightforward to envisage what a graded development of cognitive constructs such as causality or object permanence would look like. This debate is yet to be fully resolved, and many of these questions continue to be discussed to this day (e.g. Aslin, 2007, 2012; Burge, 2018; Heyes, 2014).

The ecological approach

Although the nature of infant research since the 60s has undoubtedly shifted toward increasingly quantitative and highly-controlled laboratory investigations, broad methodological questions about the meaningfulness and generalisability of lab-based findings were also being considered. In the 70s many researchers (Bronfenbrenner, 1977, 1979; J J Gibson, 1979; Herrnstein, 1977; Jenkins, 1974; Kuhn, 1978; McCall, 1977;

Neisser, 1976; see Gibbs, 1979) began to sense that something was 'left behind' as the field moved away from the observational methods of Piaget and the early baby diarists. Urie Bronfenbrenner (1977, 1979) lamented that in the pursuit of rigour, research was losing *relevance*. He argued that the emerging trend of lab-based experiments (he notes only 8% of the studies published between 1972 and 1974 were observational, while 76% were laboratory experiments) consisted of studies that were elegantly-designed but ultimately limited in scope. In his opinion, developmental psychology had become "the science of the strange behaviour of children in strange situations with strange adults for the briefest possible periods of time" (Bronfenbrenner, 1977, p. 513). It was clear to him that artificially isolating behaviours within an unfamiliar environment would produce "substantial and systematic" differences compared to the investigation of development embedded within its natural context (p. 516). Nevertheless, he viewed the perceived trade-off between 'rigour' and 'relevance' (also referred to as 'certainty vs authenticity' or 'internal vs external validity'; see Gibbs, 1979) as a false dichotomy, and proposed that a naturalistic-experimental (or 'ecological') approach was both necessary and achievable.

Bronfenbrenner's assessment was interpreted by some as a call to move developmental research away from the lab and back into the naturalistic context of the family home, although this time ensuring that research in the 'field' would be conducted with sufficient rigour (see Dunn & Kendrick, 1980 for a good example of this approach). However, Bronfenbrenner (1977) also noted that ecological validity did not *necessarily* require a change of research setting, but that the requirements of ecological research were instead highly dependent on the particular research question being considered (see also Lewkowicz, 2001). Moreover, since the earliest discussions of ecological validity (Brunswik, 1943; Lewin, 1943) a concern for the appropriateness of the stimulus materials and procedures within experiments, rather than the situation of a lab environment *per se*, has been at the forefront of this debate (see E. J. Gibson, 1969; J. J. Gibson, 1966; Lewkowicz, 2001; Lickliter & Bahrick, 2001; Neisser, 1976; Schmuckler, 2001; Walker-Andrews & Bahrick, 2001). As James J. Gibson (1979, p. 3) commented: “It is not true that the laboratory can never be like life. The laboratory must be like life!”

The essential question of ecological validity is this: are the infant behaviours recorded in response to our experimental stimuli, procedures and setting *representative* of the real-world phenomena we are purporting to study? This is an important question to pose, however it likely lacks operational precision as researchers from different theoretical perspectives often do not agree about what constitutes ‘representativeness’, and thus what components of the phenomenon are critical to retain in our investigations (see Adolph, 2020; Lewkowicz, 2001; Schmuckler, 2001). This is, of course, reminiscent of the debate discussed earlier where the field tended to divide along *theoretical* (i.e. nativist vs empiricist) party lines when debating the *methodological* question of whether looking behaviours were an appropriate index of infant cognition. The problem of ecological validity also becomes exacerbated when we move from simple questions (e.g. can infants perceive gridlines? see Atkinson et al., 1974) to more complex questions (e.g. can infants represent other’s beliefs and intentions? see Onishi & Baillargeon, 2005) as the critical contextual factors that are necessary to faithfully represent the phenomena also increase. For instance, what are the critical components that we need to preserve in order to ecologically investigate an infant’s ability to understand social cues (e.g. Hood, Willen, & Driver, 1998; Nelson et al., 1979; Onishi & Baillargeon, 2005) whilst remaining confident that our findings meaningfully generalise to the real world? In the current era of infant research such complex questions are commonplace, and the question of ecological validity remains relevant and important (Adolph, 2020; Tamis-LeMonda, Kuchirko, Luo, Escobar, & Bornstein, 2017). Today, researchers are increasingly looking to methodological advances for an answer (see Eckstein, Guerra-Carrillo, Miller Singley, & Bunge, 2017; Gredebäck, Johnson, & Von Hofsten, 2010; Hepach, Vaish, & Tomasello, 2015; Hoehl & Markova, 2018; Striano & Reid, 2006), and an important emerging theme is this: can technology be used to help us to achieve sufficient *rigour* in our methods, but without sacrificing their *relevance*?

Tracking infant eye movements

In the last few decades human society has seen unparalleled technological progress, and accordingly the methodologies which we use to investigate infant behaviours have also developed rapidly. Fantz's (1956, 1961) original preferential looking paradigm required experimenters to peep through a one-inch hole situated centrally between two visual stimuli in the ceiling of a 'looking chamber'. In order to determine an infant's visual preference, they recorded the duration (via stopwatch) that the target stimulus was mirrored upon the pupils of the infant's eyes (see Figure 1c). Alternatively, a camera could be used to photograph the infant's eyes (20 exposures per 15-second test period; i.e. 1.3 Hz), and durations were calculated from these images (Fantz, 1956). These low-tech, subjective methods appear crude compared to today's standard, but the logic and methodological challenges illustrated in this example remain unchanged today; namely to accurately determine *where* an infant is looking, and *how long* they are looking there for.

With the advent of modern eye tracking, the location and duration of infant looking can now be recorded with substantially greater accuracy, precision and efficiency. While initial attempts to measure gaze patterns used imprecise, bulky and unforgiving contraptions mounted on the head or directly on the eye (e.g. Yarbus, 1967; see Aslin & McMurray, 2004; Kowler, 2011; Schütz, Braun, & Gegenfurtner, 2011 for reviews), today eye trackers such as the EyeLink 1000+ (SR Research, 2010) can record eye movements (using corneal reflection; see Haith, 1969; Salapatek & Kessen, 1966) remotely and with excellent accuracy despite head movements; therefore enabling research with infants (see Feng, 2011; Francois, Coufal, & Chaparro, 2018; Gredebäck, Johnson, & von Hofsten, 2010; Hepach & Westermann, 2016; Oakes, 2012 for reviews). For instance, the EyeLink can record infant eye movements from both eyes with a sampling rate of 500 Hz (i.e. one sample every 2 milliseconds), with a spatial resolution of 0.01 visual degrees (i.e. approx. 0.1mm for a stimulus 600mm away), and with an average gaze position error of less than 0.5 visual degrees (i.e. less than 5mm for a stimulus 600mm away). What this means is that while Fantz could determine, with a certain degree of confidence, *if* an infant was looking at a particular stimulus, we can now accurately determine *where* on that stimulus the infant is looking, and precisely *when* their looking toward that location began and ended. These technological developments therefore allow us to ask fundamentally different questions. Instead of simply determining *if* a stimulus is perceived by infants, accurate eye tracking promises researchers the potential for investigating *how* a stimulus is perceived.

The basic premise of eye-tracking research is that the location of one's gaze is highly correlated to the location of one's attention (Buswell, 1935; James et al., 1890; Von Helmholtz, 1925), and thus by tracking a subject's gaze, we can follow the 'path' of their attention as they perceive a stimulus. Consider the physiology of looking; only a small area of the retina (the fovea) provides high-resolution vision, therefore sequences of rapid eye movements (saccades) are initiated in order to inspect regions of interest within a scene in greater detail (fixations; see Kowler, 2011; Leigh & Zee, 2006). Top-down cognitive factors such as the perceiver's intentions (Broadbent, 1958; Deutsch & Deutsch, 1963; J. J. Gibson, 1966) interact with bottom-up stimulus characteristics (e.g. visual saliency) in order to determine which regions of a scene (initially detected parafoveally) are selected for detailed inspection by the fovea. Early explorations of eye movement patterns (Noton & Stark, 1971a, 1971b; Yarbus, 1967) illustrated this interaction, with sequences of fixations over informative regions creating common ‘scanpaths’ which varied according to task objectives. For instance, when Alfred L. Yarbus (1967) instructed participants to estimate the ages of the individuals in Ropin’s painting ‘An Unexpected Visitor’ (Figure 3a), fixations were concentrated on the faces in the scene (Figure 3b), but when instructed to determine the wealth of the family, or remember the position of objects, there was much more diffuse scanning as the participant’s attention was drawn away from the people (Figure 3d). When instructed to determine how long the visitor had been away, participants showed economical scanning of the pertinent social cues in the scene (e.g. focusing on the facial expressions of relevant individuals; Figure 3e). Eye tracking can therefore be a useful indicator of top-down *cognitive* strategies as well as the low-level perceptual factors that guide visual attention, and has therefore become an established and popular method within the adult literature (Duchowski, 2017; Rayner, 1998).

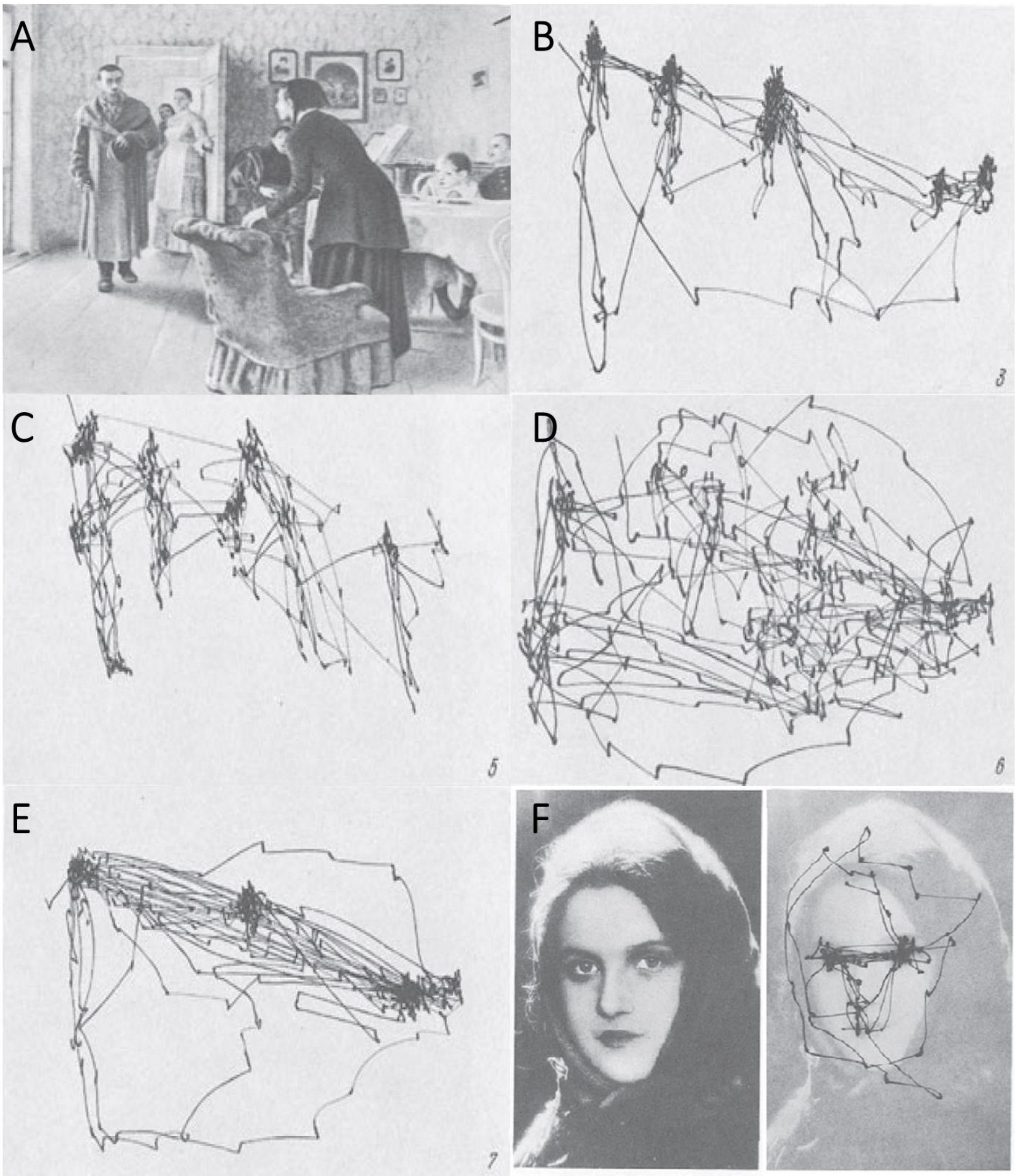

Figure 3. Early scanning patterns recorded by Alfred Yarbus (1967).

Alfred Yarbus (1967) recorded participants' patterns of scanning when viewing I. E. Repin's painting 'An Unexpected Visitor' (A). Scanpaths (B-E) demonstrate that faces capture the attention of the observer more than other objects, and within the face (F) the internal features are fixated the most. Before eye movements for Repin's painting were recorded, Yarbus asked participants to estimate the ages of the people (B); remember their clothes (C); remember the position of the people and objects in the room (D); and estimate how long the 'unexpected visitor' had been away from the family (E). These examples show that scanning strategies for complex scenes are task-specific. For instance, restricting fixations to faces in B, but scanning up and down the clothing in C. Exploring all the objects in the scene in D, but focusing intensely on the faces of the 'visitor', 'mother' and 'children' in E.

As oculomotor precision develops rapidly in the first six months of life and infant-friendly eye-tracking technologies are now widely accessible, modern infant labs are increasingly using eye-tracking to investigate infants' cognitive and perceptual abilities (see Eckstein, Guerra-Carrillo, Miller Singley, & Bunge, 2017; Gredebäck et al., 2010; Hepach, Vaish, & Tomasello, 2015; Oakes, 2012 for reviews). The main benefit of modern eye tracking, alongside the ability to determine global aspects of visual attention (e.g. stimulus preference or habituation) with greater objectivity, accuracy and efficiency, is that it also enables the investigation of specific eye-movement behaviours. Through the logic of Fantz (1956), that the location of an infant's look is an indication of their interest, we can explore looking durations toward areas of interest (AOIs) *within* a stimulus. For instance, this technique has been employed effectively within face processing research since the earliest developmental eye-tracking experiments (Hainline, 1978; Haith, Bergman, & Moore, 1977; Maurer & Salapatek, 1976; see Gredebäck, Johnson, et al., 2010). Within these studies, relative looking toward facial AOIs are compared to inform researchers about *how* infants of different ages use facial information (e.g. Chawarska & Shic, 2009; Soussignan et al., 2017; Wheeler et al., 2011; Xiao et al., 2015). One notable finding from the application of this method (Lewkowicz & Hansen-Tift, 2012; Tenenbaum, Shah, Sobel, Malle, & Morgan, 2013) is that although younger infants and adults predominantly fixate a speaker's eye-region, infants toward the end of their first year look longer toward the mouth; a shift that coincides with the emergence of speech and suggests that infants' selective visual attention facilitates language development.

Cumulative fixation durations to individual regions of the stimulus can be informative for infant research, however the rich data recorded using eye trackers also provides the opportunity to consider other metrics of infant visual attention (see Yu, Yurovsky, & Xu, 2012). Visual orienting, for instance, as indexed by an infant's 'first look’ (i.e. the target of the first saccade following stimulus onset) or their ‘speed of orienting’ (i.e. saccadic latency or cumulative time before an AOI is fixated), is another potential indicator of which aspects of the stimulus are perceptually prioritised by infants. Investigations of face orienting are particularly striking (Elsabbagh et al., 2013; Kelly, Duarte, Meary, Bindemann, & Pascalis, 2019), with even young infants demonstrating rapid and reliable detection of faces in complex naturalistic scenes. Furthermore, when combined with dynamic stimulus presentations, infant gaze shifts can be a particularly powerful index of their cognitive capabilities. If infants show consistent looking to a specific region directly following a stimulus event, we can infer that infants are *using* the information presented on the screen to guide their visual attention. For instance, our understanding of infants’ ability to use social cues such as following the direction of an adult’s gaze, has been greatly enhanced by eye-tracking paradigms (e.g. Gredebäck, Fikke, & Melinder, 2010; Itier, Villate, & Ryan, 2007; Nyström, Bölte, Falck-Ytter, & The EASE Team, 2017; Senju & Csibra, 2008; Xiao et al., 2018). Furthermore, if infants consistently orient to a stimulus region *before* an interesting event occurs, it suggests that they can use stimulus information to *anticipate* future events (McMurray & Aslin, 2004). Such ‘predictive’ gaze shifts have been used to support findings from the VoE paradigm (e.g. knowledge about objects, knowledge about other’s beliefs and goals), and provide a more detailed insight into the higher cognitive abilities of infants (Falck-Ytter, Gredebäck, & von Hofsten, 2006; Johnson, Amso, & Slemmer, 2003; Kochukhova & Gredebäck, 2010; Southgate, Senju, & Csibra, 2007; Von Hofsten, Kochukhova, & Rosander, 2007).

Often, several different eye movement measures are analysed in conjunction, and are sometimes recorded alongside other independent variables, such as infants’ spontaneous behavioural responses (e.g. reaching, facial expressions) or implicit physiological responses (e.g. pupil size, skin conductance, event-related potentials

(ERPs), heart rate). Measuring additional responses can clarify and strengthen findings from looking behaviour methods, and guard against potential ‘low-level’ interpretations of the data (see Aslin, 2007; Haith, 1998; Heyes, 2014). For instance, pupil sizes are automatically recorded by modern eye trackers alongside gaze data, and as pupillary dilations are a correlate of physiological arousal, they can therefore be used as a complementary indicator of cognitive and affective processing (e.g. Cesana-Arlotti et al., 2018; Gredebäck & Melinder, 2010; Jessen, Altvater-Mackensen, & Grossmann, 2016; Wagner, Luyster, Tager-Flusberg, & Nelson, 2016; see Eckstein et al., 2017; Hepach & Westermann, 2016; Jackson & Sirois, 2009; Laeng, Sirois, & Gredebäck, 2012 for reviews). In one notable example (Sirois & Jackson, 2012), pupillometry was used alongside the VoE ‘drawbridge’ paradigm (Baillargeon, 1987; Baillargeon et al., 1985). Looking durations were consistent with Baillargeon and colleagues’ findings, but the pupil data supported the low-level perceptual interpretation of these results, suggesting that 10-month-olds do not in fact possess object permanence. Convergent measures are therefore one way in which technological advances can help to resolve some of the methodological questions within infant looking-time research.

Recent technological developments are also enabling eye-tracking methods to be implemented in novel ways, allowing research to be conducted with greater flexibility and ecology. Improvements in head-mounted eye tracking (Franchak, Kretch, Soska, & Adolph, 2011; Franchak, Kretch, Soska, Babcock, & Adolph, 2010), are allowing researchers to gain an ‘infant’s eye-view’ as they explore their everyday environments. These technologies offer exciting new opportunities to record natural behaviours embedded in real-world environments, rather than ‘strange behaviours’ embedded in ‘strange environments’ (see Bronfenbrenner, 1977). This method has already identified several examples where results from ‘real-world’ eye-tracking differ in comparison to results from desktop-mounted flat-screen displays (e.g. Kretch & Adolph, 2015; Kretch, Franchak, & Adolph, 2014; Yu & Smith, 2013), once again reminding researchers of lingering questions concerning the ecological validity of our methods.

Ingenious ‘gaze-contingent’ eye-tracking paradigms (i.e. paradigms where the stimuli presented vary according to a participant’s eye movements; see Duchowski, Cournia, & Murphy, 2004) can also be used to allow infants to take a more active role in the experiment. Infants quickly discover novel forms of agency, and can learn to manipulate on-screen stimuli using their eyes after surprisingly brief training sequences (Deligianni, Senju, Gergely, & Csibra, 2011; Wang et al., 2012). The applications of eye tracking methods that enable young infants to actively interact with on-screen stimuli are many and varied (see Keemink, Keshavarzi-Pour, & Kelly, 2019; Powell, Wass, Erichsen, & Leekam, 2016; Vernetti et al., 2018; Vernetti, Smith, & Senju, 2017; Wass, Porayska-Pomsta, & Johnson, 2011; Wilms et al., 2010 for examples), and have numerous methodological advantages over classical conditioning and passive looking paradigms (Wang et al., 2012). In particular, many of the questions concerning the meaningfulness of infant looking are rendered irrelevant when infants’ can demonstrate agency by using their gaze to directly influence their environment.

Technological advancements therefore are providing potential answers to some of the important historical questions previously discussed, such as whether infant looking behaviours recorded in the lab generalise to the real world and also whether looking behaviours are a suitable index for assessing infant higher cognitive abilities. However, whether the contribution of technology will offer ‘real’ and lasting solutions is yet to be determined.

Current methodological challenges in infant research

The ability to automatically generate a rich and detailed record of infant responses co-occurring with dynamic stimulus presentations provides modern infancy researchers with unparalleled opportunities, but it also presents new challenges. Even short eye-tracking experiments produce large volumes of data (e.g. recording gaze position for 100 infants, one minute each, at 500Hz generates 3 million samples). In order to find useful answers to our research questions, these data have to be summarised, analysed and interpreted with care, otherwise we risk sacrificing the richness of our data, or worse distorting or misrepresenting what it tells us (see Yu et al., 2012).

This data processing challenge is amplified as we move toward more ecological paradigms, which tend to generate rich datasets but with few constraints, leading to high levels of data complexity. In such paradigms, systematically establishing even basic properties of the looking behaviours we are recording presents us with difficulties. Take dynamism, for instance; how does one determine what infants are looking at for a stimulus that changes over time? Even for static stimuli, there is considerable variation in the literature as to how AOIs are defined (e.g. compare Hills & Pake, 2013; Jones & Klin, 2013; Wheeler et al., 2011). Eye tracking data are continuous, so the number of AOI regions we choose to define, their size and precisely where we place them on the stimulus, can have a substantial impact on our results (see Caldara & Miellet, 2011; Hessels, Kemner, van den Boomen, & Hooge, 2016). Yet in *dynamic* presentations AOIs would also change in size and spatial location, rendering traditional AOIs almost entirely impractical. Furthermore, as the information within dynamic displays is presented over time, performing simple statistical comparisons that average looking durations for the whole display period may also miss vital behavioural patterns, and carving up the time course into arbitrary bins also risks distorting the data (D. J. Barr, 2008). Implementing any level of interactivity or multimodality alongside this dynamism also risks introducing additional methodological complexity, and when the measurement of infant looking is no longer constrained to a 2D screen (e.g. in ‘real-world’ head-mounted eye tracking paradigms) further complications arise as the potential variability in the visual information infants might encounter escalates dramatically. At present, increasing the naturalism of our stimuli and methods often requires researchers to resort to the subjective and labour-intensive strategy of manually coding eye-tracking data frame-by-frame.

It is clear therefore that although technological advances have increased the potential for recording detailed, ecological measurements of infant looking, practically implementing these techniques remains a challenge. As a consequence, infant researchers have been reluctant to consider using methods that have greater relevance to an infant’s real-world experiences (e.g. dynamic, interactive or multimodal displays), and our understanding of how infants perceive and understand social information is still largely based on research presenting unimodal, static and unresponsive stimuli within the ‘strange’ context of the lab setting. To take facial expressions as one example; although the perception of facial emotion has been investigated since the 1800s (e.g. Darwin, 1872; Wundt, 1909), much of what we know is founded on research using stimuli (e.g. Ekman & Friesen, 1976) that are essentially unchanged since the earliest investigations, presented within traditional global looking paradigms (e.g. Farroni, Massaccesi, Pividori, & Johnson, 2004; Field, Woodson, Greenberg, & Cohen, 1982; Nelson et al., 1979). Yet even the earliest studies identified the limitations of using static, unresponsive stimuli to investigate phenomena that are inherently dynamic and contingent (Caron, Caron, & Myers, 1985; B. M. Wilcox & Clayton, 1968), and recent work is also finding important differences for dynamic compared to static expressions (Addabbo, Longhi, Marchis, Tagliabue, & Turati, 2018; Heck, Hock, White, Jubran, & Bhatt, 2016, 2017; Soussignan et al., 2017). It is therefore imperative that new methodological and analytical procedures are developed to enable us to harness the potential of new technologies so that we are able to present ecologically relevant stimuli to infants that conserve the essential characteristics of the real-world phenomena they experience in natural environments.

Considering the methodological challenges and opportunities that exist in infant research today, the main aim of this current thesis is to assess the usefulness of gaze-contingent eye tracking alongside several novel and recently-developed analytical techniques as tools to investigate infants' abilities within a controlled lab setting whilst also ensuring our results and subsequent experimental interpretations generalise meaningfully to the real-world.

The potential methodological benefits of gaze-contingency are twofold. Firstly gaze-contingent displays allow researchers to time-lock dynamic presentations to the infant's gaze behaviour, guaranteeing that all infants are attending the same location at the beginning of an important stimulus event. This allows researchers to define a standard analysis time window of eye-tracking data that can be equated across trials and across participants. Secondly, the validity of recording infant looking is improved. The realism of video stimuli presented on the screen is dramatically increased when the on-screen events are *responsive* to the infant (see Meltzoff, 1988; Nielsen, Simcock, & Jenkins, 2008), whilst the active role that an infant plays in triggering events within the experiment also reduces the ambiguity in interpreting the meaningfulness of infant looking.

Alongside the gaze contingent paradigm, this thesis will also present 'Dynamic AOIs'; a novel tool for defining interest regions for dynamic social stimuli. This method automatically locates and tracks regions (such as facial features) as they move and change within video presentations. Other methods useful for analysing eye-tracking data without collapsing across time (mixed effects modelling, cluster permutation analysis, functional data analysis) or space (statistical heatmap analyses) are also demonstrated. Together, these methods equip developmental researchers with the tools required to present naturalistic, dynamic stimuli within infant eye-tracking paradigms.

Chapter 2

Simulating social interactions using gaze-contingent eye-tracking

---

After briefly surveying the historical context of the methods in infant research, the previous chapter highlighted the need to develop methodological strategies to allow us to conduct lab-based investigations of infant development with greater *ecology* whilst also retaining sufficient experimental *control*. Historically, infant research moved from high ecology and low control (e.g. the baby diary home observation approach), to high control and low ecology (e.g. the behaviourist approach followed later by the predominantly lab-based experimental approach in the 70s). This apparent trade-off is at the heart of discussions concerning infant methodology, and is ultimately a question of internal and external validity; are we measuring what we believe we are measuring, and do our findings accurately reflect phenomena in the real-world? To resolve this, we will need to adopt a pragmatic approach; attempting to include maximal ecology, but without compromising on rigour. We must therefore aim for the closest possible representation of real-world phenomena that current methodological and practical constraints allow. Naturally, when conducting lab-based studies we will fall short of maximal ecology. Consequently we then have to make theoretically-driven (and ideally empirically-driven) decisions about which aspects of real-world phenomena are critical and important, and thus are irreducible components that must necessarily be retained in our investigations. These decisions will be heavily dependent on the particular research question being considered. For instance, basic perceptual questions, such as determining the visual acuity of infants, are relatively simple to investigate in a lab setting without being overly concerned with ecological validity. However, how we investigate more complex and context-dependent phenomena in the lab, such as how infants learn language or understand social interactions, will require much more careful consideration. And, as Karen Adolph notes, "the best way to ensure ecological validity in structured lab tasks is to start with a rich description of real-world behaviour" (Adolph, 2020, p. 189).

In the previous chapter we also noted that analysing patterns of infant looking has been an essential methodological strategy for developmental research since the 1960s, and that recent technological innovations in remote eye tracking are enabling researchers to record eye movements with increasing precision, accuracy and automaticity. Yet clearly producing rich descriptions of eye movements contribute little to the *meaningfulness* and *relevance* of infant research if the experimental paradigm, stimuli and analyses are not 'fit for purpose'. Instead, the lasting value of our findings depend firstly on whether the looking behaviours we are recording are a true index of an infant's *internal* world, and they depend secondly on whether the stimuli we present are a true index of an infant's *external* world. Only once these two criteria are reached will our findings hold sufficient validity to be relied upon.

Development within social interactions

One important ecological factor to consider when investigating higher-order infant abilities in the lab is that infant learning and development naturally occurs within the context of social interactions. Humans are intrinsically social beings who are profoundly influenced by the continuously interrelating social structures in which they are embedded (Dunbar & Shultz, 2007). Adult social neuroscience has highlighted that human cognition is fundamentally different within social interactions (see Hari & Kujala, 2009; Schilbach et al., 2013 for reviews), and recent work from developmental social neuroscience suggests that this is also true for infants (e.g. Wass et al., 2018; see Hoehl & Markova, 2018). In fact, learning during development is often facilitated by interactions with adults (Csibra & Gergely, 2009; Vygotsky, 1978), and from birth infants are highly receptive to social information (e.g. Cooper & Aslin, 1990; Farroni, Csibra, Simion, & Johnson, 2002; Farroni, Massaccesi, Pividori, & Johnson, 2004; Farroni, Menon, Rigato, & Johnson, 2007). Yet they are not merely passive ‘absorbers’. Instead infants are active participants within these interactions (Murray & Trevarthen, 1986; Tomasello, Carpenter, Call, Behne, & Moll, 2005). Adult-infant interactions are mutual and bidirectional, and infants quickly develop expectations about their content and timing (Bertin & Striano, 2006; Striano, Henning, & Stahl, 2005, 2006). Infants also learn to use the information provided within social exchanges with increasing sophistication over the first year (see Frith & Frith, 2003; Grossmann & Johnson, 2007).

Firstly dyadic (person-person), and later triadic (person-object-person), interactions are often the contextual units within which infants gather information about the people and objects in their environments (see Csibra & Gergely, 2009; Striano & Reid, 2006). By two-to-three months infants show evidence of playing an active role (Striano et al., 2005; Wolff, 1987) within dyadic interactions, and adults and infants alternate between active and passive roles, communicating through looks, touches, expressions and vocalisations within ‘protoconversations’ (Trevarthen, 1979). These dyads, held together through mutual gaze, are the foundation for later interactions, and are an important vehicle for sharing emotions between adult and infant (Tomasello et al., 2005). Expanding these interactions to include an outside entity (e.g. an object such as a toy) allows a mechanism for infants to learn about the world within the pedagogical context of the social interaction. Triadic interactions require the monitoring of another’s attention to the point of interest (‘joint attention’), to coordinate their visual attention and perform collaborative activities. Although infants show sensitivity to triadic attention by three months (Striano & Stahl, 2005), active engagement within such interactions is thought to develop around nine months (Carpenter, Nagell, & Tomasello, 1998), though the precise age is the subject of some controversy in the literature (Flom & Pick, 2005; Frith & Frith, 2003; Reid & Striano, 2005; see Striano & Reid, 2006). Within the context of these interactions, infants develop in their understanding of the outside world whilst developing their ability to understand the goals and intentions of others (see Tomasello et al., 2005).

Social interactions are therefore an important environmental factor when considering infant development, particularly if we wish to investigate aspects of higher-order functioning which naturally occur within social exchanges. However, reciprocal adult-infant interactions contain complex and fluctuating attentional and behavioural dynamics which are difficult to represent faithfully within a lab setting. Three options therefore present themselves to researchers who wish to investigate social or cognitive development in infancy: firstly, the path most frequently trodden, we could sacrifice ecological validity in favour of experimental control. We argue that paradigms following this approach do not accurately represent the naturally-occurring phenomena frequently experienced by infants in real-world environments, and that by using these methods researchers are inadvertently distancing themselves from the phenomena they are purporting to study. Secondly, we could sacrifice experimental control and ease of analysis in favour of naturalism. Several studies, both classic (e.g. Dunn & Kendrick, 1980) and modern (e.g. Tamis-LeMonda, Kuchirko, Luo, Escobar, & Bornstein, 2017) have taken this approach. Yet the data generated from such studies, which usually involve extensive video recording of infants in their home environment, are often complex, noisy and highly labour-intensive to code and analyse. The control over extraneous variables that can be achieved within a lab setting, as well as the accuracy and automaticity of recording eye movements with modern eye-trackers, are highly beneficial factors to consider, and account for much of the reluctance to adopt this second approach. Thirdly, we can attempt a compromise, and try to retain only the essential aspects of the environment in the lab. Within this thesis, this third approach is favoured. A novel eye tracking paradigm that simulates social interactions within a lab setting will be introduced and applied to current topics of infant socio-cognitive research.

Simulating a social interaction

In order to use modern eye-tracking technology to 'simulate' naturalistic social interactions, we must first identify their essential components. While there are many aspects of interactions which could be considered important (e.g. multimodality, turn-taking, three-dimensionality, actor familiarity, the contextual relevance and congruence of behaviours, the freedom to explore within a naturalistic environment), the method we have developed in our lab (see Keemink, Keshavarzi-Pour, & Kelly, 2019) distils just three: dynamism, communicativeness and contingency. The large body of work exploring the 'still-face effect' in infancy (Tronick, Als, Adamson, Wise, & Brazelton, 1978; see Adamson & Frick, 2003 for a review) provides an ideal illustration of how infants regard interactions where these three aspects are suddenly removed. This usually leads to a sudden 'sobering' of the infant who attempts to re-engage the adult, and then after their efforts prove futile, the infant eventually becomes distressed or withdraws from the interaction. The interactions simulated in this paradigm are only the briefest of exchanges, but with the inclusion of these three critical aspects, they contain the essence of real-world interactions, whilst also allowing researchers to maintain adequate levels of experimental control.

The first basic aspect of naturalistic interactions is that they are *dynamic*. The vast majority of eye tracking studies investigating social and cognitive abilities present static images (see Krumhuber, Kappas, & Manstead, 2013). However outside of the lab infants would rarely (if ever) encounter frozen, unchanging social ‘stimuli’. Particularly when investigating communicative gestures such as facial expressions or gaze cues, static stimuli are insufficient representations of these behaviours. The differences between static and dynamic stimuli should not be understated. Previous work (Bahrick, Gogate, & Ruiz, 2002; see Walker-Andrews & Bahrick, 2001) has demonstrated that five-month-olds were able to discriminate and remember static images of faces, but when dynamic stimuli were used, infants’ attention was diverted away from facial characteristics such that only the actors’ behaviours were encoded (e.g. brushing hair or teeth). This finding has critical methodological implications, as it suggests that claims about how infants perceive and process social information from static stimuli may not generalise to real-world environments. This problem is not unique to infancy. Burton (2013), in his overview of progress within adult face recognition research, described how the field was being led astray by an over-reliance on artificial, static stimuli, reducing face-processing to simple picture matching once natural variance is removed. Recent studies indicate that dynamic and static stimuli are not equivalent as they are processed by infants in qualitatively different ways (Heck et al., 2016; Ichikawa, Kanazawa, & Yamaguchi, 2011; Lewkowicz & Hansen-Tift, 2012; Võ, Smith, Mital, & Henderson, 2012; N. G. Xiao et al., 2015; but see Widen & Russell, 2015). It is therefore essential that dynamism be retained in ecological investigations of infant social cognition.

Although dynamism is a crucial step towards ecological validity (N. G. Xiao et al., 2015), high fidelity social stimuli should arguably not just be dynamic (e.g. counting, blinking or chewing), but also *communicative*. The underlying purpose of social interaction is for one individual to share their knowledge, intentions and emotions with another. Young infants and adults communicate within dyadic ‘proto-conversations’;

where emotions are shared via looks, gestures, touches and vocalisations that occur within turn-taking sequences (Gratier et al., 2015; Legerstee, Markova, & Fisher, 2007; D. N. Stern, Hofer, Haft, & Dore, 1985; Trevarthen, 1979). Later in infancy, triadic interactions involving the adult, infant and another external entity become the vehicle through which knowledge about the world is shared (see Csibra & Gergely, 2009; Tomasello et al., 2005). The essential and connecting ingredient therefore within these early social interactions is that they are communicative, this is the aspect of interactions that is *meaningful*, and arguably this is the aspect which holds the most value for both infants and adults.

Furthermore, realistic interactions are not only dynamic and communicative, but also *contingent*. Communicative behaviours within naturalistic social interactions are not triggered according to a rigid experimental clock, but are guided by the infant's own behaviour and attempts to engage with the adult (Murray & Trevarthen, 1986). Infants are highly sensitive to contingency, and by two to three months this is an expected component of social interactions such that infants can detect even small perturbations (such as a 1 second delay) in the temporal pattern of an adult's social responses (Mcquaid, Bibok, & Carpendale, 2009; Nadel, Carchon, Kervella, Marcelli, & Râ, 1999; Soussignan, Nadel, Canet, & Gerardin, 2006; Striano et al., 2005; Striano, Henning, & Stahl, 2006). For instance, Bigelow and Birch (1999) showed that four- and five-month-olds preferred a social partner who had interacted contingently six days previously over one who hadn't (despite both responding contingently during testing). Infants therefore seek good quality interactions, and prefer individuals who display high levels of responsivity.

Reciprocity (i.e. bi-directional responsivity) is a defining characteristic of social exchanges (Bronfenbrenner, 1977), and it is therefore a requirement that on-screen interactors give temporally concordant responses to infants' behaviours if we are to simulate social dynamics effectively. Previous work that has presented dynamic and communicative social stimuli to infants via video have found differential results, including reduced responsivity, compared to live, 'real-world' experimental paradigms (R. Barr & Hayne, 1999; Hayne, Herbert, & Simcock, 2003; Kretch & Adolph, 2015), suggesting that infants do not 'believe' these interactions are genuine. These findings have been attributed to the fact that infant looking was artificially confined toward a two-dimensional display screen (e.g. Kretch & Adolph, 2015), however, there are also literature examples of infants showing *conserved* imitation of adults when they were displayed using live camera feeds in comparison to pre-recorded videos (Meltzoff, 1988; Nielsen et al., 2008). It is therefore conceivable that it is a lack of contingent reciprocity within screen-based paradigms that is substantially reducing their realism, and not the use of video stimuli *per se*. While investigations of infant responsivity to non-live contingent videos are scarce in the literature, recent investigations (Keemink, Keshavarzi-Pour, et al., 2019; Vernetti et al., 2018) note similar response rates from infants (e.g. smiles and vocalisations) compared to live interactions (e.g. Field, Goldstein, Vega-Lahr, & Porter, 1986). Taken together, these findings suggest that contingent interactions, even those simulated using a display screen, have a high degree of authenticity and believability and are a much closer representation of everyday interactions than static, unresponsive images.

The gaze-contingent paradigm

This thesis presents a novel gaze-contingent (GC) eye-tracking paradigm developed in our lab (Keemink et al., 2019; see also Vernetti et al., 2018), that is designed to simulate brief social exchanges in a lab setting (GC Social Interaction paradigm). Within GC paradigms (Duchowski et al., 2004), the participant's viewing experience is dependent on their eye movements. This approach is gaining popularity in infant research (e.g. Deligianni, Senju, Gergely, & Csibra, 2011; Vernetti, Smith, & Senju, 2017; Wang et al., 2012) because it empowers infants with a novel form of agency, allowing them to *actively* manipulate stimuli presented on the screen. Combined with dynamic stimuli, this methodology is ideal for investigating infants' social development as infants can interact with on-screen actors while their eye movements are being simultaneously recorded. The gaze-contingent manipulation not only enhances realism, but also serves an essential methodological role: using gaze-contingently activated videos ensures all infants are fixating the same location (e.g. the eye region) before the expression response begins. Scan paths for each expression therefore have a standardised start point in both time and space. This facilitates analysis across trials, participants and conditions, but also has the additional benefit of guaranteeing infant attentiveness toward the screen at the start of the analysis time window.

Procedure

Figure 1 presents a template for the GC Social Interaction paradigm. Each sequentially-presented 'interaction' begins with an attention-grabber (e.g. colourful, noisy animated circles) randomly located to the left or right of the screen (counterbalanced across trials). The attention-grabber serves to draw infants' attention away from the subsequent stimulus location, whilst also providing the opportunity to perform a drift correction. Once the attention-grabber is fixated, the first frame of the video clip (e.g. a face) appears centrally positioned on the screen. A predefined, invisible gaze-contingent boundary was placed over a salient region of the stimulus (e.g. the eye region of the face). A fixation landing in the boundary region triggers a 'dynamic social response' (e.g. the video clip of the facial animation). Thus, the on-screen actor provides a dynamic, contingent and communicative response that was initiated by the infant (e.g. looking to the eye region). The onset of the dynamic video also provides the start point for the eye-movement analysis window, where an infant's scanning in response to the actor's behaviour can be recorded.

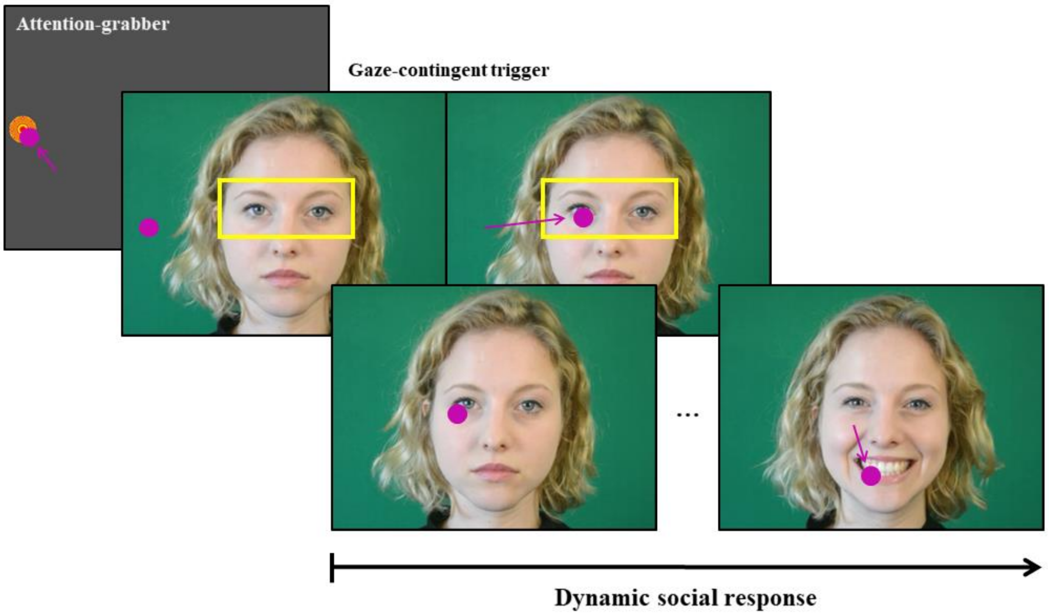

Figure 1. A template for the Gaze-Contingent Social Interaction paradigm.

Each brief interaction begins with an attention grabber (either left or right), and once fixated, the first frame of the video clip appears centrally (infants' hypothetical fixations (circles) and saccades (arrows) are represented here in pink). Once a fixation lands within a predefined, invisible gaze-contingent boundary placed over a salient region of interest (e.g. the eye region, shown in yellow here), a dynamic social response is triggered. In this example, as an infant initiates mutual gaze, the on-screen actor provides a contingent, dynamic and communicative response (e.g. smiling). Eye movements are recorded during the fixed time-window of the dynamic social response (hypothetical face scanning shown in pink).

Participants

Within the Kent Infant Lab (part of the Kent Child Development Unit at the University of Kent) there is a focus on social development across the first year of life. The lab uses a cross-sectional approach recruiting infants from three age groups: six, nine and twelve months. Participants are deemed eligible if they fall within a +/- 14 day age-range at the time of testing, and are healthy, full-term infants that have no siblings with a diagnosis of autistic spectrum disorder (research has been conducted by the lab on infant siblings of autistic children (e.g. Keemink et al., 2019), but these participants are invited only in specific circumstances, and are not included in any of the experiments within this thesis). Infants with any known visual impairments are also considered ineligible for the study and not invited for testing. The majority of participants come from Caucasian families; the ethnic population of Kent is not diverse, with the latest available census data showing that 92.7% of the population categorised themselves as white and just 1.3% of the population categorised themselves as black (Office for National Statistics, 2011). The infant's race was only included within the eligibility criteria if the experiment was specifically investigating the influence of race within social interactions (e.g. Chapter 3). A participant's data was excluded from the final analysis if fixation data were not present for all experimental trials. Such an inclusion criteria is strict (and can lead to data loss), but ensures high data quality, however participants were retained for certain analyses (e.g. mixed effects modelling) which are able to reliably handle missing data.

There were several reasons for selecting the three age groups (six, nine and twelve months) used in this thesis. Firstly, infants achieve an adult-like level of visual acuity (and also several other perceptual abilities) by six months (Braddick & Atkinson, 2011), and can clearly perceive stimuli presented at near distances (e.g. the 60cm presentation distance used in this lab). Secondly, several of the key developmental milestones investigated in this thesis (e.g. adult-like facial expression perception at seven months, other-race preferences at nine months) fall within the 6-to-12 month window. Thirdly, a thorough investigation of the topics covered in this thesis will provide a developmental trajectory across the first year of life, to see how infant functioning changes and stabilises as the perceptual system matures.

Infants were recruited from local clinics and mother-and-baby centres. Parents received an information sheet via email prior to testing and a further verbal briefing on the day before obtaining parental written consent. After testing, infants received a certificate and an age-appropriate toy as a reward. All studies within this thesis received full ethical approval from the School of Psychology's Ethics Committee.

Eye tracking

Infants were fastened in a semi-upright car seat (to minimise head movements) 60cm from a Dell 20-inch display monitor (1024 x 768 pixels). Parents sat nearby, but just behind their infant to minimise distractions. Eye movements from both eyes were recorded (500 Hz) using an SR Research Desktop-Mount EyeLink 1000+ eye tracker with a 25mm lens and 890nm illuminator operating in remote mode (spatial resolution 0.01°, average gaze position error of 0.25°). The EyeLink in remote mode is tolerant to head movements of 22x18x20cm, and has a tracking range of 32° (horizontal) x 25° (vertical). A padded target sticker placed centrally on the forehead served as a reference point for recording eye movements and head distance. Prior to the start of each experiment, a five-point calibration procedure was implemented (Experiment Builder, SR Research, Ontario, CA), using custom 'attention grabbers' to entice looking. Our experimental set-up enables us to monitor the infant's gaze and behavioural responses in real-time (see Figure 2).

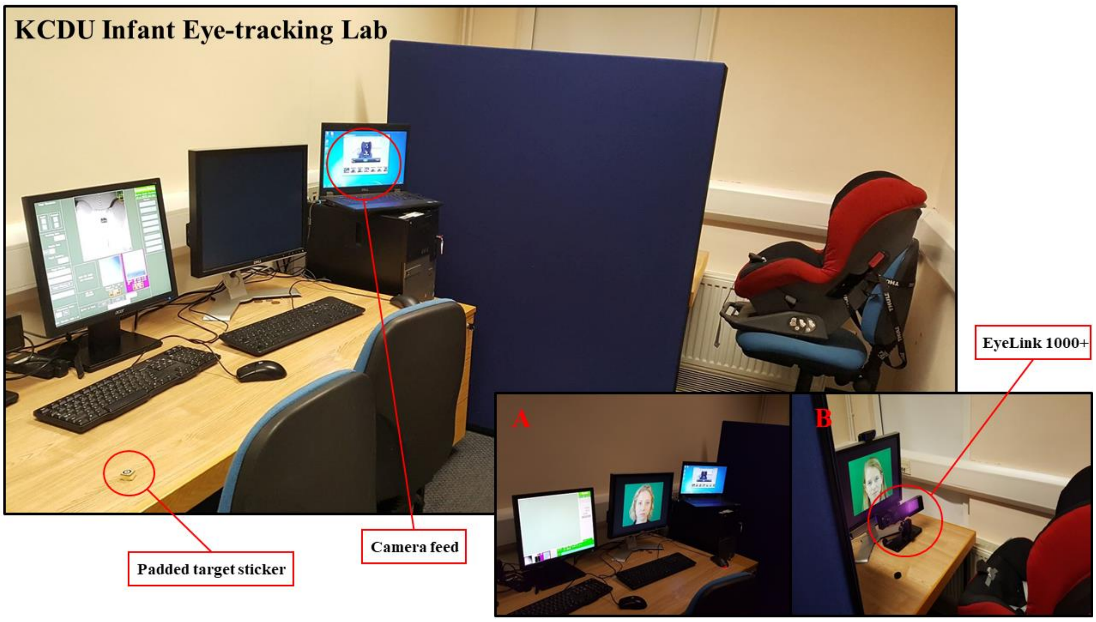

Figure 2. Infant eye tracking at Kent.

Images of the experimental set-up used in infant eye-tracking tasks at the University of Kent. Eye tracking was conducted under dim lighting (A). Infants were fastened into an upright car seat 60cm away from a display screen (B), their eye movements were tracked using the EyeLink 1000+ eye tracker operating in remote mode. A padded target sticker was used as a reference point. A large room divider was used to minimise distractions during testing, but the infant's real-time behaviour and responses could be monitored (and also recorded) using a live camera feed.

Data processing

Eye-tracking data were processed in MATLAB (Mathworks, R2017a). A custom-written velocity-based algorithm was used to identify saccades. Data were initially smoothed by applying a four-sample rolling window that returned a median average. Angular speed was computed based on four samples. Velocity values greater than 1000°/sec were judged to be biologically impossible and were removed from analysis. We set a velocity threshold of 40°/sec, with samples falling below this value identified as potential fixation samples. Time and distance between two potential fixations were calculated. If inter-fixation values were <20ms and <.03° then fixations were merged. All fixations <100ms were removed. Fixations summaries containing discrete fixations were compiled along with their duration, location and sequential number.

Following Holmqvist, Nystrom and Mulvey (2012), and discussions concerning the quality of infant eye tracking data (e.g. Wass, Forssman, & Leppänen, 2014), precision values were calculated as the root mean square (RMS) of sample-to-sample distances within computed fixations. Typical eye tracking precision was calculated separately for each age group using a large independent sample of infants tested within the lab ($N = 172$) and results were as follows: 6 months = 0.62° (*SD* = 0.09°), 9 months = 0.57° (*SD* = 0.07°) and 12 months = 0.58° (*SD* = 0.09°). The precision values presented here are equivalent to values calculated from adult eye tracking experiments (see Holmqvist et al., 2012).

Potential modifications of the GC Social Interaction paradigm

The paradigm described above outlines a basic template for the GC Social Interaction paradigm, but the applications and potential modifications of this method are many and varied. For instance, Keemink and colleagues (2019) used this basic paradigm in an amended form. Firstly, they included two dynamic social responses (a non-engaging head-turn and an engaging smile response) triggered by two different GC AOIs (the eye and mouth regions). Manipulating the type of social response and/or the 'facial trigger' for these responses allows researchers to investigate how infants learn social contingencies, and also if they can adaptively modify face scanning strategies. Furthermore, presenting infants with an actor who consistently provides positive responses alongside an actor who consistently provides negative responses, also opens up the potential to investigate infants' preferences for social behaviours, and also how they encode behaviours they experience within interactions; particularly if combined with a forced-choice visual preference test following these interactions (see Chapters 3 and 4).

Secondly, Keemink and colleagues (2019) also recorded and coded infants' behavioural responses. This step 'completes' the social interaction, and is highly informative when judging the meaningfulness of infant looking behaviours (see also Soussignan et al., 2017; Vernetti et al., 2018). For instance, do looks toward a smiling mouth correlate and co-occur with infants' reciprocal smiles and vocalisations? Such convergent approaches are becoming increasingly valuable as they can help clarify if eye movements reflect higher-level cognitive or lower-level perceptual factors (see Aslin, 2007). Moreover, most modern eye trackers automatically record pupil size alongside gaze location. Whilst there are a number of methodological issues to overcome when conducting pupillometry experiments in infancy (see Hepach & Westermann, 2016; see also Chapter 7), as pupil size is a direct correlate of physiological arousal, it also has considerable potential as a complementary measure alongside eye movements (e.g. Cesana-Arlotti et al., 2018; Jessen, Altvater-Mackensen, & Grossmann, 2016). Other measures such as heart-rate, eye blinks or ERP's have also proved useful when combined with eye tracking to narrow the range of possible interpretations (see Aslin, 2007; Eckstein, Guerra-Carrillo, Miller Singley, & Bunge, 2017; Hoehl & Markova, 2018).

Further modifications of the GC Social Interaction paradigm will depend on the particular research question being considered. For example, the applicability of this method could be widened by presenting multimodal stimuli (e.g. co-occurring speech or affective vocalisations), alternative social responses to those described here (e.g. gaze cueing, facial expressions), alternative GC triggers (e.g. gaze-cued objects triggering responses), or varying the stimulus response rate (e.g. consistent vs inconsistent responders, responders with delayed contingency). With this potential for modification, this paradigm could plausibly be applied to investigate a diverse range of research topics within infant socio-cognitive development.

Analysis

In the previous chapter, some of the challenges of analysing eye movement data for dynamic and interactive stimuli were discussed. Dynamic GC stimulus presentations vary over time, and also vary according to the participant's eye-movements, therefore accurately determining *what* an infant is looking at is an initial challenge that researchers face before further analysis can be conducted. This thesis will present several methods for connecting fixation coordinates to the 'real-world' objects of an infant's gaze. Specifically, we use traditional AOI regions, novel 'dynamic AOI' regions, and fixation 'heatmaps' to determine gaze locations. These methods will be presented in detail within the empirical chapters that follow (Chapters 3 – 7), and the general discussion chapter (Chapter 8) will compare their relative advantages and disadvantages.

Once fixation data have been coded with an AOI 'tag', they can then be collated, summarised and meaningful results can be extracted. Vast amounts of data are generated within this paradigm due to the detailed recording of eye movements across time and spatial location. Therefore care must be taken not to lose this richness, but also not to misrepresent our findings (see Yu, Yurovsky, & Xu, 2012). For instance, by manipulating multiple and arbitrarily defined AOIs or time bins there is the potential to alter results dramatically (D. J. Barr, 2008; Caldara & Miellet, 2011; Hessels et al., 2016). With this caution in mind, several analysis techniques will be implemented that allow researchers to investigate continuous data, not collapsed into time bins (e.g. mixed effects modelling, functional data analysis) or AOIs (e.g. ANOVA heatmap analysis). Again, these methods will be presented in greater detail within the empirical chapters, and considered further within the discussion chapter. Within infant research, there is a clear need for powerful analytical tools to aid researchers as they attempt to draw meaningful conclusions from the voluminous, and often complex, data routinely generated within ecological eye-tracking paradigms. Presenting potential analysis techniques alongside the GC Social Interaction paradigm is a primary aim of the current thesis.

Theoretical contribution

While the focus of this thesis is primarily methodological, and the empirical chapters will serve as 'worked examples' where these methods will be demonstrated, they will also provide a novel theoretical contribution to the infant literature. The research focuses on perceptual and behavioural cues that are naturally relevant to infants' interactions in their real-world environments. Firstly infant race and behavioural preferences (Chapter's 3 and 4) will be investigated, followed by how infants' perceive facial expressions (Chapter's 5, 6 and 7). It is important that these topics are investigated within the context of social interactions as our current theoretical understanding of these areas is founded primarily on research presenting static, unresponsive stimuli which may not necessarily generalise to real-world environments. For instance, three-month-old infants' have been shown to look longer ('prefer') images of own-race faces, and nine-month-old infants look longer at images of other-race preferences (Kelly et al., 2005; Liu et al., 2015). But would infants show such preferences in a naturalistic context where salient social information other than race is also available? To give another example, using static images newborns have been shown to have a preference for happy expressions which later transitions to a preference for fearful expressions at seven months of age (Farroni et al., 2007; Peltola, Hietanen, Forssman, & Leppänen, 2013). Facial expressions however are inherently dynamic, communicative and contingent social phenomena and infant preferences, scanning and understanding of dynamic, interactive expression stimuli might be entirely different. These questions are considered in detail within the empirical chapters alongside methodological issues. Given the novelty of the methods, the approach within this thesis will be largely exploratory, however any specific hypotheses within each topic will be clearly noted in that chapter's introduction.

Empirical 'worked examples'

The subsequent empirical chapters in this thesis will provide experimental 'worked examples' of the GC Social Interaction paradigm tailored to address specific research questions. **Chapter 3** will investigate race within simulated social interactions, and examine if these brief social exchanges are sufficient to modify infants' *a priori* preferences for own- or other-race interactors. In this version of the GC Social Interaction paradigm there will be two on-screen interactors during a familiarisation phase, followed by a traditional forced-choice visual preference test. **Chapter 4** extends the paradigm used in Chapter 3 to consider if infants show preferences for positively, socially-engaging behaviours over disinterested, non-engaging behaviours. The remaining empirical chapters will then focus on facial expression scanning and its development in infancy. In **Chapter 5**, infant responses to dynamic and static expression stimuli (happy, sadness, surprise, fear, anger and disgust) across six actors are compared, while **Chapter 6** explores developmental differences in dynamic facial expression scanning. Finally, **Chapter 7** explores the feasibility of using pupillometry within the GC Social Interaction paradigm, comparing the pupil size values from the experiment in Chapter 5 to that of another pupillometry expression experiment with increased luminance control. Several methodological topics are considered within these chapters, including the differences between dynamic and static stimuli, the utility of visual preference procedures within interactive paradigms and the usefulness of pupillometry as a companion method within infant eye tracking.

Chapter 3

Preferences for race and social behaviour within simulated social interactions

---

Abstract

Previous work presenting static faces in paired-comparison tasks suggests that infants' show an own-race preference by three months, which later transitions to an other-race preference by nine months. Nevertheless, the question remains as to whether infant race preferences emerge in naturalistic, 'real-world' settings, particularly as the differential processing of race can be attenuated in children and adults when additional social information is also presented. Experiment 1 of this chapter first replicates and extends the canonical pattern of race preferences found in the literature using dynamic stimuli. Experiment 2 then demonstrates that these race preferences are removed (six- and twelve-month-olds) or attenuated (nine-month-olds) if infants have the opportunity to 'interact' with the on-screen actors during a familiarisation phase prior to the preference test. In Experiment 3, the own- and other-race actors behave differently during familiarisation; one giving 'socially-engaging' responses, the other giving 'non-engaging' responses. In this third experiment, no preferences for either race or social behaviour emerged during the preference test. These results are discussed with reference to infants' patterns of social scanning and suggest that race preferences for static stimuli may not generalise to real-world environments.

Introduction

Infants use social information to learn about the identities and intentions of others showing socio-perceptual preferences based on certain physical characteristics such as attractiveness (A. M. Slater et al., 1998), age (Macchi Cassia, Bulf, Quadrelli, & Proietti, 2014), gender (Quinn, Yahr, Kuhn, Slater, & Pascalis, 2002) and race (Kelly et al., 2005) by 3 months of age. These perceptual preferences are based on early experiences (Bar-Haim, Ziv, Lamy, & Hodes, 2006) and are learned alongside additional sources of social information within communicative interactions. For instance, facial expressions (Farroni et al., 2007), eye gaze cues (Farroni et al., 2004) and interpersonal responsiveness (Striano et al., 2005; Striano, Henning, & Stahl, 2006) are all highly salient socio-behavioural cues that may compete with, and attenuate, an infant's preference for perceptual characteristics.

One example of socio-perceptual preference is the other-race effect (ORE), which refers to the differential processing of own- and other-race faces (see Hugenberg, Young, Bernstein, & Sacco, 2010 for a review). The ORE has been demonstrated in infancy, with the ability to recognise other-race faces declining as own-race face processing expertise is acquired (Kelly et al., 2009, 2007; Sugden & Marquis, 2017). Differential spontaneous visual preference (i.e. preferences not generated by internal goals or other stimulus characteristics) is another aspect of the ORE, and a preference for own-race faces emerges at three months (Fassbender, Teubert, & Lohaus, 2016; Kelly et al., 2005; Liu et al., 2015), driven by the asymmetrical experience of own- and other-race faces in the early social environment (Anzures et al., 2012; Gaither, Pauker, & Johnson, 2012; Sangrigoli, Pallier, Argenti, Ventureyra, & de Schonen, 2005; Singarajah et al., 2017). This preference declines and reverses across the first year, with an other-race preference present by nine months of age (Fassbender et al., 2016; Liu et al., 2015). This developmental shift from a 'familiarity' preference toward a 'novelty' preference is attributed to the infant's cumulative exposure to own-race faces and a more general tendency to seek novelty rather than familiarity at this age (Hunter & Ames, 1988; see Quinn, Lee, & Pascalis, 2019).

Most previous studies (Fassbender et al., 2016; Kelly et al., 2005; Liu et al., 2015) have used still images of neutral faces presented side-by-side for a fixed period of time. However outside of the lab infants would rarely (if ever) encounter such a situation (see Chapter 2). Consequently, it is possible that isolating race in this contrived fashion may artificially inflate its importance. Previous studies using dynamic stimuli (sequential videos of actors counting) have found infants do not show greater overall looking time (i.e. interest) for either own- or other-race faces (Liu et al., 2011; Wheeler et al., 2011). Socio-perceptual characteristics, such as race, may therefore become less salient to infants within dynamic presentations. Socio-behavioural information is also highly salient to infants (D'Entremont, Hains, & Muir, 1997; Farroni et al., 2004; Senju, Csibra, & Johnson, 2008), and recent research (Pickron, Fava, & Scott, 2017; N. G. Xiao et al., 2018) has shown that seven-month-olds are more likely to use gaze cues from an own-race adult compared to an other-race adult to anticipate the location of objects or events, suggesting that socio-perceptual and socio-behavioural cues can interact within development.

In the current study the GC Social Interaction paradigm (see Chapter 2) was used to simulate brief social exchanges in which infants can trigger socio-behavioural responses from the on-screen actor by engaging them in eye contact (see Keemink, Keshavarzi-Pour, & Kelly, 2019). Previous work in adults and older children has suggested that although race might be encoded consistently (see Cosmides, Tooby, & Kurzban, 2003 for a review), its importance can be negated in the presence of additional social information. In adults, Kurzban, Tooby and Cosmides (2001) demonstrated that when a rival visual cue (a coloured sports team vest) was made salient, categorical judgments were made using this cue rather than race. In children, Kinzler et al. (2009) found that five-year-olds prefer own-race children when choosing friends from on-screen images, but when provided with additional verbal cues, preferences shifted to accent rather than race. Additionally, while five-year-olds show explicit social preferences based on race, 10-month-old infants do not (Kinzler & Spelke, 2011). In infancy a recent study (Minar & Lewkowicz, 2017) has shown that when provided with dynamic face stimuli accompanied by concordant speech vocalisations, older infants were able to ‘overcome’ perceptual narrowing and recognise other-race faces. Accordingly, it is hypothesised that the additional social information provided within the simulated interactions will also be sufficient to remove spontaneous race preferences in infants.

Previous work has also found that infants show differential scanning of own- and other-race faces within dynamic presentations, with Caucasian infants fixating the eye-region more for own-race faces, and the mouth more for other-race faces (Pickron et al., 2017; Wheeler et al., 2011; W. S. Xiao, Quinn, Pascalis, & Lee, 2014; W. S. Xiao, Xiao, Quinn, Anzures, & Lee, 2013). Whilst it is expected that overall race preferences will be removed following simulated social interactions, differences in face scanning driven by physiognomic differences between races may be found.

It is also likely that the valence of the socio-behavioural response might influence infant preferences. Positive social feedback has been shown to be rewarding in adults (O’Doherty et al., 2003) and in infants (Farroni et al., 2007), whilst non-engaging responses (such as non-referential head-turns away, closing eyes or still-face) are less so (Scaife & Bruner, 1975), and can trigger reduced affect or distress (Bazhenova, Stroganova, Doussard-Roosevelt, Posikera, & Porges, 2007; Hains & Muir, 1996; Tronick et al., 1978). Recent work using social interactions simulated by gaze-contingent eye-tracking has found both young children and adults learned the association between a gaze-contingent social cue and its reward faster when the cue was engaging rather than non-engaging (Vernetti et al., 2017), while infants and toddlers provide a greater frequency of positive behavioural responses for socially-engaging (head turn toward, greeting, or smile), compared to non-engaging (head turn away, moan or frown) social cues, suggesting that engaging social interactions possess intrinsic reward value (Keemink, Keshavarzi-Pour, et al., 2019; Vernetti et al., 2018).

To investigate infant race preferences following simulated social interactions three experiments were conducted. The first experiment used a limited sample to confirm that well-documented spontaneous race preferences (Kelly et al., 2005; Liu et al., 2015) would extend to novel dynamic stimuli. Then, in a second experiment, this paradigm was extended to include a *familiarisation* phase where infants could experience brief interactions (via the GC Social Interaction paradigm) with the two on-screen actors who give identical social responses to see if race preferences persist in a more naturalistic context. In a third experiment, own- and other-race actors respond differently during *familiarisation*, one engagingly (mutual-gaze and smile) and one non-engagingly (head turn away), and thus pitting socio-behavioural and socio-perceptual cues against one another.

Hypotheses

It is predicted firstly that race preferences reported for static faces will be replicated in the dynamic (blinking, but unresponsive) stimuli (Experiment 1). More specifically that six-month-olds will show an own-race preference, while nine-month-olds will show an other-race preference. It is also predict that this other-race preference will extend to twelve-month-old infants. Secondly, it is predicted that race preferences will be nulled if infants have the opportunity to ‘interact’ with the on-screen stimuli during a *familiarisation* period prior to the preference test, where on-screen actors provide dynamic and contingent socio-behavioural cues (Experiment 2). Thirdly, it is predicted that infants will show a visual preference for an actor who provides positive engaging responses (smiles) during *familiarisation* over an actor that provides consistently non-engaging responses, regardless of race (Experiment 3).

Experiment 1

Method

Forty-two Caucasian infants were recruited from three age groups; six (*M* = 189.00, *SD* = 3.32, Range = 184 – 195), nine (*M* = 275.92, *SD* = 5.25, Range = 268 – 285) and twelve months (*M* = 366.85, *SD* = 4.36, Range = 360 – 376), with 14 infants in each age group (6M females = 8, 9M females = 8, 12M females = 7; see Chapter 2 for more information). Prior to testing, two colour videos were recorded, one with an African actor, the other with a Caucasian actor, using a Nikon D5200 digital camera, and cropped to 540 x 768 pixels. In both videos, the actor faced forward and maintained neutral affect. Blinking was synchronised across both videos. Both actors wore an identical red t-shirt and were seated in front of a uniform green background. Faces were matched for attractiveness and distinctiveness on a 1-10 scale as judged by 10 adult observers (see Kelly et al., 2007 for a similar approach).

On the day of testing, infants' eye movements were recorded (see Chapter 2 for details) after being presented with four 10-second *preference* trials, in which both actors were presented side-by-side, with their left-right spatial position counterbalanced across trials.

Results

A 3 (Age: 6, 9 and 12) x 4 (Trial) ANOVA for proportional fixation durations toward the own-race face was conducted. Mauchly's test indicated a violation of sphericity, and a Greenhouse-Geisser correction was applied. This analysis showed no significant effect for Trial ($F(2.39,93.21) = .141$, $p = .901$, $\eta_p^2 = .004$), or Trial x Age interaction ($F(4.78,93.21) = .141$, $p = .921$, $\eta_p^2 = .468$), but did yield a significant main effect for Age ($F(2,39) = 6.270$, $p = .004$, $\eta_p^2 = .243$). Post-hoc tests (Bonferroni) showed significantly higher own-race looking in six-month-olds (*M* = 54.25%) compared to both nine (*M* = 44.57%, *p* =.009) and twelve-month-olds (*M* = 45.21%, *p* =.016). Planned comparisons were also conducted to compare race preferences against chance (50%). These revealed six-month-olds showed a significant own-race familiarity preference (*M* = 44.57%, $t(13) = 2.558$, $p = .024$, $d = 1.42$), nine-month-olds a showed marginal other-race novelty preference (*M* = 44.57%, $t(13) = 2.070$, $p = .059$, $d = 1.15$) and twelve-month-olds a significant other-race preference (*M* = 44.57%, $t(13) = 2.290$, $p = .039$, $d = 1.27$). Experiment 1 thus successfully replicated and extended previous race preference findings (Liu et al., 2015) with dynamic stimuli (see Figure 1a).

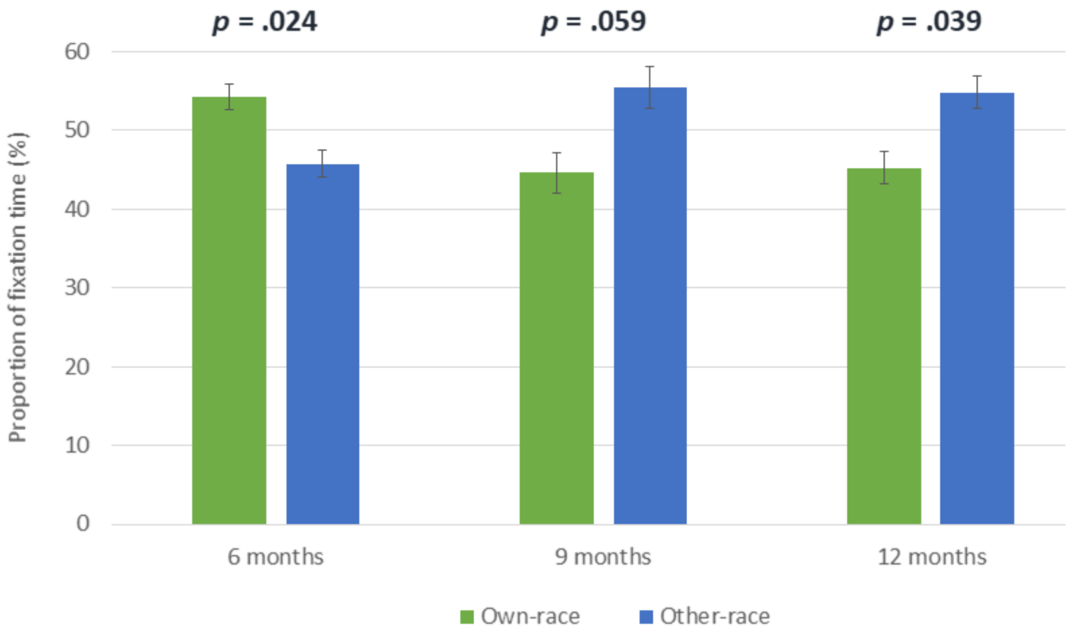

Figure 1. Race preferences in Experiment 1.

Mean proportional fixation durations towards the own- and other-race actors in six-, nine- and twelve-month-old infants from Experiment 1. Error bars represent standard error, and p values are two-tailed.

Discussion

In accordance with previous work using static stimuli (Kelly et al., 2005; Liu et al., 2015), Experiment 1 found that spontaneous race preferences emerge in a paired presentation using dynamic (blinking) stimuli. This replicated the pattern of Liu et al (2015) who showed a three-month-old own-race preference transitioning to a nine-month-old other-race preference. This also extends their findings to show twelve-month-olds also demonstrate spontaneous novelty preferences for other-race faces.

In 'real-world' environments, however, infants would not encounter individuals who were still-faced and unresponsive (see Walker-Andrews & Bahrick, 2001). In a second experiment, an interactive *familiarisation* phase was included prior to the first experiment's *preference* test to see if spontaneous race preferences would persist following brief exposure to socially-interactive actors.

Experiment 2

Method

Participants

Experiment 2 uses a 3 (Age: 6, 9 and 12 months; between-subjects) x 2 (Condition: 'Socially-engaging', 'Non-engaging'; between-subjects) x 2 (Race: Own-race, Other-race; within-subjects) design. Ninety-three infants were included in the final analysis of this experiment, and were randomly allocated to one of the two conditions (see Table 1). This equates to 35 (6 month), 28 (9 month) and 30 (12 month) infants within their respective age groups, with 48 (socially-engaging) and 45 (non-engaging) infants in each condition. This sample size is comparable to previous work in this field (e.g. Fassbender et al., 2016; Liu et al., 2015). An additional thirteen participants were excluded for not possessing data for all trials.

Table 1. Participant age, condition and gender information for Experiment 2

Age (months)	Age M (SD) (days)	Socially-engaging	Non-engaging	Male	Female
6	183.06 (15.23)	16	19	16	19
9	262.64 (11.97)	15	13	17	11
12	360.41 (12.65)	17	13	16	14

Stimuli

Preference phase stimuli were identical to Experiment 1, however, using the same video recording and editing procedures described above, eight additional videos of the two actors were created for the *familiarisation* phase (see Figure 2). In these videos, each actor was recorded performing two 'socially engaging' (SE) behaviours (open- and closed-smile), and two 'socially non-engaging' (NE) behaviours (turn away left and turn away right). These behaviours were selected to incorporate both positively and negatively-valenced interactions within the experiment (see Keemink et al., 2019). Negative expressions (e.g. fear or anger) have been previously used as a component of negative or non-engaging stimuli (Soken & Pick, 1999; Vernetti et al., 2017), however a symbol of disinterest (head-turns away) was chosen as negative expressions also have the potential to be engaging and informative (Peltola, Leppänen, Mäki, & Hietanen, 2009; Vaish, Grossmann, & Woodward, 2008).

Procedure

Familiarisation phase. The GC Social Interaction paradigm (see Chapter 2) was used to present a sequence of eight brief social interactions (see Keemink et al., 2019; Vernetti et al., 2018). Each interaction trial began with an attention-grabber randomly located to the left or right of the screen (counterbalanced across trials). Once the attention-grabber was fixated, the first frame of the video clip (i.e. the face) appeared centrally positioned in the screen (see Figure 2). A predefined, invisible gaze-contingent boundary was placed over the eye region of the face. A fixation (with a minimum duration of 100 msecs) landing in the boundary region triggered the facial animation. Thus, an engaging or non-engaging 'response' was gaze-contingently initiated by the infant. Each trial lasted five seconds. The African and Caucasian actors appeared a total of four times each and in alternate trials. The order of race presentation was counterbalanced across infants. The African and Caucasian actors responded with smiles (open x 2 and closed x 2) in the SE condition and with head-turning away (left x 2 and right x 2) in the NE condition.

Preference phase. Following completion of the *familiarisation* trials, infants were presented with two ten-second *preference* trials, in which both actors were presented side-by-side, with their left-right spatial position reversed across trials. Eye movements were recorded using the same eye-tracking apparatus and method as the first experiment.

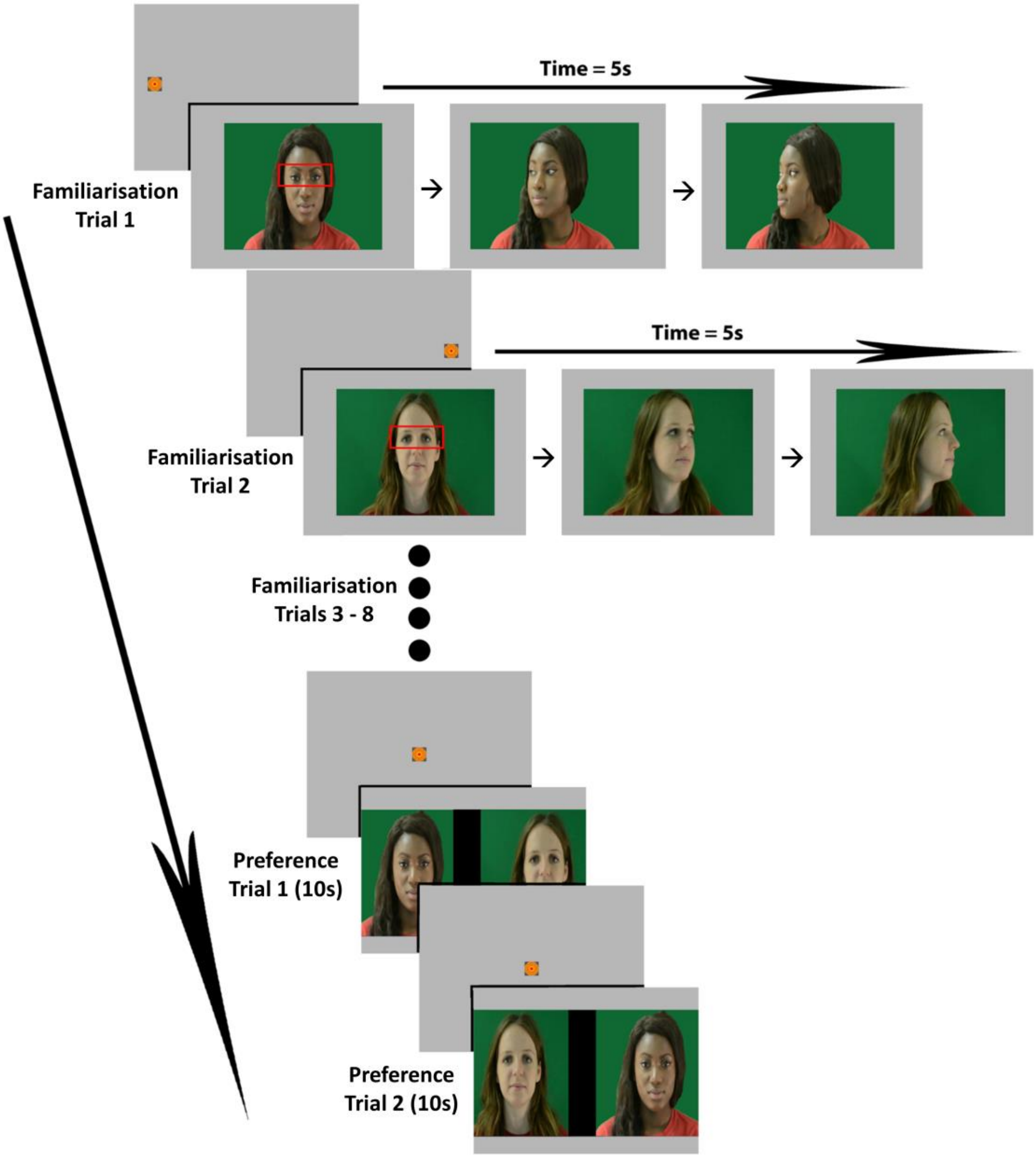

Figure 2. A flow diagram depicting the experimental procedure.

The familiarisation phase began with an attention-grabber (left or right), and once fixated, the first frame of the video clip appeared centrally. Once a fixation landed within a predefined, invisible gaze-contingent boundary placed over the eye region (shown in red here), the facial animation was triggered; either a ‘socially-engaging’ smile or a ‘non-engaging’ head-turn (shown here), depending on the condition. There were a total of eight (5-second) familiarisation trials, with African and Caucasian actors appearing four times each. Two (10-second) preference trials followed directly. Actors were presented side-by-side, with neutral affect and natural blinking. The actor’s spatial position was reversed across trials.

Data processing and visualisation

AOIs. Fixation data were computed for each stimulus (see Chapter 2 for more details). Additionally, to analyse patterns of looking within both experimental phases, fixed AOI regions were defined (Figure 3) and proportionate fixation durations within these regions were compared. To investigate face scanning, eye-region, mouth-region and gaze-region AOIs were created. Whilst the eye region is salient within both conditions, the mouth is highly salient only within the socially-engaging (smiling) condition, and the gaze regions are highly salient only within the non-engaging (turning away) condition.

Heatmaps. Fixation density heatmaps (Figures 5 and 7) were also produced to visualise infant looking patterns. Heatmaps were produced in MATLAB (MathWorks, R2018a). The process involved first producing descriptive fixation maps by summing fixation durations for each pixel ‘coordinate’. These fixation distributions were then smoothed with a Gaussian kernel and computed into z-scores. The resulting z-score maps were cropped to include only AOI-region fixations (i.e. face and gaze), controlling for the inflation of z-scores that can occur when the stimulus background is included.

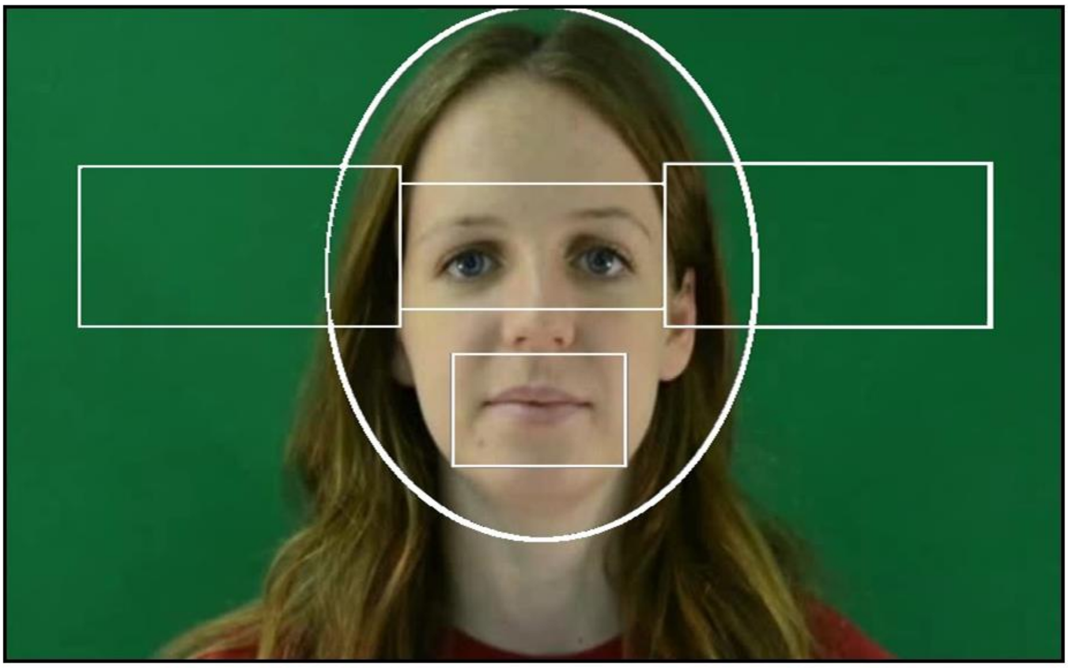

Figure 3. Areas of Interest.

The areas of interest (AOIs) used during the familiarisation phase are shown here in white; including eye (190 x 110 pixels), mouth (115 x 100 pixels), face (Ellipse: 280 x 450 pixels), and both left and right gaze regions (220 x 140 pixels each). AOI dimensions and positions were adjusted (maximum 10 pixels per dimension) to account for differences in facial morphology.

Results: Race preferences

Familiarisation phase: Eye region onset

Infants rapidly triggered the gaze-contingent response during the *familiarisation* phase, with average first eye region fixation onset across familiarisation trials under one second (*M* = .93s, *SD* = .59s). Eye region onset did not significantly differ across stimulus race (F (1,91) = .054, p = .817, η_p^2 = .001), but did differ across participant age (F(2,91) = 3.43, p = .037, η_p^2 = .070). Post-hoc tests (Bonferroni) indicate twelve-month-olds (*M* = 1.10s) were significantly slower to fixate the eye-region compared to nine-month-olds

(M = .73s, p = .031). This age effect may reflect known reductions in eye-region looking for twelve-month-old infants (Tenenbaum et al., 2013).

Familiarisation phase: Duration

A 3 (Age: 6, 9 or 12 months) x 2 (FaceRace: Caucasian or African) x 2 (Condition: SE or NE) x 4 (Trial) mixed ANOVA conducted on looking time duration showed no significant main effect of FaceRace ($F(1,87) = .465$, $p = .497$, $ŋ_p^2 = .005$), or Age ($F(2,87) = 1.214$, $p = .302$, $ŋ_p^2 = .027$). However, there was a significant main effect of Trial ($F(3,261) = 7.774$, $p < .001$, $ŋ_p^2 = .082$) with significantly reduced looking durations for the third (M = 3.57s, $p < .001$) and fourth (M = 3.70s, p = .001) trials compared to the first trial (M = 4.20s). Infants show decreased attention over time to repetitively presented stimuli (Fantz, 1964). Additionally, there was a significant main effect of Condition ($F(1,87) = 4.192$, $p = .044$, $ŋ_p^2 = .046$), indicating longer looking toward socially-engaging stimuli (M = 4.03s) relative to non-engaging stimuli (M = 3.65s). All interactions were non-significant (all $p > .15$).

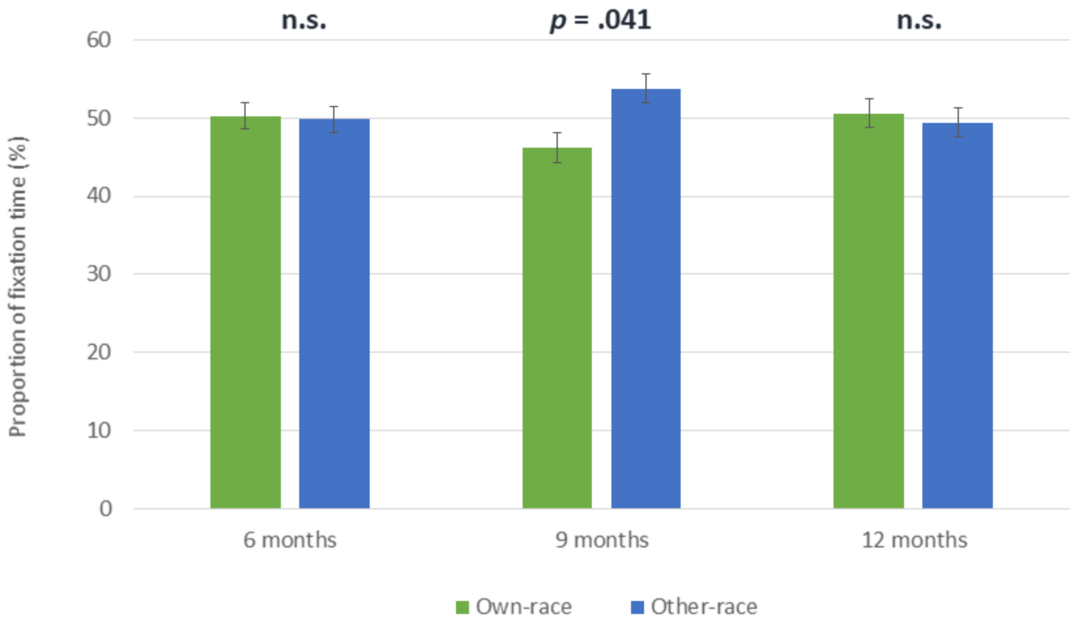

Figure 4. Race preferences in Experiment 2.

Mean proportional fixation durations towards the own- and other-race actors in six-, nine- and twelve-month-old infants from the preference phase in experiment 2. Error bars represent standard error, and p values are two-tailed.

Preference phase

A 2 (FaceRace: Caucasian or African) x 2 (Condition: SE or NE) x 3 (Age: 6, 9 or 12 months) ANOVA was conducted for stimulus looking time during the *preference* phase, finding no main effect for Age ($F(2,174) = 1.568$, $p = .211$, $ŋ_p^2 = .018$), Condition ($F(1,174) = .196$, $p = .659$, $ŋ_p^2 = .001$) or FaceRace ($F(1,174) = .022$, $p = .884$, $ŋ_p^2 < .001$). Furthermore, all interactions were non-significant (all $p > .09$). Proportional looking durations (% of total) toward the own-race stimulus were analysed via a 2 (Condition) x 3 (Age) ANOVA finding no significant effects of Age ($F(2,87) = 1.464$, $p = .237$, $ŋ_p^2 = .033$) or Condition ($F(1,87) = 1.905$, $p = .171$, $ŋ_p^2 = .021$), and the interaction was non-significant ($F(2,87) = .293$, $p = .746$, $ŋ_p^2 = .007$). Proportional looking to the own-race stimulus (OwnRace *M* = 49.15%; OtherRace *M* = 50.85%, *SD* = 9.97%) was also compared against chance (50%), yielding no significant race preferences overall (*t*(92) = -.825, *p* = .842, *d* = .172), or within six (*M* = 50.23%, *t*(32) = .123, *p* = .903, *d* = .044) and twelve-month (*M* = 50.57%, *t*(31) = .332, *p* = .742, *d* = .119) age groups, though nine-month-olds did show significantly higher proportional looking toward the other-race stimulus (OwnRace *M* = 46.24%, *t*(27) = -2.146, *p* = .041, *d* = .826). Race preferences were therefore removed following *familiarisation* for six- and twelve-month-olds, but were somewhat more robust within the nine-month age group (Figure 4).

Results: Social interest and scanning behaviour

In this experiment, conditions differed according to the way the on-screen actors responded to an eye-region fixation during a *familiarisation* trial; with infants receiving either a turning away or smiling response. This manipulation necessarily produced different patterns of face-scanning in participants. For instance, mouth region dwell time (time fixating the mouth region AOI as a proportion of total face region looking; see Figure 3) was higher for the SE condition (SE Mouth *M* = 20.59%, NE Mouth *M* = 5.32%; *t*(69.60) = 5.331, *p* < .001, *d* = 1.12), reflecting higher looking toward the smiling mouth. Also, looking durations toward gaze regions were higher during the NE condition (SE Gaze *M* = 3.15%, NE Gaze *M* = 26.40%; *t*(48.01) = -9.830, *p* < .001, *d* = 2.06), reflecting the tendency for infants to follow the gaze as actors turned away. Infants were highly attentive to these social behaviours, and the majority exhibited at least one instance of ‘smile-looking’ (mouth-region fixating during smile response; *N* = 42, 87.50% of SE condition) or ‘gaze-following’ (congruent gaze-region fixating during head turn response; *N* = 40, 88.89% of NE condition) across *familiarisation* (see Figure 7 for heatmaps depicting social scanning).

AOIs: Familiarisation phase

To determine if infants showed differential scanning of own- and other-race faces (see Liu et al. 2011; Wheeler et al. 2011), proportional fixation durations toward three AOI regions (Eyes, Mouth and Gaze, each as a % of total, see Figure 3) were compared using a 2 (FaceRace) x 2 (Condition) x 3 (AOI) x 3 (Age) mixed ANOVA. This analysis found no significant effect of Condition ($F(1,261) = .003$, $p = .955$, $ŋ_p^2 < .001$) or Age ($F(2,261) = 1.811$, $p = .165$, $ŋ_p^2 = .014$), but yielded a significant main effect of FaceRace ($F(1,261) = 5.164$, $p = .019$, $ŋ_p^2 = .021$), and a significant FaceRace x AOI interaction ($F(2,261) = 9.603$, $p < .001$, $ŋ_p^2 = .069$). Post-hoc comparisons uncovered significantly higher eye-region dwell time for the own-race face ($M = 37.28\%$), compared to the other-race face ($M = 29.36\%$; $F(1,261) = 23.714$, $p < .001$, $ŋ_p^2 = .083$) while mouth and gaze region dwell time did not significantly differ between races (both $p > .30$, see Table 2).

Table 2. Differences in *familiarisation* phase AOI looking (% of total fixation duration) produced in response to race and socio-behavioural cues (estimated marginal means)

	Socially-engaging Condition		**Non-engaging Condition**	
	Own Race	**Other Race**	**Own Race**	**Other Race**
Eyes	40.46	35.40	34.11	22.73
Mouth	18.53	20.91	5.44	6.38
Gaze	3.33	2.92	26.64	26.22

The ANOVA also produced a significant main effect of AOI ($F(2,261) = 50.393$, $p < .001$, $ŋ_p^2 = .279$), a reflection of known preferences for the eye-region (Haith et al., 1977). A significant AOI x Condition interaction ($F(2,261) = 40.823$, $p < .001$, $ŋ_p^2 = .238$)

was also found, reflecting clear differences in social scanning behaviours produced between conditions (see Table 2). In particular, there was an increased interest in the eye (p = .002) and mouth (p < .001) regions within the socially-engaging condition, but increased interest in gaze-regions during the non-engaging condition (p < .001). These large and general differences between conditions broadly correspond to the patterns of smile-looking (SE condition) and gaze-following (NE condition) that were noticed from many participants in the lab (see Chapter 2 Figure 2).

While there was no significant main effect of Age, there was a significant Age x AOI interaction (F(4,261) = 2.719, p = .030, ŋ$_p^2$ = .040). Post-hoc analyses indicate that there was a developmental progression of increased mouth-looking in older infants (6M mouth M = 7.55%, 9M mouth M = 11.92%, 12M mouth M = 18.96%; F(2,261) = 4.492, p = .012, ŋ$_p^2$ = .033), with twelve-month-olds showing significantly greater looking toward the mouth region compared to six-month-olds (p = .009), which replicates previous findings (Tenenbaum et al., 2013) and is thought to reflect an adaptive allocation of visual attention to facilitate speech development (Lewkowicz & Hansen-Tift, 2012). All other interactions were non-significant (all p > .13).

AOIs: Preference phase

A 2 (FaceRace) x 2 (Condition) x 3 (Age) x 2 (AOI: Eyes and Mouth) ANOVA investigating proportionate looking to facial features (% of Face) during the *preference* trials yielded no main effects of Condition (F(1,174) = .001, p = .981, ŋ$_p^2$ < .001), Age (F(2,174) = .055, p = .946, ŋ$_p^2$ = .001), or FaceRace (F(1,174) = 1.078, p = .301, ŋ$_p^2$ = .006), however there was a significant main effect of AOI (F(1,174) = 216.407, p < .001, ŋ$_p^2$ = .554) reflecting greater looking to the eyes (M = 32.12%) compared to the mouth (M = 7.33%). And a significant AOI x Age interaction (F(2,174) = 3.301, p = .039, ŋ$_p^2$ =

.037), reflecting greater looking to the mouth in twelve-month-olds (6M mouth *M* = 5.56%, 9M mouth *M* = 5.95%, 12M *M* = 10.46%). There were also significant FaceRace x Condition (*F*(1,174) = 4.070, *p* = .045, $ŋ_p^2$ = .023), and FaceRace x Condition x Age (*F*(2,174) = 3.852, *p* = .023, $ŋ_p^2$ = .042) interactions. Post-hoc (Bonferroni) comparisons indicate that within the NE condition there was greater looking toward the facial features of the other-race face (OwnRace *M* = 17.89%, OtherRace *M* = 21.51%, *p* = .035), particularly within the nine-month-old age group (9M OwnRace *M* = 16.41%, 9M OtherRace *M* = 22.83%, *p* = .029).

Table 3. Differences in *preference* phase AOI looking (% of face fixations) produced for each condition (estimated marginal means)

	Socially-engaging Condition		**Non-engaging Condition**	
	Own Race	**Other Race**	**Own Race**	**Other Race**
Eyes	33.79	29.03	30.32	35.34
Mouth	6.85	9.29	5.46	7.69

There was also a significant three-way interaction between FaceRace, Condition and AOI (*F*(1,174) = 4.443, *p* = .036, $ŋ_p^2$ = .025; all other interactions *p* > .11). Post-hoc comparisons indicate significant race differences in eye-region looking, with greater own-race eye-region looking in the SE condition (*p* = .043), but greater other-race eye-region looking in the NE condition (*p* = .038; see Table 3). Increased other-race eye-looking in the non-engaging condition may be explained by the *overall* increase in nine-month-old looking for other-race facial features within this condition, not differences in face-region scanning *per se* (see Table 3). Descriptive heatmaps in Figure 5 show similar looking to the eye-region across both conditions and races. Nevertheless, increased interest in the eye-region (or facial features in general) during the NE condition may reflect nine-month-olds increased motivation to ‘trigger’ a head turn from the other-race actor. Previous work has shown that gaze-following and race interact (Pickron et al., 2017; N. G. Xiao et al., 2018), while the race preference analysis from the previous section also indicates resiliency in nine-month-old other-race preferences (see Figure 4).

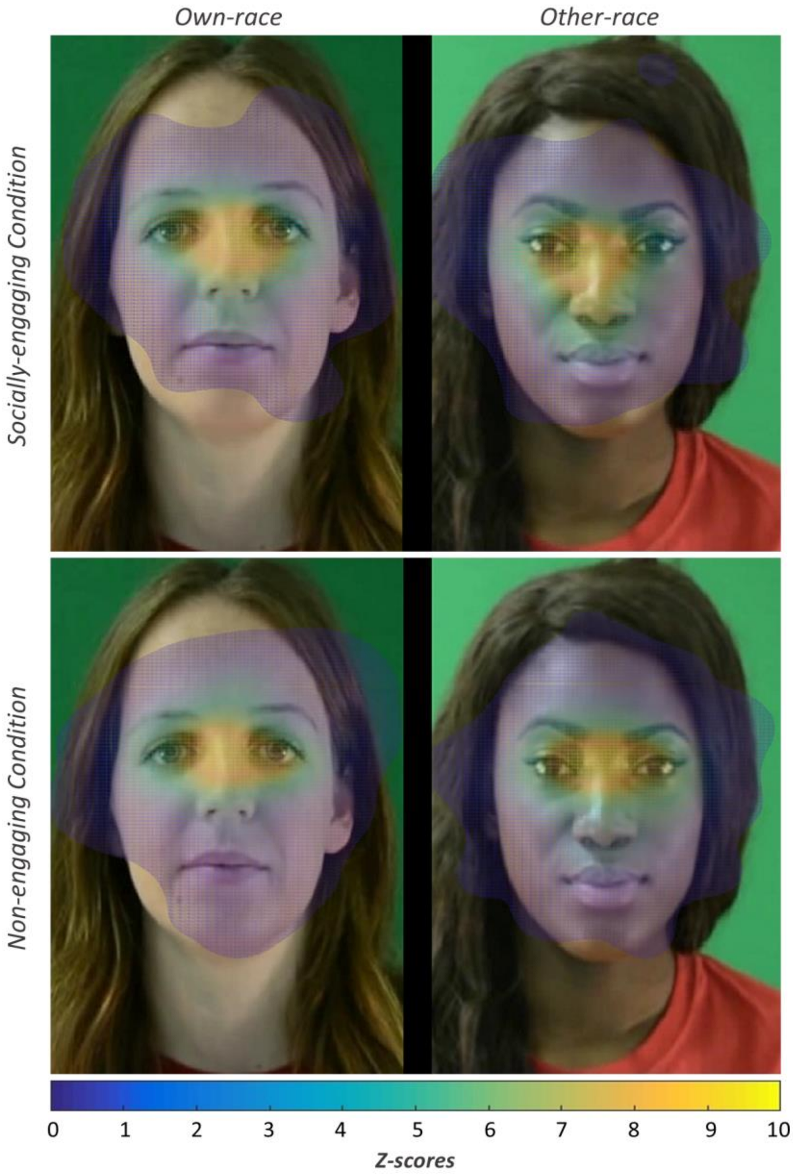

Figure 5. Descriptive fixation density heatmaps illustrating comparative looking during the preference phase. Fixation location and duration data across both preference trials were collated (trial 2 location data were transposed along the vertical axis as stimulus race was left-right counterbalanced across the two trials), and are displayed here as z-scores collapsed across Age. Data from the socially-engaging condition (top; N = 45) and non-engaging condition (bottom; N = 48), are presented separately.

Discussion

The results from Experiment 2 found that six- and twelve-month infants' *a priori* race preferences were nulled after an interactive *familiarisation* period in which actors contingently responded with identical social behaviours; either engagingly (condition 1) or non-engagingly (condition 2). Nine-month-olds, however did show proportionally greater looking toward the other-race stimulus, suggesting race preferences at this age show greater *resilience*. There was also some indication within the AOI analyses that nine-month-olds may have increased motivation to 'trigger' a head-turn response from the other-race actor (see Pickron et al., 2017; N. G. Xiao et al., 2018).

Experiment 2 also found that during *familiarisation* infants looked longer at the stimuli and the actors' facial features when they were responding engagingly compared to when they were responding non-engagingly. Maintaining mutual gaze and smiling has been shown to hold visual attention and induce positive affect in infants (Haviland & Lelwica, 1987; Kuchuk, Vibbert, & Bornstein, 1986; La Barbera, Izard, Vietze, & Parisi, 1976), while disengaging from the social interaction by averting gaze is not preferred by infants and triggers an attentional shift away from the face (Farroni, Menon, & Johnson, 2006; Keemink, Keshavarzi-Pour, et al., 2019; Scaife & Bruner, 1975).

Experiment 3 will investigate whether the type of socio-behavioural response given by the on-screen actors is sufficient to induce a test-phase visual preference, despite differences in the actor's race. During the *familiarisation* phase of this experiment, own- and other-race actors will respond to infants differently; either engagingly or non-engagingly. A first prediction for this experiment is that, as in Experiment 2, infants will be highly attentive toward the socio-behavioural responses of the actors, nullifying *a priori* race-preferences. A second prediction is that infants will show greater attention toward the socially-engaging actor during *familiarisation* as in Experiment 2, but will additionally show a preference for the socially-engaging actor during the subsequent *preference* phase, regardless of race.

Experiment 3

Method

Participants

Ninety-three healthy, full-term infants (6-12 months old) from Caucasian families were included in the final analysis of this experiment across two conditions and three age groups (see Table 4). An additional six participants were excluded due to missing data.

Table 4. Participant number for age, condition and gender in Experiment 3

Age (months)	**Age M (SD) (days)**	**Own-race SE**	**Own-race NE**	**Male**	**Female**
6	185.93 (9.48)	12	15	10	17
9	268.76 (10.61)	16	24	21	19
12	364.19 (13.35)	12	14	14	12

Stimuli and procedure

The stimuli and procedure were identical to Experiment 2 (see Figure 2) in all aspects except that the own- and other-race actors will now give differential responses during *familiarisation*. In the first condition, the own-race (Caucasian) actor gives socially-engaging (smile: open x 2, closed x 2) responses throughout, while the other-race

(African) actor gives non-engaging (head-turn away: left x 2, right x 2) responses throughout (OwnRace SE). In the second condition, these roles are reversed with a non-engaging own-race actor and a socially-engaging other-race actor (OwnRace NE). As before, the *familiarisation* phase consisted of eight trials (5s each), alternating between races. This was then followed by a *preference* phase consisting of two trials (10s each), counterbalancing for side.

Results: Race preferences

Familiarisation phase

A 3 (Age: 6, 9 or 12 months) x 2 (FaceRace: Caucasian or African) x 2 (Condition: OwnRace SE or OwnRace NE) x 4 (Trial) mixed ANOVA conducted on looking time duration showed no significant main effect of FaceRace ($F(1,87) = 1.461$, $p = .230$, $ŋ_p^2 = .017$), Condition ($F(1,87) = .115$, $p = .736$, $ŋ_p^2 = .001$) or Age ($F(2,87) = .607$, $p = .547$, $ŋ_p^2 = .014$), and the critical FaceRace x Condition interaction was also non-significant ($F(1,87) = 1.691$, $p = .197$, $ŋ_p^2 = .019$).

There was a significant main effect of Trial ($F(3,261) = 8.748$, $p < .001$, $ŋ_p^2 = .091$). Bonferroni-corrected comparisons suggest that infants became less interested in the stimuli toward the end of the familiarisation phase (due to decreasing interest; see Fantz, 1964) as there was significantly less looking in trial 4 compared to trial 3 ($p = .021$), trial 2 ($p = .046$) and trial 1 ($p < .001$, see Table 5). Trial did not significantly interact with FaceRace ($F(3,261) = .671$, $p = .571$, $ŋ_p^2 = .008$), Condition ($F(3,261) = 1.955$, $p = .121$, $ŋ_p^2 = .022$) or Age ($F(6,261) = 1.442$, $p = .199$, $ŋ_p^2 = .032$), but the Trial x Condition x Age interaction was significant ($F(6,261) = 2.203$, $p = .043$, $ŋ_p^2 = .048$). Post-hoc comparisons indicate that this effect was driven by a significant difference within the twelve-month age group only, who showed longer trial 4 looking for the second condition, (OwnRace NE Trial 4 M = 3.86s), compared the first condition (OwnRace SE Trial 4 M = 2.83s, p = .025). All other contrasts and interactions were non-significant (all p > .09).

Table 5. Mean fixation durations (secs) during the *familiarisation* phase of Experiment 3

		Trial 1	**Trial 2**	**Trial 3**	**Trial 4**
C1	**Own Race SE**	4.07	3.95	4.12	3.57
	Other Race NE	4.11	4.10	3.91	3.58
C2	**Own Race NE**	4.57	3.87	3.67	3.62
	Other Race SE	3.91	3.63	3.81	3.51

Preference phase

A 2 (FaceRace: Caucasian or African) x 2 (Condition: SE or NE) x 3 (Age: 6, 9 or 12 months) ANOVA was used to investigate fixation durations toward the stimulus across both preference trials. There was no significant main effect of FaceRace (*F*(1,174) = .146, *p* = .703, $ŋ_p^2$ = .001), and FaceRace did not significantly interact with Condition (*F*(1,174) = .185, *p* = .667, $ŋ_p^2$ = .001) or Age (*F*(2,174) = .246, *p* = .782, $ŋ_p^2$ = .003), and the FaceRace x Age x Condition interaction was also non-significant (*F*(2,174) = 2.191, *p* = .115, $ŋ_p^2$ = .025). Comparing proportional looking durations (% of total) toward the own-race stimulus against chance (50%) also yielded no significant race preferences overall (*M* = 48.77%, t(92) = 1.172, *p* = .244, *d* = .244), or within individual age and condition groups (all *p* > .16, see Figure 6).

There was no main effect of Condition (*F*(1,174) = .868, *p* = .353, $ŋ_p^2$ = .005), but there was a significant main effect of Age (*F*(2,174) = 4.530, *p* = .012, $ŋ_p^2$ = .049), and a significant Age x Condition interaction (*F*(2,174) = 3.492, *p* = .033, η_p^2 = .039). Post-hoc comparisons indicate that six-month-olds (*M* = 3.67s) showed significantly longer overall looking compared to nine-month-olds (*M* = 3.03s, *p* = .013), and marginally longer looking compared to twelve-month-olds (*M* = 3.12s, *p* = .077). For the interaction, age differences were significant within the OwnRace SE condition only (*F*(2,174) = 5.743, *p* = .004, η_p^2 = .062), and condition differences within the six-month group only (*F*(1,174) = 5.544, *p* = .034, η_p^2 = .025). Six-month-olds therefore showed significantly longer *preference* phase looking during the OwnRace SE condition (*M* = 4.03s) compared to nine (*M* = 3.19s, *p* = .039) and twelve-month-old (*M* = 2.86s, *p* = .004) infants.

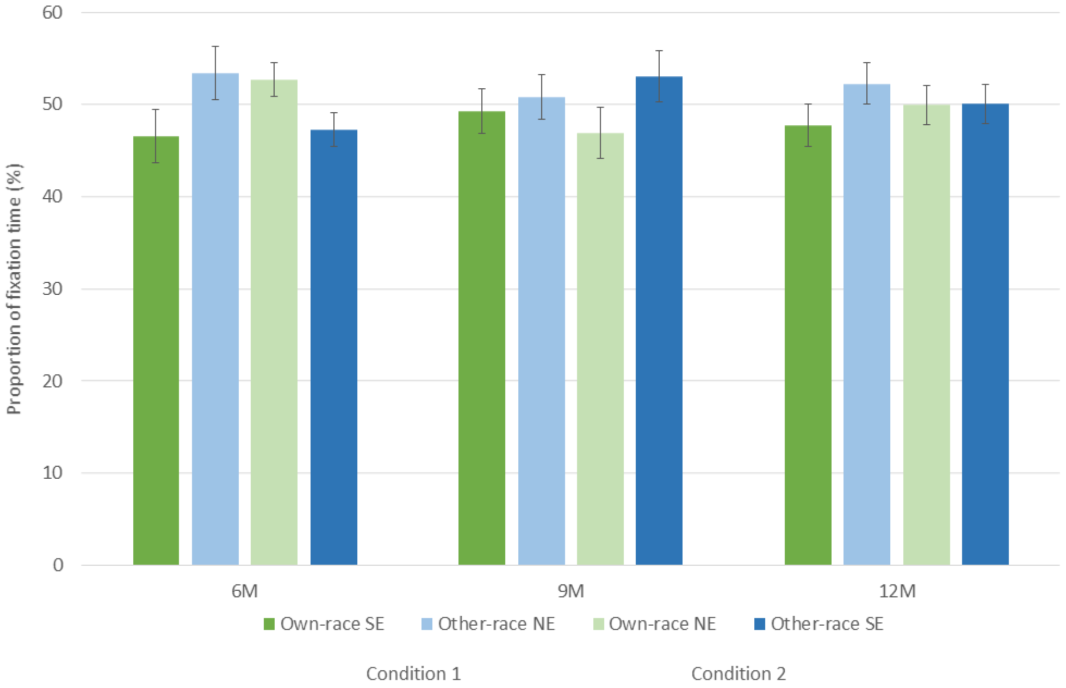

Figure 6. Race preferences in Experiment 3.

Mean proportional fixation durations towards the own-race (green) and other-race (blue) actors for the from the preference phase in experiment 3, across age and condition groups. Bold colours represent actors that responded engagingly (smiles) and faded colours represent actors that responded non-engagingly (head-turn away) during the familiarisation phase. Error bars represent standard error, all comparisons against chance (50%) were non-significant (all p > .16).

Results: Developmental differences in scanning

Experiment 2 found that infants show particular scanning behaviours in response to the actor's socio-behavioural cues; namely smile-looking and gaze-following (see Figure 7). To attempt to replicate these findings, but also to explore developmental differences in social scanning, a 2 (FaceRace: Caucasian or African) x 2 (Condition: OwnRace SE or OwnRace NE) x 3 (AOI: Eyes, Mouth or Gaze) x 3 (Age: 6, 9 or 12 months) mixed ANOVA was conducted for proportional (% of total) looking durations. This analysis yielded a main effect of AOI ($F(2,261) = 87.447$, $p < .001$, $ŋ_p^2 = .401$), and a significant FaceRace x Condition x AOI interaction ($F(2,261) = 134.183$, $p < .001$, $ŋ_p^2 = .507$). These large effects correspond to the differential patterns of face scanning necessarily produced by infants in response to the on-screen actor's contingent behaviours. Post-hoc (Bonferroni-corrected) comparisons indicate that all Condition and FaceRace contrasts produced significant differences in AOI looking (all $p < .007$), with socially-engaging responses attracting higher mouth- and eye-region looking and non-engaging responses attracting higher gaze-region looking (see Table 6 and Figure 7). The ANOVA was repeated for Behaviour (SE or NE) instead of FaceRace, which indicated that there were no significant race differences in face scanning as the critical Behaviour x Condition x AOI interaction was non-significant ($F(2,261) = .731$, $p = .482$, $ŋ_p^2 = .006$).

Table 6. Differences in *familiarisation* phase AOI looking (% of total fixation duration) produced in response to race and socio-behavioural cues (estimated marginal means)

	Condition 1		**Condition 2**	
	Own Race SE	**Other Race NE**	**Own Race NE**	**Other Race SE**
Eyes	45.59	24.33	27.33	42.64
Mouth	15.10	3.78	5.32	17.84
Gaze	3.99	30.26	33.74	3.05

While there were no significant main effects for FaceRace (F(1,261) = 1.426, p = .233, ŋ$_p^2$ = .005), Condition (F(1,261) = .556, p = .456, ŋ$_p^2$ = .002) or Age (F(2,261) = .657, p = .519, ŋ$_p^2$ = .005), there was a significant Age x AOI interaction (F(4,261) = 2.697, p = .031, ŋ$_p^2$ = .040). This age difference was significant for the eye-region (F(2,261) = 4.275, p = .015, ŋ$_p^2$ = .032), with twelve-month-olds showing significantly less eye-region looking (M = 14.68%) compared to six-month-olds (M = 19.60%, p = .015), and marginally less eye-region looking compared to nine-month-olds (M = 18.18%, p = .090, see Figure 7). There was also a marginal FaceRace x Age x AOI interaction (F(4,261) = 2.053, p = .087, ŋ$_p^2$ = .031). Post-hoc comparisons indicated that nine-month-olds showed significantly higher own-race eye-region looking (M = 20.38%) compared to other-race (M = 15.98%, p = .009), and twelve-month-olds showed significantly higher own-race gaze-region looking (M = 10.59%) compared to other-race (M = 6.45%, p = .042). These patterns can be viewed in Figure 7. The FaceRace x Age x AOI x Condition interaction (F(4,261) = .634, p = .639, ŋ$_p^2$ = .010) and all other interactions (all p > .12) were non-significant.

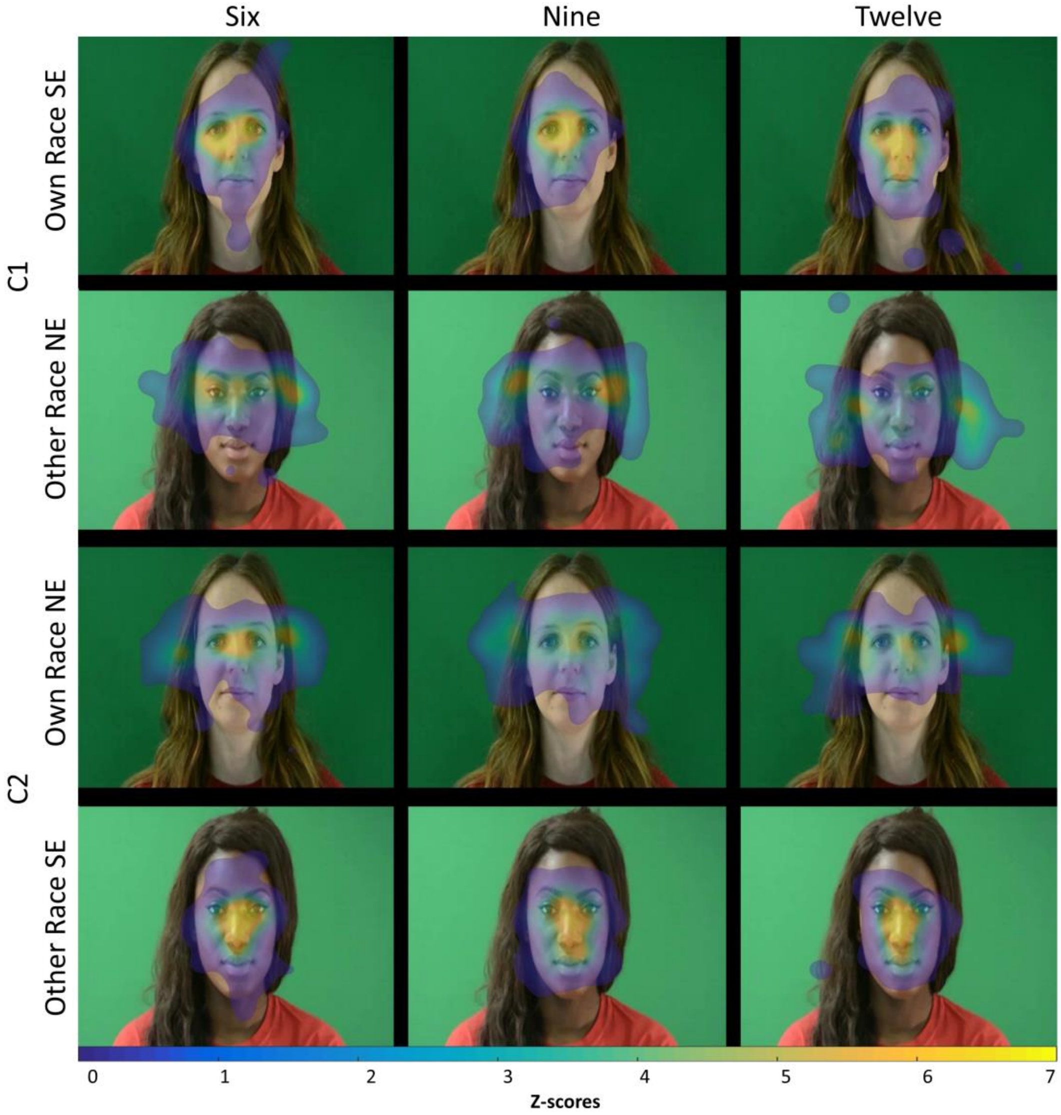

Figure 7. Development heatmaps depicting social scanning.

Descriptive heatmaps displaying fixation data (duration and location, z-scores) for the familiarisation phase of experiment 3 (N = 93) are shown here overlaying neutral stills of the on-screen actor. In condition 1 (C1, N = 40), the own-race actor gave socially-engaging (smiling) contingent responses and the other-race actor gave non-engaging (head-turn away) contingent responses. In condition 2 (C2, N = 53), these behaviours were reversed. Heatmaps are displayed for both own- and other-race actors in both conditions across three age groups: six, nine and twelve months.

Discussion

In Experiment 3, own- and other-race on-screen actors performed competing positive (socially-engaging) and negative (non-engaging) socio-behavioural responses during an interactive *familiarisation* phase, before being presented side-by-side in a *preference* phase. Similar to the previous experiment, Experiment 3 found that when infants experienced these brief interactions during *familiarisation*, proportional looking between actors in the *preference* phase was nulled. However, in this experiment the lack of preference occurred despite the actors' performing *different* behavioural responses during *familiarisation*. Note that as it became clear infants were not showing significant preferences for race or social behaviour, Experiment 3 was ended prematurely to prioritise data collection for other experiments, hence the relatively low sample sizes (see Table 4).

While six-month-olds did show longer *preference* looking overall in the first condition, given that six-month-olds have shown own-race preferences (Experiment 1), it is tempting to conclude that this is due to the engaging own-race actor in this condition. However, proportional looking values indicate that it was in fact the non-engaging other-race actor that drew a higher proportion of their interest (see Figure 6). These findings suggest that the non-engaging 'head-turn away' response is likely not as negative or aversive to infants as was first predicted. Infants were again highly attentive toward *both* behavioural responses, showing the same social scanning patterns (smile-looking and gaze-following) as in Experiment 2 (see Figure 7). It is therefore conceivable that infants are interpreting the contingent head-turn response as an informative behavioural cue, rather than a communication of disinterest (c.f. Keemink, Keshavarzi-Pour, et al., 2019), and thus do not show a preference for either behaviour.

Experiment 2 found differences in face scanning dependent on the race of the on-screen actor, with greater upper-face looking for own-race and greater lower-face looking for other-race during the *familiarisation* interactions. Experiment 3 further explored differences in face scanning, investigating also its development from six to twelve months. Firstly, the analysis of scanning behaviours found less eye-region looking in twelve-month-olds compared to six-month-olds (compare 6M and 12M heatmaps in Figure 7), a finding which complements that of Experiment 2 where older infants showed greater mouth-region looking, which is thought to facilitate language learning (Lewkowicz & Hansen-Tift, 2012; Tenenbaum et al., 2013). Secondly, twelve-month-olds show comparatively less gaze-region looking for the other-race face. However the heatmaps (Figure 7: 12M OwnRace NE vs OtherRace NE), suggest that in this case, like in Experiment 2, the gaze-following trajectory for twelve-month-olds was spatially lower for other-race, and therefore some fixations may have fallen outside of the boundary of the gaze-region AOIs (Figure 3). Finally, this experiment found that nine-month-olds show increased eye-region looking for own-race. This effect can be explained by consulting the descriptive heatmaps (Figure 7) which depict nine-month-olds as demonstrating cleaner social scanning (i.e. smile-looking and gaze-following) for the other-race actor. This finding is consistent with previous work suggesting infants' scanning of social cues interact with race (Liu et al., 2011; Pickron et al., 2017; Wheeler et al., 2011; N. G. Xiao et al., 2018).

General Discussion

After replicating and extending previously reported spontaneous infant race preferences (Kelly et al., 2007, 2005; Liu et al., 2015; Sugden & Marquis, 2017) in Experiment 1, these preferences were comprehensively removed for six- and twelve-month infants (but persisted in nine-month-olds) after they were able to briefly interact with on-screen actors during a *familiarisation* phase. During the *familiarisation* interactions, infants looked longer within the socially-engaging condition. However, when, in Experiment 3, actors gave differential socio-behavioural responses during *familiarisation*, infants in all age groups showed no race preferences, but also did not show any preferences for engaging (mutual gaze, smiling) behaviours over non-engaging (turning away) behaviours as was predicted. There were also no asymmetries in looking during *familiarisation*.

A first conclusion from these experiments is that other-race preferences are somewhat more resilient in nine-month-olds. Previous literature (Fassbender et al., 2016; Liu et al., 2015) suggests that race preferences are most robust within three- (own-race preference) and nine-month-old (other-race preference) infants, so it is therefore consistent with this literature that within Experiment 2 proportional looking bias for other-race remained within the nine-month age group. Perhaps the more general transition from familiarity to novelty preference at this age (see Quinn et al., 2019) provides a clue as to why other-race preferences persist for nine-month-olds only. Interestingly, interactions between race and social scanning were also present within the nine-month age group. For instance in the *preference* phase of Experiment 2, nine-month-olds showed increased other-race eye-region looking, though only for the non-engaging condition. Unlike *familiarisation,* the *preference* phase stimuli did not gaze-contingently animate, thus increased eye-region looking may reflect an increased ‘motivation’ to trigger a dynamic response from the other-race actor. While in Experiment 3, nine-month-olds showed increased own-race eye-region looking during *familiarisation*, reflecting much cleaner social scanning (i.e. smile-looking and gaze-following) for the other-race actor (see Figure 7). Taken together, these findings suggest socio-perceptual and socio-behavioural cues interact, and that nine-month-olds show increased sensitivity toward the social behaviours of other-race adults.

Nevertheless it is important to note that nine-month other-race preferences showed a reduced effect size in Experiment 2 (where actors gave identical behavioural responses), and were further reduced to a non-significant trend within Experiment 3 where the saliency of the actors' socio-behavioural responses was enhanced (actors gave differential behavioural responses). Moreover, six- and twelve-month-old race preferences were nulled for both Experiment 2 and Experiment 3. These findings suggest that the importance of race in guiding visual preference may have been overstated previously, and that preferences may have reduced strength in real-world environments where faces are dynamic, interactive and communicative. Instead, socio-behavioural information is likely to be prioritised, while 'race' may not necessarily be encoded. This interpretation is consistent with evidence from studies conducted with a variety of age groups showing that 'race effects' of categorisation, preference and recognition can all be reduced or removed when additional socially salient information is provided (Kinzler, Shutts, DeJesus, & Spelke, 2009; Kinzler & Spelke, 2011; Kurzban et al., 2001). It is also consistent with work exploring the disruption of the ORE within recognition paradigms. For instance, training with several static face exemplars over several weeks (Anzures et al., 2012; Heron-Delaney et al., 2011) or directly within a test setting (Sangrigoli & De Schonen, 2004) is sufficient to remove an own-race recognition advantage. A recent study (Minar & Lewkowicz, 2017) has also shown that by using 'richer' stimuli, in this case dynamic face stimuli accompanied by concordant speech vocalisations, older infants were able to 'overcome' perceptual narrowing and recognise other-race faces. This chapter's findings extend this pattern of behaviour to infant race preferences.

The reduction in race preferences could be due to the reduced salience of race when presented alongside naturally co-occurring social cues (e.g. dynamic facial gestures), nevertheless there is an alternative explanation for these findings. It is conceivable that the inclusion of a *familiarisation* phase also may have inadvertently habituated infants to both face identities. This scenario would result in a reduction of infants' attention to both on-screen actors during the *preference* phase. Nevertheless, such a disinterest in the stimulus could still be asymmetrical, preserving race preferences.

Experiments 2 and 3 of this chapter introduced the GC Social Interaction paradigm to the infant race processing literature. This novel methodology permits infants to actively engage with a stimulus, which is not only a radical departure from the typically used 'static face', but critically it is also moves us closer to simulating infant's day-to-day experiences (see Chapter 2). The efficacy of the stimuli in extracting social interest and awareness is evidenced by the scan patterns evoked during the *familiarisation* phase of Experiments 2 and 3 in which socio-behavioural cues interacted with race to determine the focus of an infants' attention. Nearly all infants showed smile-looking for both actors, but relative to own-race, other-race face fixations were often spatially lower. Similarly, infants tracked the shifting gaze of both on-screen actors, but according to a lower spatial trajectory for the other-race actor. Such findings are consistent with differences in face scanning across faces of different races reported previously (Pickron et al., 2017; Wheeler et al., 2011; W. S. Xiao et al., 2014, 2013). However, the meaningfulness of these race differences in face scanning should again not be overstated. Social scanning patterns were automatic for most, with only minor adjustments made for race – perhaps due to facial morphological differences of the face stimuli used in these studies. Given that were no significant asymmetries in looking duration to own- and other-race faces during these interactions, it seems reasonable to infer that face processing for both actors was functionally equivalent;

and that it was the social behaviour, not the race of the actors, that was of primary interest to infants.

Although the experiments within this chapter implement the GC Social Interaction paradigm during *familiarisation*, they still utilise a force-choice preference task to measure infant race preferences. However, it is not clear how these methodologies might interact. For instance, an infant familiarised with responsive social partners might expect responsiveness from the actors during the *preference* test, and thus might seek to 'trigger' a response by looking back and forth between the actors. The visual preference result may therefore not truly represent infants' intrinsic biases. Further, this paradigm relies on spontaneous infant preferences as infants were not systematically habituated to one actor, and the interpretation of such preferences can be challenging. Infants can show intrinsic preferences for both familiarity and novelty, and these biases can also shift across development (Quinn et al., 2019). When infants are presented with multiple cues, such as is the case here, the task of determining why infants are looking longer to a stimulus becomes increasingly more complex (see Chapter 8 for a more detailed discussion of the limitations of visual preference paradigms). Future work might consider using the active role afforded to infants within gaze-contingent paradigms to assess asymmetries in race preference within social interactions. For instance, presenting responsive on-screen actors simultaneously would allow researchers to investigate if infants were more willing to initiate interactions with own- or other-race adults, rather than relying on spontaneous visual preference alone.

Conclusion

The experiments in this chapter have shown that race preferences are nulled for infants aged six and twelve months after a *familiarisation* period in which they could ‘interact’ with on-screen actors. Nine-month other-race preferences showed greater resiliency, but were also attenuated with the increasing salience of the actors’ social behaviours. And though differences in facial scanning were found, with fixations dropping spatially lower for other-race faces, these data show that face scanning was functionally equivalent across race, with social behaviour being the primary determinant of infants’ visual attention. Therefore, given that infants experience dynamic and interactive faces in their everyday life, and perceptual and behavioural cues naturally co-occur, it is conceivable that infant race preferences, particularly for six- and twelve-month-olds, may not necessarily emerge in real-world environments.

Chapter 4

Do infants show a preference for socially-engaging behaviours?

Abstract

Chapter 3 investigated infants' preferences for race and social behaviours following brief interactions with on-screen actors and found no preferences for positive, engaging behaviours relative to non-engaging behaviours. This chapter further explores the reason for this unintuitive finding, firstly by removing the influence of race (Experiment 1), secondly by reducing the motion and communicativeness of the non-engaging cue (Experiment 2), and finally by removing the contingency of the non-engaging cue (Experiment 3). The results indicate that infants' preferences for social behaviours remain at approximately chance level, despite these alterations. Possible methodological reasons for these results are discussed.

Introduction

Infants demonstrate sensitivity to a variety of social behaviours including facial expressions (Addabbo et al., 2018; Bayet et al., 2017; Farroni et al., 2007), eye gaze cues (Brooks & Meltzoff, 2005; Farroni et al., 2004) and interpersonal responsiveness (Striano et al., 2005; Striano, Henning, & Stahl, 2006). Previous work has also shown that it is behaviour, and not facial characteristics, that are readily encoded, remembered and discriminated by infants (Bahrick et al., 2002).

The previous chapter found that infant face scanning was highly predicted by the contingently-triggered dynamic behaviours of the on-screen actors, with both socially-engaging (smiling, mutual gaze) and non-engaging (turning away) behaviours capturing their attention. It was also found (Experiment 2, Chapter 3) that infants' showed greater interest in the stimulus when the actor was responding engagingly compared to when they were responding non-engagingly, and looked longer toward the facial features of the socially-engaging actor. Yet a visual preference for the actor performing engaging behaviours (Haviland & Lelwica, 1987; Kuchuk et al., 1986; La Barbera et al., 1976) over the actor performing non-engaging behaviours (Farroni et al., 2006; Hains & Muir, 1996; Scaife & Bruner, 1975) did not emerge during the *preference* phase whilst the on-screen actors also differed according to race (Experiment 3 of the previous chapter). Instead, potential behavioural preferences, alongside spontaneous race preferences found in the literature (Fassbender et al., 2016; Kelly et al., 2005; Liu et al., 2015) and in Experiment 1 of the previous chapter, were nulled following brief exposure to both engaging and non-engaging behavioural cues.

The following chapter investigates infants' spontaneous preferences for social behaviours within the GC Social Interaction paradigm (see Chapter 2). Recent work using similar paradigms (Keemink, Keshavarzi-Pour, et al., 2019; Vernetti et al., 2018, 2017) has found that infants preferentially trigger socially-engaging behaviours from on-screen actors and provide a greater frequency of positive responses (e.g. smiles) for interactive faces relative to interactive toys (Vernetti et al., 2018), and for socially-engaging faces relative to non-engaging faces (Keemink, Keshavarzi-Pour, et al., 2019). Faces, and particularly smiling faces, possess intrinsic hedonic reward value (e.g. O'Doherty et al., 2003), which might be heightened within simulated social interactions relative to static stimuli (Kilts, Egan, Gideon, Ely, & Hoffman, 2003; Labar, Crupain, Voyvodic, & McCarthy, 2003; Sato, Fujimura, & Suzuki, 2008). It is therefore predicted that when the potentially confounding factor of race is removed, infants will show spontaneously longer looking toward an actor that responds socially-engagingly over an actor that responds non-engagingly, and will encode this preference such that a greater proportion of looking will be directed toward this actor during a subsequent visual preference test.

Experiment 1

Method

Thirty-seven Caucasian infants from three age groups were included in the final analysis of this experiment (6, 9 and 12 months +/- 14 days; Table 1). An additional fourteen participants were excluded for having incomplete data (see Chapter 2 for further participant, eye tracking and data processing information). The stimuli and procedures for this experiment were identical to those used in Experiment 3 of the previous chapter except the other-race actor's stimuli were exchanged for those from a second own-race actor (see Figure 2, Chapter 3), which were again matched for attractiveness and distinctiveness (Kelly et al., 2007). As before, infants completed both a *familiarisation* and a *preference* phase. The *familiarisation* phase consisted of eight brief social interactions with two on-screen Caucasian actors (four with each). The actor's behavioural responses were triggered by infants fixating within a predefined, invisible gaze-contingent boundary placed over the eye region. One actor responded consistently with smiles (socially-engaging 'Smiling' response; 2 x OpenSmile, 2 x ClosedSmile), while the other actor consistently turned their head to one side (non-engaging 'TurnAway' response; 2 x TurnLeft, 2 x TurnRight). Each trial lasted 5 seconds. To counter-balance across actors, infants were allocated to one of two conditions. In the first condition Actor 1 responded engagingly and Actor 2 responded non-engagingly, in the second condition these roles were reversed.

Following *familiarisation*, infants were presented with two ten-second *preference* trials, in which both actors were presented side-by-side, with their left-right spatial position reversed in the second trial. During this phase the actors were unresponsive, and maintained neutral affect and natural blinking.

Table 1. Participant information for Experiments 1-3

	Age (months)	**Age M (SD) (days)**	**Condition N**		**Gender N**	
			One	**Two**	**Male**	**Female**
Experiment 1	6	194.67 (5.92)	6	3	4	5
	9	283.81 (11.82)	9	7	7	9
	12	361.25 (21.55)	7	5	8	4
Experiment 2	6	198.85 (21.34)	7	6	7	6
	9	269.86 (12.30)	13	8	12	9
	12	360.82 (13.61)	5	12	8	9
Experiment 3	6	190.67 (12.12)	12	-	6	6
	9	270.30 (7.89)	10	-	8	2
	12	378.56 (21.48)	9	-	5	4

Results

Familiarisation phase

To investigate whether infants looked longer toward the actor that responded engagingly during *familiarisation,* a 2 (Actor: Actor 1 or Actor 2) x 2 (Condition: Actor 1 SE or Actor 1 NE) x 4 (Trial) mixed ANOVA was performed for total fixation duration. No significant main effects of Actor ($F(1,35) = .232$, $p = .633$, $ŋ_p^2 = .007$) or Condition

(*F*(1,35) = .019, *p* = .891, $ŋ_p^2$ = .001) were found, but the ANOVA yielded a significant main effect of Trial (*F*(3,105) = 3.372, *p* = .021, $ŋ_p^2$ = .088) in which infants showed significantly less looking in trial 4 (*M* = 3.57s) compared to trial 2 (*M* = 4.30s, *p* = .015), reflecting a known tendency for infants to habituate to repetitive stimuli (Fantz, 1964). Trial did not show any significant interactions (all *p* > .34) and the critical interaction between Actor and Condition was also non-significant (*F*(1,35) = .102, *p* = .751, $ŋ_p^2$ = .003).

Preference phase

A 2 (Condition) x 2 (Behaviour: OpenSmile vs TurnAway) between subjects ANOVA was conducted for fixation durations during the *preference* phase finding no significant main effect of Condition (*F*(1,70) = .827, *p* = .366, $ŋ_p^2$ = .012) or Behaviour (*F*(1,70) = .270, *p* = .605, $ŋ_p^2$ = .004), and no significant Condition x Behaviour interaction (*F*(1,70) = .303, *p* = .584, $ŋ_p^2$ = .004). Overall, proportional looking toward the socially-engaging actor during the *preference* phase was not significantly different to chance (*M* = 48.96%, *t*(36) = -.702, *p* = .487, *d* = .234).

Discussion

Experiment 1 found that infants do not look longer toward on-screen actors that give engaging, smiling responses compared to those who disengage and turn away from them. They also did not show any visual preference based on behaviour during the non-responsive *preference* phase. These findings with own-race actors replicate those of the previous chapter (Experiment 3) and indicate that infants find engaging and disengaging behaviours equally interesting or informative. It became clear during testing that infants were not showing preferences for socially-engaging behaviour, so the experiment was concluded prematurely, hence the small sample size (*N* = 37, Table 1).

This result is unintuitive and inconsistent with previous literature (Haviland & Lelwica, 1987; Kuchuk et al., 1986; La Barbera et al., 1976), however it is conceivable, given that infants follow the gaze of the non-engaging actor (see previous chapter, Figures 5 and 7), that the head-turn response is being interpreted by infants as a communicative gaze cue rather than a symbol of disinterest. To test this hypothesis, Experiment 1 was repeated but with a different non-engaging behavioural cue. Instead of turning away to the left or right once infants make eye contact, the non-engaging actor instead closed their eyes. This gesture cannot be interpreted as a communicative gaze cue, and has minimal motion, and therefore may better represent a 'non-engaging' response. For Experiment 2 it is hypothesised that infants will show a preference toward actors that perform smiling, engaging behaviours, over those that give a non-engaging, closed-eye response.

Experiment 2

Method

Fifty-one Caucasian infants were included in the final analysis of this experiment (see Table 1 and Chapter 2 for further information), with an additional nine infants excluded for having missing data. The procedures from Experiment 1 were repeated after making two stimulus alterations. Firstly, for the engaging response four 'OpenSmile' responses were used in order to maximise the intensity of these interactions. Secondly, the non-engaging 'TurnAway' response was exchanged for a 'ClosedEyes' response in which actors simply closed their eyes for the remainder of the trial when gaze-contingently triggered.

Results

Familiarisation phase

A 2 (Behaviour: OpenSmile or ClosedEyes) x 2 (Condition: Actor 1 OS or Actor 1 ClosedEyes) x 4 (Trial) ANOVA for fixation durations during *familiarisation* found no significant main effect of Behaviour ($F(1,49) = 1.118$, $p = .296$, $ŋ_p^2 = .022$). Marginally significant effects were found for Trial ($F(3,147) = 2.567$, $p = .057$, $ŋ_p^2 = .050$), Condition ($F(1,49) = 3.758$, $p = .058$, $ŋ_p^2 = .071$) and the Behaviour x Trial interaction ($F(3,147) = 2.599$, $p = .054$, $ŋ_p^2 = .050$). All other interactions were non-significant (all $p > .30$).

There was significantly less looking in trial 4 ($M = 3.62$s) compared to trial 1 ($M = 4.06$s, $p = .047$), suggesting infants became less interested in the stimuli towards the end of the *faniliarisation* phase. For the interaction with Behaviour, post-hoc tests indicate that infants looked marginally longer toward the open-smile stimulus compared to the closed-eyes stimulus in the first (OpenSmile $M = 4.39$s, ClosedEyes $M = 3.74$s, $p = .054$) trial, but no significant difference in looking during the second and third trials (both $p > .09$). Longer overall looking durations were also found in the second condition (Actor 1 OS, $M = 4.00$s) compared to the first condition (Actor 1 CE, $M = 3.57$s, $p = .058$).

Preference phase

Looking durations during the *preference* phase were analysed using a 2 (Condition) x 2 (Behaviour) between subjects ANOVA which found no significant main effect of Behaviour ($F(1,98) = .157$, $p = .846$, $ŋ_p^2 = .009$) or Condition ($F(1,98) = .157$, $p = .693$, $ŋ_p^2 = .002$), and no significant Condition x Behaviour interaction ($F(1,98) = 1.171$, $p = .282$, $ŋ_p^2 = .012$). Overall, proportional looking toward the open-smile actor during the *preference* phase was not significantly different to chance ($M = 48.76\%$, $t(50) = -.968$, $p = .338$, $d = .274$).

Discussion

As in the first experiment, Experiment 2 also found that infants do not show preferential looking toward actors who have interacted engagingly over those who have interacted non-engagingly. The closed-eyes modification did little to alter *preference* phase looking, but there were marginal asymmetries in looking duration based on behavioural response over the course of the *familiarisation* phase. In particular infants showed greater interest to the smiling actor compared to the eye-closing actor when the stimuli were first presented. However, this initial interest quickly faded and was not encoded with actor identity such that a visual preference emerged in the subsequent *preference* phase. This experiment was again cut short (N = 51) once it became clear that infants' visual preferences were at chance level.

Given that it is the actor's behavioural responses that are most predictive of infant scanning (see previous chapter), it is surprising that infants do not seem to clearly prefer one behaviour over another, not even for low-level stimulus properties such as motion (Vinter, 1986). One possible explanation for these findings could be that any behavioural response, regardless of its perceived communicative value, is of interest to infants as long as it is contingent. Previous work has shown the importance of contingency in social interactions, and how infants consistently prefer contingent over non-contingent social partners (Bigelow & Birch, 1999; Murray & Trevarthen, 1986; Striano, Henning, & Vaish, 2006). In a third experiment this hypothesis was tested by again adapting the non-engaging stimulus. In Experiment 3, non-engaging actors were now unresponsive; they maintained neutral affect and random blinking for the duration of their *familiarisation* trials, while socially-engaging actors continued to respond contingently with smiles.

Experiment 3

Method

Thirty-one Caucasian infants were included in the final analysis of this experiment (see Table 1 and Chapter 2 for further information), with an additional nine infants excluded for having missing data. The procedures from the previous experiment were again repeated in an altered form. In this experiment instead of closing their eyes, the non-engaging actor is now unresponsive during *familiarisation*. The actor remains neutral, with natural blinking, for the entirety of the trial, while the socially-engaging actor again responds contingently with open smiles as in the previous experiment. In order to maximise our sample, only one condition was included in this experiment given that no significant differences have emerged between Actor 1 and Actor 2 in the previous two experiments.

Results

Familiarisation phase

A 2 (Behaviour: OpenSmile or Unresponsive) x 4 (Trial) ANOVA was conducted for *familiarisation* phase fixation durations finding no significant main effect for either Behaviour ($F(1,30) = .076$, $p = .785$, $ŋ_p^2 = .003$) or Trial ($F(3,90) = 2.209$, $p = .092$, $ŋ_p^2 = .069$), and no significant Behaviour x Trial interaction ($F(3,90) = .595$, $p = .620$, $ŋ_p^2 = .019$). There was a non-significant trend of shorter looking durations toward the end of the *familiarisation* phase (trial 1 $M = 3.92$s, trial 4 $M = 3.27$s). Overall looking durations toward both types of *familiarisation* stimuli were approximately equal (OpenSmile $M = 3.69$s, Unresponsive $M = 3.64$s).

Preference phase

An independent groups t-test found no significant difference in looking durations between *preference* phase stimuli (OpenSmile M = 3.30s, Unresponsive M = 3.24, $t(60) = .222$, $p = .825$, $d = .057$). Proportional looking durations toward the OpenSmile stimulus did not significantly differ compared to chance (M = 50.47%, $t(30) = .308$, $p = .760$, $d = .112$).

Table 2. Mean looking durations toward *preference* phase stimuli

	Experiment 1 ($N = 37$)		Experiment 2 ($N = 51$)		Experiment 3 ($N = 31$)	
	OpenSmile	TurnAway	OpenSmile	ClosedEyes	OpenSmile	NoResponse
FD (s)	3.33	3.51	3.54	3.73	3.30	3.24
FD (%)	48.96	51.04	48.76	51.24	50.47	49.53

FD = Fixation Duration

Combined analyses: Developmental differences

Developmental differences in behavioural preference (OpenSmile vs NonEngaging: TurnAway, ClosedEyes and Unresponsive) were investigated after collapsing across the three experiments in this chapter ($N = 119$; six-month $N = 34$, nine-month $N = 47$, twelve-month $N = 38$). A 3 (Age: six, nine or twelve) x 2 (Behaviour: OpenSmile or NonEngaging) between subjects ANOVA yielded no significant main effects of Age ($F(2,232) = .445$, $p = .641$, $\eta_p^2 = .004$) or Behaviour ($F(1,232) = .645$, $p = .423$, $\eta_p^2 = .003$), and no significant interaction ($F(2,232) = .187$, $p = .829$, $\eta_p^2 = .002$). Comparing proportional looking to the open-smile actor against chance (50%) found no significant differences overall (*M* = 49.27%, *t*(118) = -.894, *p* = .373, *d* = .165) or within individual age groups (all *p* > .33). See Table 2 for a summary.

General Discussion

This chapter investigated infant visual preferences for behavioural responses given by two own-race (Caucasian) on-screen actors. In the **first experiment**, it was found that infants did not prefer the actor that gave contingent socially-engaging (mutual gaze, smiling) responses during familiarisation over the actor who gave contingent non-engaging responses (head-turn away). This replicated the results found for the cross-race experiments in the previous chapter, and suggests that infants found contingent socially-engaging and non-engaging responses equally as interesting. Speculating that the head-turn cue might be socially-informative to infants as a gaze-cue, in a **second experiment** the non-engaging behaviour was modified from ‘head-turn away’ to a less informative non-engaging response (closing their eyes). Surprisingly, infants still showed no preference for the socially-engaging actor. Reasoning that perhaps the lack of visual preference was due to both responses being equally contingent, a **third experiment** modified the non-engaging behavioural response to an unresponsive dynamic neutral face (natural blinking). It was again found that infants showed no visual preferences based on the actor’s socio-behavioural responses.

The three experiments in this chapter have found that, regardless of behaviour, infants’ *preference* following a simulated social interaction is approximately at chance. Known preferences for race were also nulled following *familiarisation* (see previous chapter). It seems highly unlikely that differences in both socio-behavioural factors, such as the nature or contingency of an interaction, and socio-perceptual factors, such as race, during *familiarisation* would have no impact on an infant’s visual preference.

Furthermore, given that socially-engaging stimuli have been shown to evoke pupillometric arousal responses in infants (Jessen et al., 2016), to elicit a high frequency of positive reciprocal responses from infants' within simulated interactions (Keemink, Keshavarzi-Pour, et al., 2019) and static images of smiling faces are often visually preferred by infants (Farroni et al., 2007; La Barbera et al., 1976), it seems unlikely that infants do not possess inherent preferences for contingent, dynamic, positive social behaviours over static, neutral and unresponsive behaviours. The root cause of these results is therefore considered to be methodological in nature.

The preferential looking paradigm is foundational to infant research (Fantz, 1956; Teller, 1979; see Chapter 1). A statistically significant difference in looking between two exemplars, presented side-by-side, suggests infants can discriminate between them, and thus reveals information about their cognitive or perceptual abilities. As the preferential looking paradigm relies on infants possessing an *a priori* or 'spontaneous' attentional bias to the stimuli, this method is often enhanced by preceding the preference test with a period of habituation to one exemplar. Thus a preference may be induced via the relative novelty of the unhabituated stimulus, and therefore increasing the likelihood that infants will demonstrate discrimination (Fantz, 1964). The present case did not attempt to induce a novelty preference toward one exemplar, but to see instead, using an interactive *familiarisation* phase, whether previously reported spontaneous visual preferences (both for behaviour and race) would emerge following brief social encounters with the actors.

Yet regardless of the nature of the social interactions during *familiarisation*, including pronounced behavioural asymmetries between the two actors, no preferences emerge. It may be the case that *any* interactivity or dynamic responsiveness within stimuli prior to a preferential looking procedure using static versions of the same stimuli negates visual preferences. Infants may be anticipating interactivity in the preference stimuli and therefore divide their looking equally between the static images in an attempt to ‘trigger’ a response. Alternatively, infants’ attention to the dynamic and contingent behaviours during *familiarisation* might disrupt or negate infants’ encoding of the actors’ identities, such that they are not recognised in the following forced-choice preference task (c.f. Bahrick et al., 2002). If either of these explanations are true, then this work has important methodological implications for how and when to use preferential looking procedures in infant testing, particularly as we attempt to increase the ecological validity of our methods by including dynamic or interactive stimuli. Future work might explicitly test these conclusions, and should perhaps consider alternatives to a static preferential looking task when using dynamic and interactive stimuli, or will perhaps utilise a combination of methods, such as recording behavioural responses and/or pupil size alongside looking behaviours (see Houston-Price & Nakai, 2004).

In conclusion, this chapter has investigated infants’ spontaneous preferences for social behaviour following brief simulated interactions with own-race actors. It was found that infants did not show attentional asymmetries according to an actor’s social behaviour during *familiarisation*, and also did not show any visual preference for socially-engaging behaviours in a following preferential looking task. These findings highlight potential limitations of using classical preferential looking tasks alongside interactive stimuli.

Chapter 5

Do infants scan static and dynamic facial expressions differently?

Abstract

Despite the inherent dynamism, communicativeness and contingency of ‘real-world’ social gestures, much of what we know about how infants’ perceive facial expressions relies on investigations that have presented static face images. One likely reason for this concerns the methodological challenges that arise when incorporating naturalistic stimuli within infant eye-tracking paradigms. In this chapter, infant scanning of dynamic and interactive videos of the six basic emotional expressions are compared to static images of the same stimuli. In doing so this chapter provides researchers with useful tools for analysing eye-tracking data; including a novel method (‘Dynamic AOIs’) and several existing analysis methods seldom used in infant research (heatmap and time-course analyses). This chapter finds clear, qualitative differences in how infants’ scan static and dynamic expressions, and suggests that more naturalistic stimuli should be used in future research.

Introduction

As researchers, we are trained to strive for sufficient experimental control. However, in doing so, there is a danger that we can distance ourselves from the very phenomena we are purporting to study. This is especially relevant when exploring how infants perceive and understand social information. By presenting unimodal, static and unresponsive stimuli within the ‘strange’ context of the lab setting, there is a risk that we might fail to capture even the essential components of the dynamic, communicative and interactive ‘real-world’ phenomena (such as facial expressions, gestures or eye gaze cues) frequently experienced by infants in their natural environments (see Walker-Andrews & Bahrick, 2001). Chapter 3 found that visual preferences for race using static stimuli (Kelly et al., 2005; Liu et al., 2015) were highly attenuated within the GC Social interaction paradigm. This chapter investigates infant facial expression perception, and compares whether the stimuli that we use, either static or dynamic, influences how infants scan facial expressions.

How humans perceive facial expressions is an enduring question (see Darwin, 1872; Wundt, 1909), and there has been much work investigating its developmental trajectory. This research has most often used static stimuli depicting faces at peak expressive amplitude, such as the now classic ‘Ekman faces’ (Ekman & Friesen, 1976; Ekman et al., 1987). However, even the earliest studies have identified the limitations of using static stimuli to investigate phenomena that are inherently dynamic (Caron et al., 1985; B. M. Wilcox & Clayton, 1968). Burton (2013), in his overview of progress within adult face recognition research, described how the field was being led astray by an over-reliance on artificial, static stimuli, reducing face-processing to simple picture matching once natural variance is removed. Walker-Andrews and Bahrick (2001) also argue that the use of dynamic, multimodal stimuli is critical for understanding how infants perceive and remember social information such as facial expressions, and question the generalizability of highly artificial methods (see also Lewkowicz, 2001; Schmuckler, 2001).

Recent studies in both adults (Richoz, Lao, Pascalis, & Caldara, 2018; see Krumhuber, Kappas, & Manstead, 2013 for a review) and infants (Addabbo et al., 2018; Godard, Baudouin, Schaal, & Durand, 2016; Heck et al., 2016; Soussignan et al., 2017) are increasingly using dynamic stimuli, and are discovering that dynamic social cues are processed and encoded in fundamentally different ways (Bahrick, Gogate, & Ruiz, 2002; see Chapter 2). Other work has shown that static stimuli may underestimate infant's abilities, with infants demonstrating sensitivity to facial emotions at earlier ages when realistic, dynamic displays are used (Addabbo et al., 2018; Heck et al., 2016; Montague & Walker-Andrews, 2001; Soussignan et al., 2017). Infants show a preference for (Ichikawa et al., 2011) and attend longer to dynamic expressions (B. M. Wilcox & Clayton, 1968), particularly to internal features (Haith et al., 1977). Dynamic faces are also scanned differently by infants across the first year of life (Soussignan et al., 2017; N. G. Xiao et al., 2015), with older infants showing increased attention toward 'diagnostic regions' (such as the mouth for happy and the nose for disgust; see Gosselin & Schyns, 2001; Smith, Cottrell, Gosselin, & Schyns, 2005) of dynamic facial expressions (Soussignan et al., 2017). If infants are selectively attending to facial regions that disambiguate between expressions (e.g. the wide eyes of fear distinguish these expressions from surprise), it suggests that they may be able to discriminate and make categorical distinctions between these expressions (Gosselin & Schyns, 2001; see Quinn et al., 2011).

This chapter will compare scanning differences between static and dynamic facial expression stimuli for all six basic expressions (happy, sad, surprise, fear, anger and disgust). This will be a substantial contribution to the literature given that no previous infant study has considered all six of the basic expressions, and whilst some researchers have begun to implement dynamic stimuli (e.g. Addabbo et al., 2018; Heck et al., 2016), contingency has thus far been overlooked. This is critical given the universality and biological salience of all six of these expressions (Ekman et al., 1987, but see Jack, Sun, Delis, Garrod, & Schyns, 2016), which are dynamically, but also contingently, communicated in natural environments. This chapter will also introduce several novel and adapted methodological techniques that have been developed by the author to overcome the challenges of using more complex and naturalistic stimuli. Facial expressions are inherently dynamic phenomena, therefore if clear differences between static and dynamic do emerge, then it suggests that a transition toward more complex and dynamic stimuli is warranted if we are to make meaningful and generalizable claims about infants' everyday processing of facial expressions.

Hypotheses

In this experiment, it is predicted that infants will attend to and scan dynamic and static facial expressions differently. More specifically, it is predicted that infants will show greater interest, and therefore look longer toward, dynamic stimuli (cf. B. M. Wilcox & Clayton, 1968) and that for dynamic there will be more precise and coordinated looking toward diagnostic regions (cf. Soussignan et al., 2017). For instance, as the smile emerges within a *dynamic* happy expression, it is expected that there will be a steeper peak in looking toward the mouth, while a flatter and more continuous pattern of mouth-region looking across time is expected for the *static* happy expression.

Methods

Participants

One hundred and eighteen infants (6, 9 and 12 months, see Table 1 for participant details and Chapter 2 for further information) were assigned to either the 'dynamic' (*N* = 77) or the 'static' (*N* = 41) condition. An additional 31 infants were excluded from the Area of Interest (AOI) analysis for not possessing fixation duration data for all expression trials, but were retained for the heatmap and time-course analyses. The sample size used in this experiment is comparable to infant eye-tracking studies in this field (e.g. Hunnius & Geuze, 2004; Soussignan et al., 2017).

Table 1. Participant information for the initial and reduced cohorts

	Age (months)	Age M (SD) (days)	Condition N		Gender N	
			Dynamic	Static	Male	Female
Initial	6	191.17 (9.78)	33	10	21	22
	9	274.77 (12.69)	32	24	29	27
	12	366.76 (13.42)	36	14	25	25
Reduced	6	192.51 (15.32)	28	10	19	19
	9	273.78 (13.73)	20	17	21	16
	12	366.24 (11.62)	29	14	23	20

Stimuli

Eighteen expression videos were recorded (Nikon D5200 digital camera) with six different actors (3 male, 3 female) such that each actor contributed three videos (1 neutral, 2 expressive), with each of the six core expressions (happy, sad, surprise, fear, anger, disgust) being recorded twice (by 1 male, and 1 female) and neutral six times (1 per actor, see Figure 1). Each video (720 x 576 pixels) was edited to be three seconds in length; beginning with neutral affect and ending at peak expressive amplitude. All actors wore an identical black t-shirt, and were seated in front of a uniform green background. Fifty-one adult observers were asked to firstly recognise the expression displayed in the video, then to rate how representative (1-5 scale) this video was for the target expression (i.e. how close do our stimuli match the facial expression we are trying to represent?). Expressions were highly recognised (*M* = 84.57%, *SD* = 14.99%), and received high representativeness ratings (*M* = 3.62, *SD* = .38, see Table 2). The images used in the static condition were stills taken from these videos, when the expression was judged to be at ‘peak’ amplitude (i.e. the point at which the expression reaches its highest intensity; c.f. Ekman & Friesen, 1976; see Figure 1).

Figure 1. The six basic expressions.

A selection of expressive stimuli used within the static condition (Top) that were created by taking stills from the dynamic expression videos at 'peak' expressive amplitude. The expressive stimuli used in this chapter are compared to classic expression stimuli (Bottom; Ekman & Friesen, 1971). The six 'basic' expressions are used (Ekman et al., 1987): happy, sad, surprise, fear, anger and disgust; from left to right respectively.

While previous studies have used artificially-generated (Soussignan et al., 2017) or quasi-dynamic (Ruba, Johnson, Harris, & Wilbourn, 2017) facial expression stimuli, this study took the novel step of using real human faces. Though this may sacrifice some control, realism and variance within a stimuli set can be seen as a methodological strength, as it better reflects an infant's everyday experience, and prevents the collection of exemplar-specific data. According to Burton (2013), a lack of realism and variance is a potential drawback of much of the stimuli currently used in face research. Burton recommends using stimuli that capture multiple instances, shot with several different cameras, in order to produce a generalised representation of a face. Whilst not all of these recommendations have been applied here, the stimuli used in this experiment are video clips of real faces in motion, shot across multiple identities and both genders, which should provide rich and naturalistic exemplars that can relate to an infant's everyday experience.

Table 2. Mean recognition accuracy and representativeness ratings for the static expression stimuli

	Neutral	Happy	Sad	Surprise	Fear	Anger	Disgust
Accuracy (%)	86.93	96.08	94.12	92.16	66.67	85.29	67.65
Rating (1-5)	3.73	3.62	3.02	3.92	3.78	3.43	3.12

N = 51 adult observers; Accuracy chance level = 14.29%

Procedure

The experimental stimuli were presented within the GC Social Interaction paradigm (see Chapter 2). The experiment consisted of 18 trials, one for each expression video, presented in a fully randomised order. For the **dynamic condition** each trial consisted of a brief social interaction. Each interaction began with an attention-grabber located to the left or right of the screen (counterbalanced across trials), and once fixated, the first frame of the expression video (the actor, facing forward with neutral affect) appeared centrally on the screen. An invisible gaze-contingent boundary was placed over the eye region, and a fixation (minimum duration 100 msecs) within this region triggered the facial animation (i.e. the playing of the expression video). Infants, therefore, contingently triggered the on-screen actor to respond with one of the six basic emotional expressions (or with neutral affect in the control trials) by engaging them in eye contact. Trials ended after five seconds if the eye region was not fixated. If the eye region was fixated rapidly, the three-second expression videos paused on the last frame to ensure each trial reached its five-second duration. Trials within the **static condition** were not dynamic or interactive. Each trial instead presented a static image from the expression video (at peak expression) for the full five-second duration. Figure 2 provides a graphical representation of the procedure used in both conditions.

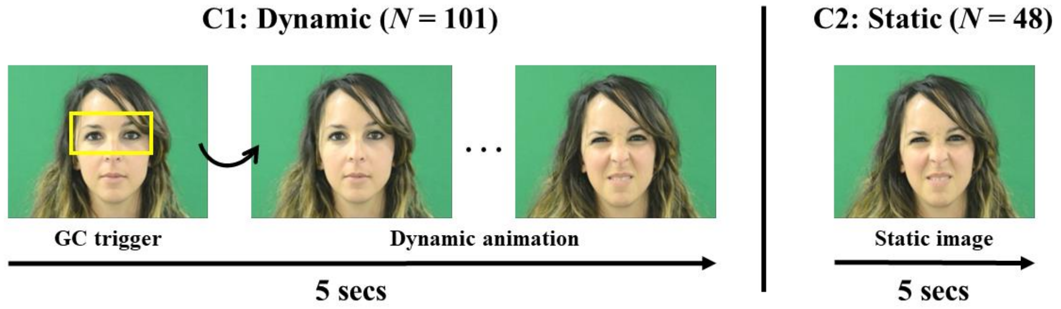

Figure 2. An illustration of the experimental procedure.

The GC Social Interaction paradigm was used to investigate infants' scanning of dynamic (condition 1; 'C1') and static (Condition 2; 'C2') facial expressions. 18 trials in both conditions lasted 5 seconds each, and were presented in a fully randomised order. In the dynamic condition, a 3-second expression animation was contingently-triggered by infants when they fixated the actor's eye-region (yellow boundary). In the static condition, an image of the same expression was presented for the full 5-second duration.

The GC Social Interaction paradigm used here and in previous chapters is novel for facial expression research, and allows the investigation of dynamic expressions embedded within contingent interactions. Although there is a growing appreciation within the literature that dynamism is an essential component of creating ecologically valid stimuli (Heck et al., 2016; Krumhuber et al., 2013; Quinn et al., 2011; N. G. Xiao et al., 2015), contingency has thus far largely been overlooked. Infants are highly sensitive to social contingency (Bigelow & Birch, 1999), and by two to three months this is an expected element of social interactions (Mcquaid et al., 2009; Soussignan et al., 2006; Striano et al., 2005; Striano, Henning, & Stahl, 2006). Allowing infants to interact with the on-screen actor through gaze-contingent responses increases realism, but is also methodologically important. Using gaze-contingently activated videos ensures all infants are fixating the same location (eye-region) before the expression response begins. Scan paths for each expression therefore have a standardised start point in both time and space. In this case, the analysis window begins when infants' first fixate the eye region (in both static and dynamic conditions), and lasts for three seconds. This therefore facilitates analysis across trials, participants and conditions, but it also provides the additional benefit of guaranteeing infant attentiveness toward the expression animation.

Analytical tools

Dynamic AOIs. In Chapter 3, traditional fixed AOIs were used to identify *what* an infant was looking at on the screen (e.g. which facial region). However, fixed AOIs are problematic when using dynamic stimuli (Hessels et al., 2016), and are often inadequate for identifying precisely where an infant is looking (see Chapter 1 for a discussion). Considering this, custom MATLAB scripts were written by the author to identify interest regions directly from the video stimulus (Figure 3, Video Figure 1).

This method uses colour and luminance information within video frames to define background, hair, upper torso, face and facial features (see Kolkur, Kalbande, Shimpi, Bapat, & Jatakia, 2017 for a similar skin detection method). After pre-defining a face 'midpoint' and skin tone range for each actor, numerical matrices defining six AOIs (Background, hair/torso, upper-face skin, upper-face feature, lower face skin, lower face feature) were automatically computed for each video frame (or for just the single image used for the static condition). The script first identifies the region of the video frame that corresponds to the green background, before identifying pixel regions that are within the actor's skin tone range. This denotes the background, and differentiates the face from the torso and hair. This ‘skin mask’ defines the outer contour of the face, but also inversely defines the location of inner facial features such as the mouth, nose, eyes and eyebrows. These feature regions are ‘coalesced’, by smoothing (using a Gaussian filter) and dilating (by applying a size 1, disk-shaped structural element) feature regions, and extracting the larger connected components (by applying a binary area filter) to form discrete AOIs. Feature AOIs are then divided into ‘upper’ and ‘lower’ categories either side of the midline (centred upon the bridge of the nose). Specific ‘Dynamic AOIs’ were thus generated for each expression video that can accurately accommodate for variations in the size and location of interest regions between actors and across time (see Hessels, Benjamins, Cornelissen, & Hooge, 2018 for an alternate approach).

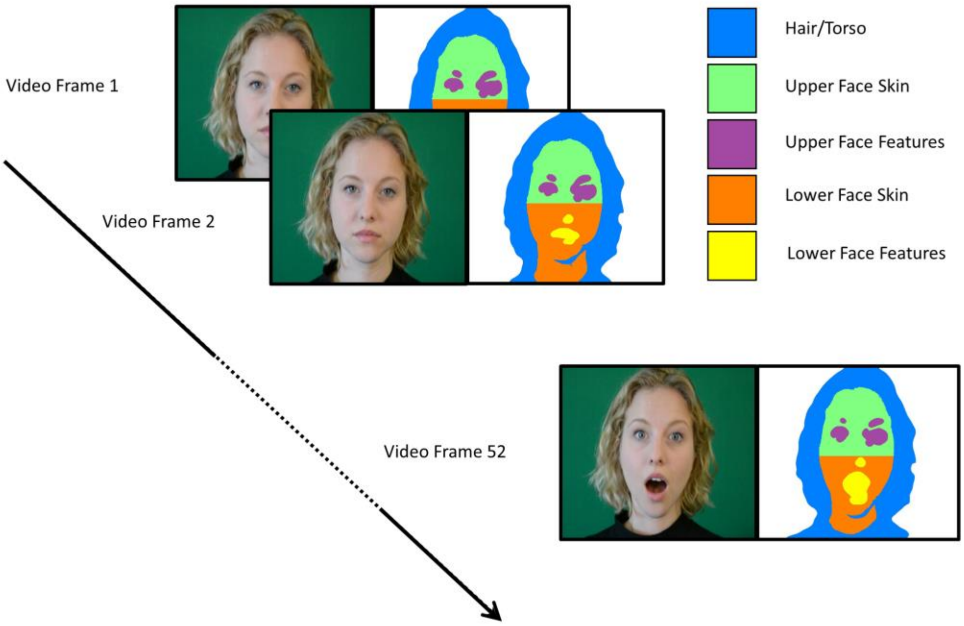

Figure 3. Dynamic AOIs.

Areas of interest (AOIs) were generated for the facial expression animations using colour and luminance information present within each video frame. After pre-defining a facial midpoint and skin tone range for each actor, numerical matrices defining six AOI regions were automatically computed for each video frame. These numerical matrices when combined to form "dynamic AOIs", which are depicted here using colour labels (background [white], hair and torso [blue], upper face skin [green] and features [purple], lower face skin [orange] and features [yellow]), alongside an image of the video frame from which they were generated.

Statistical heatmaps. Descriptive fixation heatmaps (see Caldara & Miellet, 2011) were produced by summing all fixation durations within a trial for each pixel 'coordinate', smoothing these with a Gaussian kernel and computing a matrix of z-scores from the resulting values. Fixations on the background were removed to prevent z-score inflation. Difference maps were created by subtracting one map from another (e.g., dynamic minus static). This process was repeated for each participant individually, and concatenated into a multidimensional array of differential z-score matrices. These arrays were then used within MATLAB’s independent groups t-test function (‘ttest2’; MathWorks, Statistics and Machine Learning Toolbox) to produce a heatmap of t-values. Once a critical t-score was identified (after applying a Bonferroni correction), clusters of significant differences between conditions could be identified. To compare across age (see Chapter 6), this process was adapted to generate one-way ANOVA heatmaps. Finally, to display data descriptively across time, ‘dynamic heatmaps’ were created by generating a series of heatmaps depicting fixation data time-locked to each video frame (see Video Figures 2 and 3 for examples).

Temporal analyses. When using dynamic stimuli, on-screen information changes over time. Time therefore becomes a relevant variable, so collapsing data across this dimension may obscure important patterns (D. J. Barr, 2008). Here we used growth-curve analysis to incorporate time as a predictor within mixed-effects regression models using the ‘lmer’ function from the ‘lme4’ R package (Bates, Maechler, & Bolker, 2014; Mirman, Dixon, & Magnuson, 2008). For these analyses, three seconds of fixation data from each trial (from the gaze-contingent onset till the end of the expression animation) were aggregated to 20ms time slots forming a binomial dataset for each AOI (Fixating AOI = 1, Not fixating AOI = 0). Mean proportions were computed, however to account for their bounded nature, they were then adjusted to the log-odds scale using an empirical logit (E-log) transformation:

$$\eta = ln\left(\frac{\phi + \varepsilon}{1 - \phi + \varepsilon}\right)$$

Where η represents the likelihood of fixating a particular AOI, ϕ represents a given proportion and ε is a constant value equal to 0.5 (D. J. Barr, 2008). Groups were also sum-coded (e.g. Condition: Dynamic = +.5, Static = -.5) so that the intercept represented mean log-odds. The curvilinear relationship of time course on AOI looking was then modelled (across 60ms time bins) using orthogonal power polynomials (via the ‘eyetrackingR’ package; Dink & Ferguson, 2015; see Mirman et al., 2008). These are capable of representing complex functional forms by incorporating higher order components (e.g. quadratic, cubic, quartic). Within these models the intercept represents effects irrespective of time, the slope reflects a unidirectional change over time, the 2nd order (quadratic) term represents a symmetrical double change (i.e. the rise and fall of a curve) over time, while 3rd and 4th order (cubic and quartic) terms reflect three and four changes over time respectively, capturing any steepness or asymmetry of the curve around the inflection point. The change in deviance (ΔD) based on the deviance statistic (-2LL; minus two times the log-likelihood) was used to assess whether each additional parameter significantly improved model fit.

Cluster Permutation Analysis (Dink & Ferguson, 2015; Maris & Oostenveld, 2007) is a complementary temporal analysis method that can be used to generate potential time-windows of significant divergence between conditions. In the first pass, this method performs a statistical test for each time bin (e.g. every 60ms), and significant bins are grouped to form temporal ‘clusters’ of significant difference. To control for multiple comparisons, a histogram of test statistics (a permutation distribution) is generated by repeatedly ‘shuffling’ the data and clustering significant time bins. The first-pass clusters are compared against this distribution to obtain a *p*-value (using the Monte Carlo estimate). Thus significant temporal divergence clusters can be identified using this method.

Analysis plan

This chapter will explore the question of whether infants scan dynamic and static facial expressions differently. Firstly, a traditional analysis will be performed by collapsing across time and AOI region, comparing total fixation time toward static and dynamic expression stimuli. Secondly, the proportion of the total stimulus looking time fixated upon individual AOI regions (face, lower face, and lower face features) will be computed using the Dynamic AOI's. The face AOI was defined as all stimulus fixations that were not located on the background or hair/torso, while the lower face AOI was defined as any fixation within the lower face skin or feature regions (see Figure 1). As the eye region is the default start point for each expression, and the mouth-region shows the greatest change when dynamic, the analysis will focus on the lower face. The results from the upper face AOIs will be only briefly reported as they are expected to be the inverse of lower face. Substantial overlap is also expected between lower face and lower feature analyses, however both will be included in order to investigate how dynamic AOIs can be best used to optimally capture face scanning. Thirdly, the analysis will then focus on one actor's stimulus set in greater detail to explore the utility of heatmaps, mixed-effects modelling and cluster permutation analysis for analysing the rich data generated from naturalistic eye-tracking paradigms (results from the remaining trials will be only briefly summarised for the sake of space). Together these methods will indicate precisely *where* and *when* significant differences in face scanning emerge between conditions.

Results

Looking durations toward the stimulus

A 2 (Condition: dynamic and static) x 6 (Expression: happiness, sadness, surprise, fear, anger, disgust) mixed ANOVA for stimulus fixation durations (FDs) yielded a main effect of Expression ($F(5,580) = 5.503$, $p < .001$, $\eta_p^2 = .045$) and a main effect of Condition

($F(1,116) = 9.564$, $p = .002$, $\eta_p^2 = .076$), with higher average looking times toward dynamic ($M = 4.13$s) compared to static ($M = 3.78$s) stimuli (Figure 4a). Pairwise comparisons reveal that infants looked significantly longer toward dynamic happy ($p = .008$), sad ($p = .034$) and disgust ($p = .001$) expressions. Expression and Condition did not show a significant interaction ($F(5,580) = 1.550$, $p = .172$, $\eta_p^2 = .013$).

Looking durations toward AOI regions

Face. Proportionate fixation durations toward the face region (% of stimulus FD) were investigated using a 2 (Condition) x 6 (Expression) mixed ANOVA, yielding a significant main effect of Expression ($F(4.11,477) = 16.157$, $p < .001$, $\eta_p^2 = .122$) but no significant main effect of Condition (Dynamic $M = 91.86\%$, Static $M = 90.34\%$; $F(1,116) = 1.621$, $p = .206$, $\eta_p^2 = .014$), or interaction ($F(4.11,477) = 1.785$, $p = .129$, $\eta_p^2 = .015$). Pairwise comparisons indicate a higher proportion of looking toward the dynamic face for the angry expression only ($p = .044$, Figure 4b).

Lower face. To analyse differential looking toward the lower face region (% of stimulus FD), a 2 (Condition) x 6 (Expression) mixed ANOVA was conducted finding significant main effects for both Expression ($F(5,580) = 29.542$, $p < .001$, $\eta_p^2 = .203$) and Condition ($F(1,116) = 5.658$, $p = .019$, $\eta_p^2 = .047$), though no significant interaction ($F(5,580) = .670$, $p = .646$, $\eta_p^2 = .006$). The proportion of time spent fixating the lower face was higher for dynamic stimuli ($M = 42.60\%$), compared to static stimuli ($M = 32.73\%$, Figure 4c). The same analysis for upper face also yielded a significant main effect of Condition ($F(1,116) = 4.194$, $p = .043$, $\eta_p^2 = .035$) in the opposite direction (Dynamic $M = 49.26\%$, Static $M = 57.61\%$). Lower face looking was significantly different between conditions for happiness ($p = .019$) and anger ($p = .007$), with a marginal effect for fear ($p = .067$).

Lower face features. Proportional looking toward lower face features (% of stimulus FD) were analysed using a 2 (Condition) x 6 (Expression) mixed ANOVA. This analysis yielded a significant main effect of Expression (*F*(5,580) = 24.829, $p < .001$, $ŋ_p^2 = .176$), a marginal effect of Condition (*F*(1,116) = 3.887, $p = .051$, $ŋ_p^2 = .032$), but no significant interaction (*F*(1,116) = .166, $p =.684$, $ŋ_p^2 = .001$). Pairwise comparisons indicated significantly higher proportion of looking toward the lower features of the dynamic angry expression ($p = .044$), and marginally so for happy ($p = .063$) and disgust ($p = .060$, see Figure 4d). Conversely, upper face features showed an opposite trend (Dynamic $M = 21.94\%$, Static $M = 26.55\%$, *F*(1,116) = 2.848, $p =.094$, $ŋ_p^2 = .024$), with a significant difference between conditions for the happy expression only (Dynamic Happy $M = 22.26\%$, Static Happy $M = 30.57\%$, $p = .049$). It is also noted that the lower face feature AOIs captured substantially lower proportions of fixations compared to the lower face AOIs (compare Figure 4c and d).

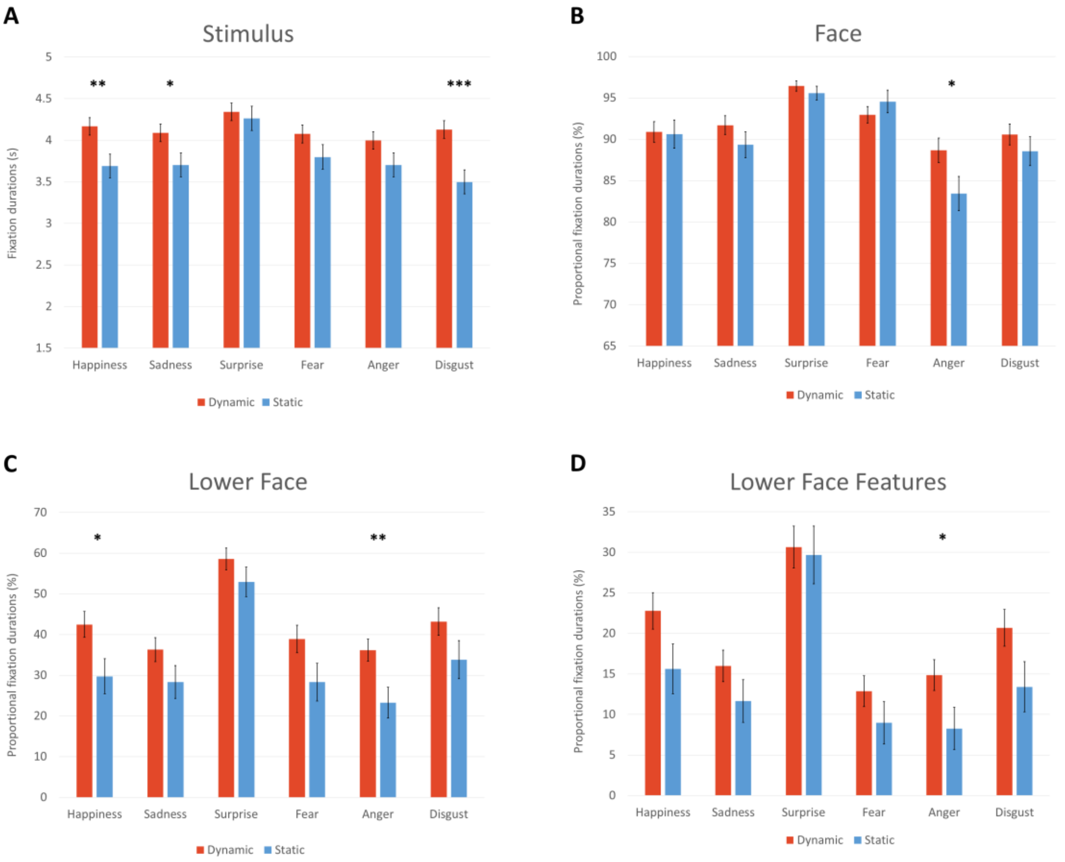

Figure 4. Looking durations to the stimulus and AOIs for all six expressions

Infant visual attention toward the whole stimulus (A) and AOI regions (B-D) divided by Condition (dynamic, red and static, blue) and Expression (happiness, sadness, surprise, fear, anger and disgust). Mean fixation durations (seconds) toward the entire stimulus (A) were used to compute proportional looking time (%) to the face (B), lower face (C) and lower feature (D) AOIs. Significant pairwise comparisons between conditions are identified (* p < .05, ** p < .01, *** p < .001, two-tailed). Error bars are standard error of the mean.

Summary

In line with classic work (B. M. Wilcox & Clayton, 1968), these analyses find that compared to static images, infants look longer toward dynamic and contingent representations of expressions, indicating greater interest in this type of stimuli. The effect was not uniform across expressions, however, with particularly strong effects for happiness and disgust, but little difference between dynamic and static surprise (Figure 4a). The Dynamic AOI analyses find firstly that there are clear differences in scanning between expressions, but also that infants show greater interest in the lower facial regions when viewing dynamic expressions, particularly so for happiness and anger.

Happiness and anger: Heatmap analysis

Differences in scanning behaviours between conditions can be captured in heatmap presentations and further investigated using time-course analyses. The entire participant cohort (Dynamic $N = 101$, Static $N = 48$), were used for these analyses and the data from one actor who performed neutral, happy and angry expressions are reported here (Figure 5; see Figures 6 and 7 for other expressions). The heatmaps (Figure 5a-c), depict spatial differences in face scanning between conditions as z-scores (Dynamic = red and Static = blue), with significant regions ($p < .05$) outlined in black. For the static neutral stimulus (Figure 5a), looking clustered more toward the centre, perhaps driven by blinking in the dynamic neutral trial increasing the salience of the peripheral eye-region. For dynamic happy and angry expressions, however, the heatmaps show a significant lower face bias (Figures 5b and 5c). For both expressions, significance clusters for the dynamic condition emerge around the mouth, chin and nose regions, and in the upper face and eye-region for the static condition (see Video Figures 2 (angry) and 3 (happy) for side-by-side comparisons of dynamic and static conditions).

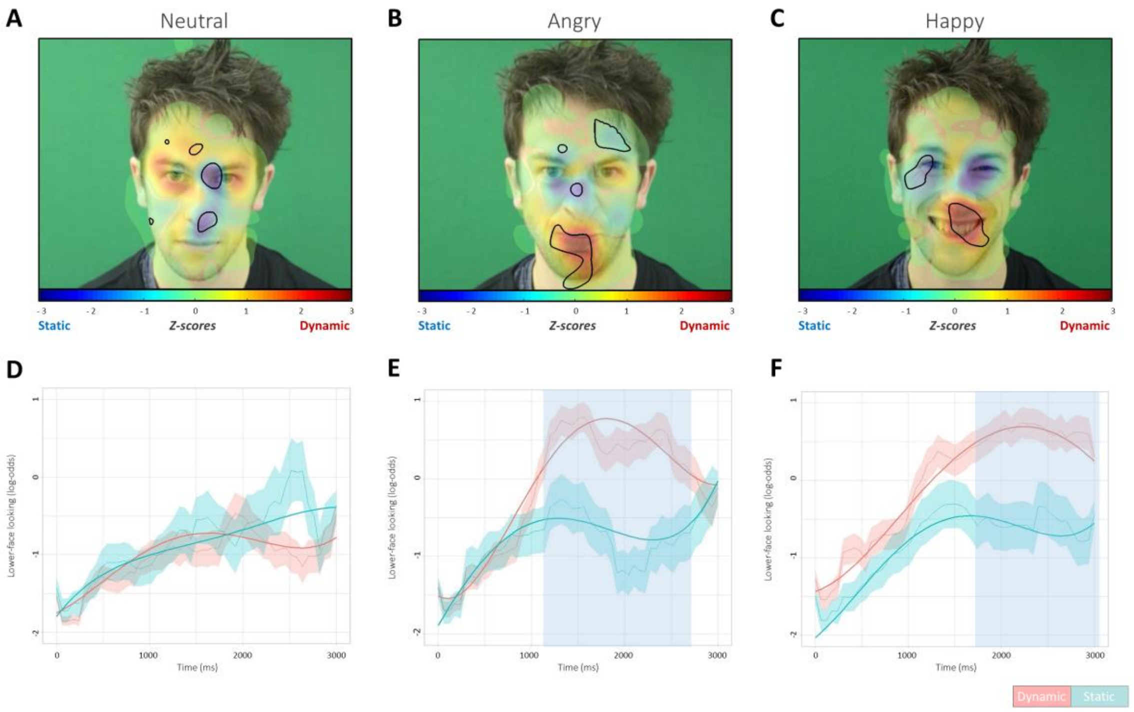

Figure 5. Heatmaps and time-course analyses for neutral, happy and angry expressions.

Statistical heatmap and time-course analyses for neutral (A, D), angry (B, E) and happy (C, F) trials for one actor. The heatmaps (A-C) depict spatial differences in face scanning between conditions as z-scores, with positive values denoting greater looking for the dynamic condition (orange/red) and negative values denoting greater looking for the static condition (blue). Regions of significant difference between conditions ($p < .05$, two-tailed) are outlined in black. Differences in lower-face looking between dynamic (red) and static (blue) conditions across the expression animation time-window (3 seconds) were modelled using polynomial growth curves and analysed using multilevel logistic regression (D-F). For each trial, the log-likelihood of fixating the lower face was computed across 60ms time-bins, plotted here (faint line) with the standard error of the mean (coloured border). These data were then modelled using linear, quadratic, cubic and quartic polynomial time terms (solid line). Cluster permutation analysis was used to identify time-windows of significant differences between conditions (blue/grey shaded region).

Happiness and anger: Temporal dynamics

Building on the AOI and heatmap analyses that show increased lower face looking for dynamic happy and angry expressions (see Figures 4c and 5b-c), here *temporal* differences in face-scanning between conditions are investigated. These analyses model the log-likelihood of lower face looking across time (Figure 5d-f) using growth curves (Mirman et al., 2008) and multilevel logistic regression (D. J. Barr, 2008). The curves generated from the power polynomials overlay the fixation data in Figure 5d-f; and the model estimates, standard errors and t-scores are shown in Table 3. In these models the empirical log-likelihood of fixating the lower face AOI was the dependent variable (LowerFace), with Condition (Dynamic = .5, Static = -.5) and Time as predictors. Time was defined by four polynomial terms: linear (t1), quadratic (t2), cubic (t3) and quartic (t4). Random effects of Participant were also included within the model, allowing Participant to vary according to the intercept and all polynomial time terms. This was formalised in Wilkinson notation as:

$$\begin{aligned} LowerFace \sim\ & Condition(t1 + t2 + t3 + t4) \\ & + (1 + t1 + t2 + t3 + t4 \mid Participant) \end{aligned}$$

The model (see Table 3) identified a significant negative intercept (all $p < .01$) and positive linear effects for all three expressions (all $p < .001$). These effects indicate that while overall there was greater upper-face looking (intercept-only), there was also a general increase in lower face fixations across time (linear). Of all facial features, the eye-region is attended the most (see Itier & Batty, 2009 for a review), and face scanning tends to begin with fixating the eye-region before moving downward (Schyns, Petro, & Smith, 2007). This pattern has also been amplified by the gaze-contingent manipulation that ensured the analysis time-window began with an eye-region fixation. Significant negative quadratic effects (all $p < .001$), suggest the time course of lower face looking followed an "inverted-U" curve, with interest in the lower face first rising, before decreasing again toward the end of the trial, though this pattern was less prominent in the neutral trial (neutral $t = -3.38$, angry $t = -8.06$, happy $t = -6.96$). Anger also showed a significant main effect for the quartic time term (*p* = .004), suggesting that the peak in lower face looking occurred toward the midpoint in time, flanked by “troughs” either side.

In accordance with the heatmap and AOI analyses, there were significant main effects of Condition, indicating increased lower face looking for dynamic happy (*p* = .001) and angry (*p* = .005) expressions compared to static, but not for dynamic neutral (*p* = .768). For happiness, Condition significantly interacted with the linear term (*p* = .019), and for anger, Condition significantly interacted with linear, quadratic and cubic temporal terms (all *p* < .02). The significant linear interactions indicate that both dynamic happy and angry expressions show a greater increase in lower face looking across time compared to static. The negative Condition interactions with the higher temporal terms for anger suggest dynamic lower face looking followed a stronger inverted-U pattern (quadratic interaction), with its peak shifted more toward the latter half of the trial (cubic interaction) compared to static. A lack of significant interactions between Condition and the higher-order polynomials for happy suggests that the main difference across time between the groups was linear in nature.

Table 3. Model fit, estimates and t-values for lower-face looking within neutral, angry and happy trials

	Model Fit			Parameter estimates			
Fixed Effect	-2LL	*ΔD*	*P*	*β*	*SE*	*t*	*P*
Neutral							
Intercept	16393	–	–	–.974	.101	–9.654	**< .0001**
Condition	16393	0	.5264	–.063	.213	–.296	.7677
Linear	16064	329	**< .0001**	2.041	.386	5.285	**< .0001**
Quadratic	15962	102	**< .0001**	–.985	.290	–3.393	**.0009**
Cubic	15955	7	**.0090**	.260	.295	.884	.3784
Quartic	15952	3	.0729	.133	.196	.678	.4988
Condition: Linear	15938	14	**.0002**	–.277	.808	–.342	.7328
Condition: Quadratic	15923	15	**.0001**	–.838	.614	–1.365	.1747
Condition: Cubic	15923	0	.9303	–.269	.621	–.433	.6659
Condition: Quartic	15917	6	**.0126**	.044	.413	.107	.9151
Angry							
Intercept	20786	–	–	–.345	.103	–3.332	**.0011**
Condition	20780	6	**.0096**	.629	.219	2.878	**.0047**
Linear	20070	710	**< .0001**	2.935	.379	7.752	**< .0001**
Quadratic	19546	524	**< .0001**	–2.927	.364	–8.035	**< .0001**
Cubic	19546	0	.6387	–.125	.285	–.439	.6611
Quartic	19502	44	**< .0001**	.734	.256	2.869	**.0048**
Condition: Linear	19416	86	**< .0001**	1.945	.801	2.428	**.0166**
Condition: Quadratic	19335	81	**< .0001**	–2.744	.775	–3.543	**.0005**
Condition: Cubic	19260	75	**< .0001**	–1.890	.605	–3.121	**.0022**
Condition: Quartic	19254	6	**.0117**	.766	.544	1.409	.1612
Happy							
Intercept	21728	–	–	–.320	.109	–2.946	**.0038**
Condition	21717	11	**.0008**	.770	.232	3.325	**.0011**
Linear	20538	1179	**< .0001**	3.883	.388	10.015	**< .0001**
Quadratic	20199	339	**< .0001**	–2.042	.293	–6.967	**< .0001**
Cubic	20184	15	**.0002**	–.320	.289	–1.106	.2706
Quartic	20179	5	**.0219**	.311	.222	1.403	.1628
Condition: Linear	20091	88	**< .0001**	1.981	.831	2.384	**.0185**
Condition: Quadratic	20091	0	.9515	–.327	.629	–.520	.6040
Condition: Cubic	20062	29	**< .0001**	–1.117	.618	–1.808	.0728
Condition: Quartic	20060	2	.1771	–.393	.477	–.823	.4119

Using cluster permutation analysis (Maris & Oostenveld, 2007; Wendt, Brand, & Kollmeier, 2014) time windows in which the groups significantly diverged were identified (highlighted in blue/grey in Figure 5d-f). Following this analysis, one time-cluster for happy (1680 – 3000ms, p = .010) and one for angry (1080-2760ms, p < .001), but no neutral clusters, were identified (see Figures 6 and 7 for divergence analysis for the other expressions).

Summary

Here condition differences in face-scanning were further investigated for happy and angry expressions using statistical heatmap comparisons and mixed-effects modelling. These techniques enabled eye-tracking data to be analysed without collating across spatial regions (heatmaps) or across time (mixed-effects modelling, cluster permutation analysis). The heatmap analyses indicate that condition had little effect on scanning for the neutral trial, but clear differences emerged for happy and angry trials, with dynamic expressions attracting greater lower face looking. These findings are consistent with the previous AOI analyses, but here precise spatial regions of significant difference are identified, for instance increased mouth and chin looking for dynamic anger and increased mouth and nose looking for dynamic happiness.

Mixed effects modelling and cluster permutation analyses were then used to describe the time course of this lower face bias for dynamic happy and angry expressions. Infants in both conditions began fixating the eye-region, and lower face looking then rose as infants scanned other regions of the face. However, at approximately the midpoint of the expressive trials (1500ms), the analyses identified a significantly higher peak in lower face looking for the dynamic condition as the emergence of the expression drove infants' visual attention toward lower facial regions (see this illustrated in the dynamic heatmaps in Video Figures 2 and 3).

Further differences between conditions

Thus far heatmap and temporal analyses have been presented for one actor only, who performed neutral, happy and angry expressions. Figures 6 and 7 give an overview of heatmap and temporal analyses for the other expressions. In the temporal plots within these figures, cluster permutation analyses were used to identify time-windows of significant differences between groups. Firstly, in Figure 6a, we see an averaged map for the six **neutral** trials, with Figure 6d depicting the pattern of lower face looking to the neutral face trials over time. From these Figures we can see that neutral face scanning shows comparatively even looking across space (Figure 6a) and time (Figure 6d) for both dynamic and static trials.

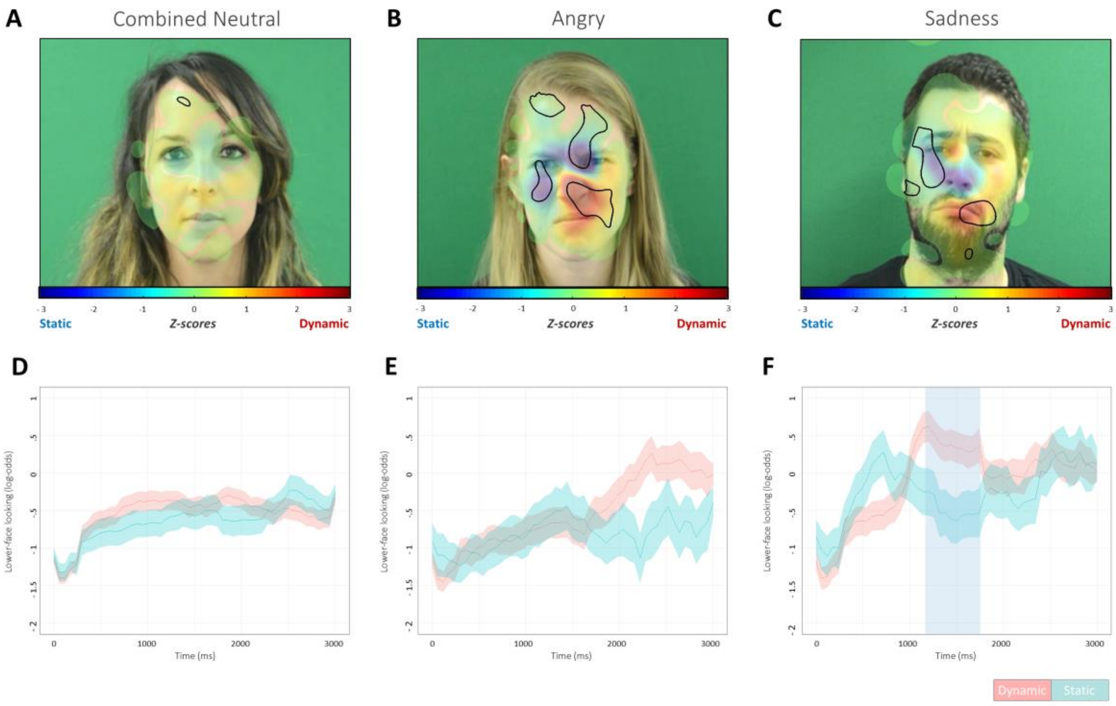

Figure 6. Heatmaps and time-course analyses for combined neutral, angry and sad expressions.

Statistical heatmaps (A-C) and time-course analyses (D-F) for the six neutral trials combined (A, D), an alternative angry trial (B, E) and a sad trial (C, F). In graphs D to E, the log-likelihood of lower-face looking is plotted across time, dynamic in red, static in blue. Time windows of significant differences between groups that were identified using cluster permutation analyses are highlighted in grey. See Figure 5 for more information.

The alternate **anger** trial (Figure 6b and 6e) to the example given earlier illustrates the trial-specific nature of infant expression scanning. For this actor's angry expression, there is a general trend of increased lower looking for dynamic (Figure 6e; red) as reported previously, but this dynamic region was located comparatively higher, and to the right (observer perspective) of a static region (blue). As such, no significant time-window of greater lower-looking emerges for the dynamic trial according to the pre-defined facial midpoint (bridge of the nose). This example illustrates how idiosyncrasies present within naturalistic dynamic expressions can have a strong effect on infant face scanning.

The **sad** expression trial (Figure 6c and 6f), also depicts idiosyncratic differences between conditions. This actor's performance of the sad expression included a slight head tilt toward the right. This slight difference in the location of facial features within the dynamic trial led to significant left-right condition differences which are depicted in Figure 6c. The temporal analyses, however, highlight a time-window of significant difference between conditions (1140 – 1800ms, *p* = .04), with a later and slightly higher peak in lower face looking for the dynamic condition. A similar pattern of a more pronounced later peak in lower face looking is also evident for **surprise** (1500 – 1980ms, *p* = .03; Figure 7a and 7d).

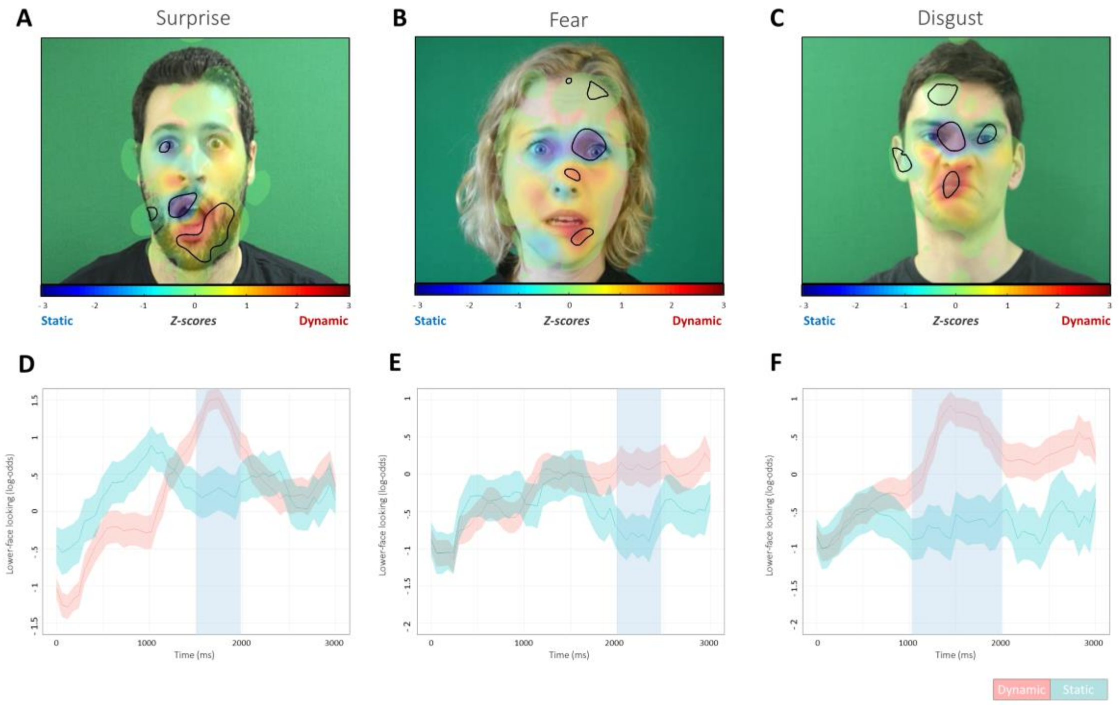

Figure 7. Heatmaps and time-course analyses for surprise, fear and disgust expressions.

Statistical heatmaps (A-C) and time-course analyses (D-F) for surprised (A, D), fearful (B, E) and disgusted (C, F) sample trials. See Figure 5 for more information.

For **fear** (Figure 7b and 7e), increased lower face looking is again evident for the dynamic expression. The eye-region is considered most diagnostic for fear (Jack, Garrod,

& Schyns, 2014; M. L. Smith et al., 2005), so overall the pattern of looking toward the lower face is flatter compared to the other expressions. However, there are clusters of significantly higher looking toward the eye-region for static and around the mouth for dynamic (Figure 7b). The temporal analyses (Figure 7e) indicate that while there is a noticeable drop in lower face looking for static as infants re-fixate the eye-region (1980 – 2460, $p < .05$), the dynamic lower face remains salient. For **disgust**, there is a clear lower face bias for dynamic (1080 – 2040, $p = .02$), with looking toward the static expression remaining predominantly in the eye-region.

To conclude, it is clear that the *where* and *when* of infant facial expression scanning is highly dependent on whether the expression stimuli are dynamic or static. Whilst particular patterns of looking are clearly modulated by differences between actors and the idiosyncrasies of their expression performance, there is also a general trend of increased looking toward the lower face for dynamic expressions.

Discussion

For each of the complementary analyses above (AOI, heatmap, time-course), there are significant effects of stimulus type on infant facial expression scanning. Replicating classic findings (B. M. Wilcox & Clayton, 1968), dynamic stimuli hold infant’s attention longer than static stimuli. It was also found that the lower face and its features becomes much more salient when an expression is dynamic. In particular, the AOI analyses highlighted a pronounced lower face bias for dynamic happy and angry expressions. These two expressions were then further explored using statistical heatmaps and mixed-effect modelling to provide a more detailed picture of how infants scan static and dynamic expressions differently.

This experiment demonstrates that when expressions are dynamic, infant scanning patterns are *reactive*, with attention being driven toward the lower face as expression-related changes emerge. As one would predict, face scanning for static expressions is less varied across time. Lower face looking for static expressions also peaks earlier in the trial and remains lower overall. Dynamic stimuli provide additional diagnostic motion cues which are disproportionately communicated through the lower facial features (relative to upper facial features; see Krumhuber et al., 2013 for a review), which enhances the salience of this facial area in comparison to static stimuli. For instance, work with adult participants has highlighted the diagnostic value of lower facial action units (e.g. “Lip Stretcher” or “Upper Lip Raiser”) for categorising dynamic fear and angry expressions (Jack et al., 2014), while categorising static fear and anger relies almost entirely on the eye region (M. L. Smith et al., 2005). The findings in this chapter are consistent with this literature, as infants show greater looking toward the lower face for dynamic compared to static expressions.

From the findings in this chapter, it is clear that infants scan static and dynamic facial expressions differently. Infants would rarely, if ever, experience a still expression, frozen at peak expressive amplitude, and such an occurrence would likely be recognised as unnatural by infants (Adamson & Frick, 2003; Bertin & Striano, 2006). It is therefore unsurprising that this and other recent work (Addabbo et al., 2018; Heck et al., 2016; Richoz et al., 2018; Soussignan et al., 2017) have found different results with dynamic stimuli, calling into question the generalisability of much of the work in this field and necessitating the development of methods that enable researchers to use dynamic social stimuli.

Alongside these general differences, this experiment also demonstrates that particular patterns of infant face scanning are highly sensitive to individual differences in facial morphology and expression production (e.g. compare scanning differences for anger trials; Figure 5b and 6b). Recent work has shown that the traditionally held view of a "default" scanning pattern between the eyes and mouth is not accurate. This pattern is an artefact of averaging across participants who instead show stable but idiosyncratic and task-specific face scanning strategies (Arizpe, Walsh, Yovel, & Baker, 2017; Kanan, Bseiso, Ray, Hsiao, & Cottrell, 2015; Mehoudar, Arizpe, Baker, & Yovel, 2014). The lack of a default "triangular" scanning pattern is especially true when observing dynamic faces, where motion cues direct attention to salient regions of the face (Lewkowicz & Hansen-Tift, 2012; Võ et al., 2012). Our work is consistent with these findings; infants demonstrate "bespoke" scanning patterns which are dependent not only on the facial expression being observed, but also on the actor's individual facial morphology and expression performance.

This chapter has also demonstrated several methodological advances such as dynamic AOIs, statistical heatmaps and time-course analyses. These techniques enable researchers to present more naturalistic and ecologically-valid stimuli to infants, and have thus enabled the investigation of how responses to biologically relevant stimuli (such as those used here) unfold early in ontogeny. The combination of naturalistic stimuli, high-quality eye-tracking data and sophisticated analyses are all important advances, although it is evident from the data presented here that succinct reporting of results and interpretation is challenging. Yet we should not be surprised or deterred by this fact; social interactions are inherently complex and accordingly we should anticipate the need for equally complex solutions to reliably dissect and understand such behaviours. Nevertheless, identifying and/or creating effective methods and analysis tools that permit us to quantify and understand complex behaviours should be a challenge that we all relish as researchers. It is imperative from both an intellectual and an ethical perspective that we strive to produce results in our labs that can generalise outside to the world in which our participants live and interact. Future studies should therefore carefully consider which methodologies are most relevant and informative to their study design and research questions.

Chapter 6

The development of dynamic facial expression scanning

Abstract

This chapter explores the development of expression scanning from six to twelve months of age. The previous chapter demonstrated that infants’ scan dynamic and static facial expressions differently, and demonstrated useful methodological tools for analysing eye-tracking data. In this chapter these tools are used to investigate age-related differences in infants’ *interest* and *scanning* for the six basic facial expressions. Previous work with static stimuli suggests that seven-months of age is a key developmental threshold for facial expression perception. The results from this chapter suggest that infants across all age groups show differential interest and scanning for expressions, and demonstrate precise looking toward ‘diagnostic’ regions. Nevertheless, the analyses also suggest subtle age-related differences, including relative increases in interest and scanning precision toward negative facial expressions (e.g. anger, fear disgust). Yet, a question remains as to how informative these differences in looking patterns are about infants’ *understanding* of an expression’s meaning.

Introduction

The previous chapter found that infants scan static and dynamic facial expressions differently, and thus concluded that a transition to more naturalistic, dynamic facial expression stimuli is warranted. The focus was therefore on differences between conditions (i.e. dynamic versus static), however highly significant differences in scanning across the expressions (happy, sad, surprise, fear, anger, disgust) were also reported.

Consequently, in the current chapter the data from Chapter 5 will be further analysed, focusing on the development of dynamic facial expression scanning across the first year of life.

Using static stimuli and variations of the standard visual preference paradigm, researchers have uncovered much about how our ability to recognise facial expressions develops across ontogeny. Infants are able to perceive differences in (*discriminate*) certain facial expressions (e.g. happy vs fear) from birth (Farroni et al., 2007; Field et al., 1982). By seven months, infants can identify many of the six 'core' expressions (Ekman, 1993; Ekman et al., 1987) within discreet emotion categories (*categorisation*); being able to generalise across individuals or variations in expressive amplitude (Bornstein & Arterberry, 2003; Kotsoni, Haan, & Johnson, 2001; Kuchuk et al., 1986; Nelson et al., 1979; Ruba et al., 2017). At this age they also develop several other hallmarks of perceptual maturity such as orientation specificity (Kestenbaum & Nelson, 1990), cultural diversity (Geangu et al., 2016) and an adaptive preference for fearful expressions (Peltola et al., 2013). After seven months, the ability to perceive facial emotion is further fine-tuned (see Leppänen & Nelson, 2009 for a review), including being able to use affective expressions to inform about objects or events in their environment (*social referencing*; see Walden & Ogan, 1988).

Previous work using static stimuli has therefore identified seven months of age as an important developmental threshold at which infants transition toward an adult-like ability to process facial expressions. Yet recent work suggests that sophisticated processing of facial emotion may emerge earlier in infancy when more realistic, dynamic displays are used (Addabbo et al., 2018; Heck et al., 2016; Montague & Walker-Andrews, 2001; Soussignan et al., 2017). Dynamic faces are scanned differently by infants (N. G.

Xiao et al., 2015), and one study has reported developmental differences in diagnostic scanning for dynamic expressions (Soussignan et al., 2017). Following this work, this current chapter will explore whether six-, nine- and twelve-month-olds differ in their interest and scanning of dynamic facial expressions.

Methods

Analysis plan

This chapter further analyses the data from Chapter 5. Details about participant information (see Table 1), stimuli and methodological procedures can be found in Chapter 5. As in the previous chapter, Dynamic AOIs, statistical heatmaps and time-course analyses will be used, but here we will investigate differences across Age (six, nine and twelve months) and Expression (happy, sad, surprise, fear, anger, disgust), using data from the dynamic condition only (*N* = 101, see Table 1). This chapter will begin with a collapsed analysis of looking time toward individual facial regions using the Dynamic AOIs. Differences in *interest* (face) and *scanning* (lower face and feature looking) across expressions will be reported, with a particular emphasis on how these looking patterns change from six to twelve months. Temporal analyses and Dynamic AOIs will then be combined to model developmental changes in infant looking toward the face (interest) and the lower-face features (scanning). As before, 31 infants were excluded from the collapsed analysis for not possessing fixation duration data for all expression trials, but were retained for the heatmap and time-course analyses.

Table 1. Participant information for the initial and reduced cohorts (see Chapter 5)

	Age (months)	**Age M (SD) (days)**	**N Dynamic Condition**
Initial	6	191.17 (9.78)	33
	9	274.77 (12.69)	32
	12	366.76 (13.42)	36
Reduced	6	192.51 (15.32)	28
	9	273.78 (13.73)	20
	12	366.24 (11.62)	29

Within these developmental models, Age, Expression and Time will be used as predictors. Given the complexity of these models, the analysis will report firstly intercept-only differences across Time, Age and Expression. **Time**, defined by orthogonal power polynomials, describes the general pattern of looking across the trial. **Age** identifies general differences in AOI attention for two age comparisons: 1. Six-months versus nine- and twelve-months combined, and 2. Nine-months versus twelve-months. **Expression** compares overall looking durations toward each of the six expressions against neutral. Following this, the analysis will then consider any developmental differences in looking patterns across Time regardless of Expression (Age x Time interactions), and across Expression regardless of Time (Age x Expression interactions). Given the focus on developmental differences, reporting Expression x Time interactions are beyond the scope of these analyses and Age x Expression x Time interactions will only be reported for significant Age x Expression interactions. Information regarding how each additional parameter and interaction improves model fit (based on the Δ*D*) will be included within model summary tables.

Following the *temporal* analyses, statistical heatmaps will then be used to investigate fine-grained *spatial* differences in face scanning across age groups. Firstly, descriptive heatmaps summarising mean looking across the face for each individual age group will be produced. Secondly, one-way ANOVA maps will be created depicting significant regions of difference across the three age groups (six, nine and twelve months). Finally, post-hoc Bonferroni-adjusted t-test maps will be used to conduct direct comparisons between age groups (nine minus six; twelve minus nine; twelve minus six), and highlight particular facial regions which show developmental differences in face scanning.

As mentioned above, previous work has identified seven months of age as a key transition period in facial expression perception. At this age infants are thought to not only show greater precision discriminating and categorising emotional expressions, which in adults is associated with *scanning* diagnostic regions (Gosselin & Schyns, 2001), but also to show greater *interest* in negatively-valenced expressions such as fear, anger and disgust. It is therefore predicted that these patterns will emerge within this investigation of developmental differences in dynamic facial expression *interest* and *scanning*.

Results

AOI analysis

Face. A 6 (Expression) x 3 (Age) ANOVA for proportional looking durations toward the face AOI yielded significant main effects for Expression ($F(4.1,303.7) = 7.130$, $p < .001$, $\eta_p^2 = .088$), Age ($F(2,74) = 5.05$, $p = .009$, $\eta_p^2 = .120$), and a significant Age x Expression interaction ($F(8.2,303.7) = 2.059$, $p = .038$, $\eta_p^2 = .053$). Pairwise comparisons

(Bonferroni) indicate greater face-looking for surprise (*M* = 96.54%) in comparison to all other expressions (Combined *M* = 91.01%, all $p < .03$), but also for fear (*M* = 93.10%) compared to anger (*M* = 88.99%, $p = .045$). Overall, six-month-olds spent proportionally less time on the face (*M* = 88.89%) compared to nine (*M* = 93.39%, $p = .046$) and 12-month-olds (*M* = 93.65%, $p = .015$). Planned contrasts indicate that age differences emerged for the angry expression ($F(2,74) = 8.36$, $p = .001$, $\eta_p^2 = .184$), with significantly higher looking for nine ($p = .004$) and twelve-month-olds ($p = .001$) compared to six-month-olds. A significant Age effect also emerged for sadness ($F(2,74) = 3.387$, $p = .039$, $\eta_p^2 = .084$) and a marginally significant effect for disgust ($F(2,74) = 2.968$, $p = .058$, $\eta_p^2 = .074$), again following the trend of reduced face-looking in six-month-olds.

Lower face and features. Two 6 (Expression) x 3 (Age) ANOVAs were used to investigate proportional looking toward the lower face and lower-face features. A significant main effect of Expression was found for both lower face ($F(5,370) = 16.474$, $p < .001$, $\eta_p^2 = .182$) and feature ($F(5,370) = 12.732$, $p < .001$, $\eta_p^2 = .147$) ANOVAs, though no significant main effects of Age (both $p > 23$) or interactions with Age were found (both $p > 23$). Performing the same analyses for upper face and features also found no significant main effects of Age (both $p > .47$) or Expression x Age interactions (both $p > .52$). Post-hoc tests reveal that surprise attracted more lower-face looking (Surprise *M* = 57.87%) than all other expressions (Combined *M* = 38.98%, all $p < .001$), and more lower-feature looking (Surprise *M* = 29.88%) compared to all other expressions except happiness (Combined *M* = 17.09%, Happiness $p = .159$, all other $p < .01$). Fear also attracted less lower-feature looking (Fear *M* = 12.13%) in comparison to disgust (Disgust *M* = 19.97%, $p = .024$) and happiness (Happiness *M* = 22.71%, $p = .002$).

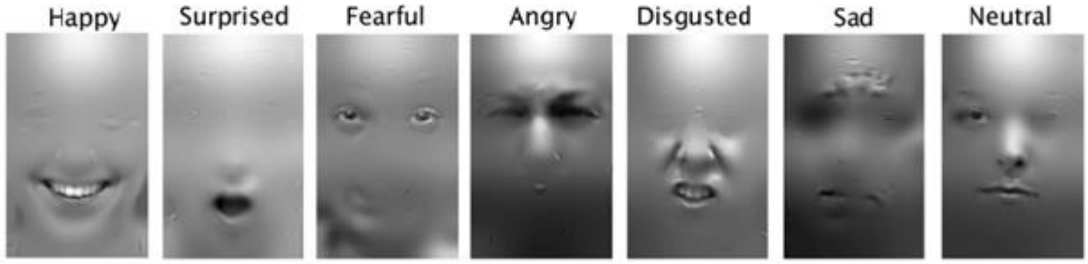

Figure 1. Diagnostic regions for identifying static facial expressions in adult human observers.

Adapted and reproduced from Smith and colleagues (2005).

Summary

The above analysis indicates that infants show developmental differences in expression *interest* (face looking time), but expression *scanning* (lower face and feature looking time) is largely determined by the stimulus, with few differences across age. Overall, infants showed the most interest in the surprised face, and relatively less interest in the angry face. There was also clear looking toward diagnostic regions across all age groups (Smith, Cottrell, Gosselin, & Schyns, 2005), with surprise attracting the greatest amount of lower feature looking, followed by happiness and disgust, while fixations for the fearful expression were largely confined to the upper features (c.f. Figure 1). Nevertheless, compared to older infants, six-month-olds spent less time on the face overall, and less time looking at sad, angry and disgusted faces in particular.

Temporal dynamics

The AOI analyses found developmental differences in expression *interest* but not *scanning*. Here these results are explored in more detail by modelling the temporal profile of face interest (Table 2), and lower face and feature scanning (Tables 3 and 4) for all participants within the dynamic condition (N = 101) using the template formula below (Wilkinson notation). Trials without fixation data for the stimulus region were automatically discarded from these analyses.

$$AOI \sim Age * Expression * Time \; + \; (1 + Time \mid Participant/Trial)$$

The dependent variable was the empirical log-odds of proportional fixation durations within each AOI (either face, lower face or lower-face feature looking). All three models used Age, Expression and Time as fixed effects. Each expression was compared against neutral, while two contrasts were created for Age; the first comparing the average of nine and twelve-month-olds against six-month-olds (C1: 6M = -.66, 9M = .33, 12M = .33), and the second comparing twelve-month-olds against nine-month-olds (C2: 6M = 0, 9M = -.5, 12M = .5). Time was defined by three slope terms; linear, quadratic and cubic for face looking. A fourth term (quartic) was added for lower face and lower-face features to capture the steepness of the curves within these data (see Figure 2).

A maximal random effects structure was attempted, including nested Participant and Trial random effects for the intercept, Expression and all Time polynomials. However, including Expression as a parameter lead to non-convergence in all models and it was therefore removed (see Baayen, Davidson, & Bates, 2008; Barr, Levy, Scheepers, & Tily, 2013).

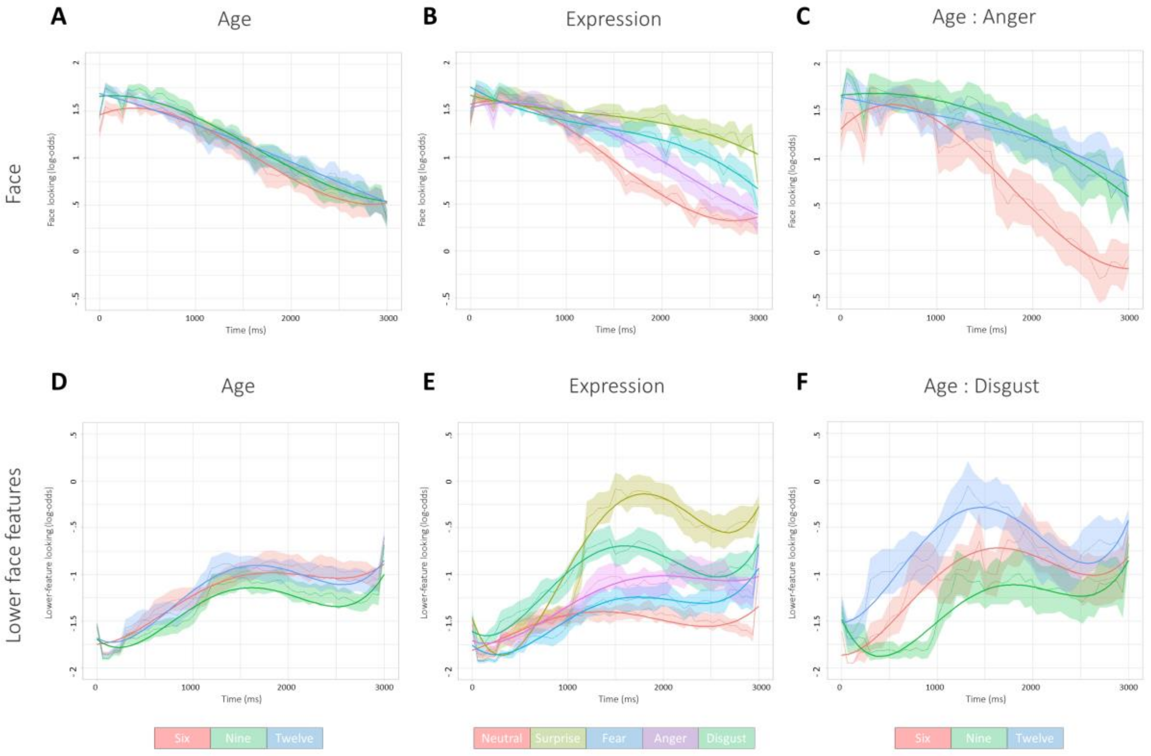

Figure 2. Temporal differences in face and lower-feature looking across Age and Expression.

The log-likelihood of fixating the lower face was computed across 60ms time-bins, plotted here (faint line) with the standard error of the mean (coloured border). Three polynomial terms were used to model face-looking across time (linear, quadratic and cubic), and four for lower-feature looking (linear, quadratic, cubic and quartic). The log-likelihood of fixating face **(A-C)** and lower-features **(D-F)** is plotted divided by Age: six (red), nine (green) and twelve (blue) months, and Expression: neutral (red), surprise (yellow/green), fear (blue), anger (purple) and disgust (green). For clarity, disgust has been omitted from the face-looking Figure **(B)**. In **C** and **F**, data from two significant Age x Expression interaction effects are plotted, displaying developmental differences in face looking for the angry expression **(C)** and lower-feature looking for the disgusted expression **(F)**.

Temporal dynamics: Face

All parameters significantly improved model fit with the exception of Age, though Age interactions with Expression and Time polynomials did improve fit (Table 2). Overall, we can see that there was a substantial negative linear effect of Time (β = -3.418, *t* = -15.614) on face looking, with flattened tails on each end (significant cubic effect). This effect suggests that infants' interest in the face (across all Age groups) declined over time (negative linear effect), though face-looking at the start and end of trials was comparatively flatter (cubic effect). This pattern can be viewed in Figure 2a.

There were significant positive effects of all expressions compared to neutral (all *p* < .02), though expressions varied in the size of their effect. Heightened interest in the face was strongest for surprise (β = .445, *t* = 5.942), and weakest for anger (β = .188, *t* = 2.571, see Figure 2b). Relative to older infants (Age C1: 6M vs 9 and 12M), six-month-olds show greater interest in the surprised face compared to neutral (Age C1 x Surprise: β = -.413, *t* = -2.554, *p* = .011), and less interest in the angry face compared to neutral both overall (Age C1 x Anger: β = .398, *t* = 2.537, *p* = .011) and linearly across time (Age C1 x Anger x Linear: β = 2.510, *t* = 3.754, *p* < .001, Figure 2c). A marginally significant Age x Expression interaction also suggests twelve-month-olds showed greater interest in the disgusted face compared to nine-month-olds (Age C2 x Disgust: β = .332, *t* = 1.853, *p* = .064).

Table 2. Model fit, estimates and t-values for developmental differences in face looking

	Model Fit			Parameter estimates			
Fixed Effect	-2LL	*ΔD*	*P*	*β*	*SE*	*t*	*P*
Intercept	300360	–	–	.961	.055	17.563	**< .0001**
Time							
Linear	295668	4692	**< .0001**	–3.418	.219	–15.614	**< .0001**
Quadratic	295559	109	**< .0001**	.022	.137	.159	.8740
Cubic	295519	40	**< .0001**	.635	.114	5.577	**< .0001**
Age	295519	0	.8068				
C1: (9 + 12M) vs 6M				.072	.118	.611	.5421
C2: 12M vs 9M				–.063	.134	–.473	.6365
Expression	294347	1172	**< .0001**				
Happy				.295	.072	4.090	**< .0001**
Sad				.334	.074	4.488	**< .0001**
Surprise				.445	.075	5.942	**< .0001**
Fear				.301	.074	4.043	**< .0001**
Anger				.188	.073	2.571	**.0102**
Disgust				.361	.074	4.861	**< .0001**
Age: Time	294227	120	**< .0001**				
Expression: Time	293423	804	**< .0001**				
Age: Expression	292895	528	**< .0001**				
Age C1: Happy				.002	.155	.012	.9906
Age C1: Sad				.043	.159	.268	.7889
Age C1: Surprise				–.413	.162	–2.554	**.0108**
Age C1: Fear				.044	.160	.277	.7821
Age C1: Anger				.398	.157	2.537	**.0113**
Age C1: Disgust				–.122	.161	–.760	.4473
Age C2: Happy				–.084	.176	–.476	.6340
Age C2: Sad				.177	.183	.966	.3340
Age C2: Surprise				.047	.182	.258	.7961
Age C2: Fear				.247	.182	1.360	.1741
Age C2: Anger				.001	.179	.008	.9939
Age C2: Disgust				.332	.179	1.853	.0641
Age: Expression: Time	292499	396	**< .0001**				
Age C1: Surprise: Linear				–.729	.690	–1.056	.2909
Age C1: Surprise: Quadratic				.769	.563	1.366	.1720
Age C1: Surprise: Cubic				.222	.449	.493	.6221
Age C1: Anger: Linear				2.510	.670	3.754	**.0002**
Age C1: Anger: Quadratic				.596	.546	1.092	.2752
Age C1: Anger: Cubic				–.643	.436	–1.475	.1404

N: Participants = 101, Trials = 1659, Observations = 84609

Temporal dynamics: Lower features

The results from the lower-face feature model (Table 3) indicate a more complex pattern across time (Figure 2d). Overall, there was a general increase in lower-feature looking across time (Linear: β = .440, *t* = 2.400), with an ‘inverted U’ pattern (Quadratic: β = -.467, *t* = -3.016). Significant third and fourth order temporal terms for this model largely reflect the asymptotic tails at the beginning and end of the time window (see Mirman et al., 2008). Lower-feature looking was also not significantly predicted by Age (all *p* > .37), or any Age x Time interactions (all *p* > .26, Figure 2d). This suggests that, like face interest, the general pattern of face scanning showed few developmental differences across time. Infant looking toward the lower features was low at the start of the trial (the time window began with a look to the eye region), and peaked toward the midpoint as the dynamic expression emerged, before reducing again toward the end of the trial.

However, this pattern of infant face scanning varied according to Expression (Figure 2e). Expressions with diagnostic lower features, such as surprise, happy and disgust (see Figure 1), showed strong positive effects compared to neutral (Surprise: β = .731, Happiness: β = .485, Disgust: β = .469; all *p* < .001), while expressions with diagnostic upper-face features showed much weaker effects compared to neutral (Fear: β = .091, *p* = .238; Anger: β = .264, *p* < .001). Compared to older infants, six-month-olds showed greater interest in the lower features of surprise overall (Age C1 x Surprise: β = -.341, *t* = -2.061, *p* = .039), and across time (Age C1 x Surprise x Linear: β = -1.815, *t* = -2.578, *p* = .010; Age C1 x Surprise x Quartic: β = .896, *t* = 2.198, *p* = .028). This pattern can be visualised in Figure 8biii. Also, twelve-month-olds showed greater looking toward the lower features of the disgusted expression compared to nine-month-olds (Age C2 x Disgust: β = .576, *t* = 3.122, *p* = .002), displaying a more pronounced ‘inverted U’ curve across time (Age C2 x Disgust x Quadratic: β = -1.504, *t* = 2.337, *p* = .020). This can be viewed in Figure 2f.

Table 3. Model fit, estimates and t-values for developmental differences in lower-feature looking

	Model Fit			Parameter estimates			
Fixed Effect	-2LL	*ΔD*	*P*	*β*	*SE*	*t*	*P*
Intercept	250123	–	–	−1.511	.056	−27.002	**< .0001**
Time							
Linear	247991	2132	**< .0001**	.440	.183	2.400	**.0168**
Quadratic	247340	651	**< .0001**	−.467	.155	−3.016	**.0027**
Cubic	247339	1	.3340	.471	.130	3.614	**.0003**
Quartic	247085	254	**< .0001**	.248	.110	2.265	**.0241**
Age	247083	2	.3038				
C1: (9 + 12M) vs 6M				.035	.121	.288	.7734
C2: 12M vs 9M				.123	.137	.898	.3704
Expression	244684	2399	**< .0001**				
Happy				.485	.074	6.520	**< .0001**
Sad				.305	.076	4.002	**< .0001**
Surprise				.731	.077	9.494	**< .0001**
Fear				.091	.077	1.182	.2375
Anger				.264	.075	3.500	**.0005**
Disgust				.469	.076	6.172	**< .0001**
Age: Time	244618	66	**< .0001**				
Expression: Time	243212	1406	**< .0001**				
Age: Expression	242811	401	**< .0001**				
Age C1: Happy				.012	.160	.074	.9408
Age C1: Sad				−.059	.163	−.364	.7161
Age C1: Surprise				−.341	.165	−2.061	**.0394**
Age C1: Fear				−.268	.165	−1.623	.1049
Age C1: Anger				−.011	.162	−.066	.9471
Age C1: Disgust				.030	.164	.184	.8539
Age C2: Happy				.024	.181	.132	.8951
Age C2: Sad				−.191	.187	−1.018	.3089
Age C2: Surprise				.065	.188	.345	.7305
Age C2: Fear				.137	.188	.730	.4656
Age C2: Anger				.144	.184	.780	.4358
Age C2: Disgust				.576	.185	3.122	**.0018**
Age: Expression: Time	242488	323	**< .0001**				
Age C1: Surprise: Linear				−1.815	.704	−2.578	**.0100**
Age C1: Surprise: Quadratic				−.632	.573	−1.104	.2700
Age C1: Surprise: Cubic				−.083	.507	−.164	.8700
Age C1: Surprise: Quartic				.896	.408	2.198	**.0281**
Age C2: Disgust: Linear				−.554	.797	−.696	.4866
Age C2: Disgust: Quadratic				−1.504	.644	−2.337	**.0196**
Age C2: Disgust: Cubic				.942	.569	1.656	.0980
Age C2: Disgust: Quartic				.054	.457	.118	.9060

N: Participants = 101, Trials = 1659, Observations = 72641

Temporal dynamics: Lower face

Results for the lower face overall largely mirror those for lower features (see Table 4). However, we see no significant linear increase in lower-face looking across time (p = .37), but do see the inverted parabola quadratic pattern (p < .001). The natural increase in lower-face looking across time was therefore likely offset by reduced face interest in general (lower-face looking was calculated as a percentage of total looking, not of face looking only). However, there was still a midpoint peak in lower-face looking coinciding with the emergence of the dynamic expression (quadratic effect). Like in the lower-feature analyses, there were also no significant effects of Age (both p > .10), or Age x Time interactions (all p > .17).

For lower face, there were significant positive effects of all expressions compared to neutral (all p < .007). Expressions with diagnostic lower regions showed large effects (Surprise β = .773, Happiness β = .543, Disgust β = .512), while expressions where the lower face is less useful for decoding expressions showed smaller effects (Fear β = .287, Anger β = .255, Sadness β = .235). Lower face also did not show a significant Age C2 x Disgust interaction (β = .196, t = .931, p = .352). Instead, lower face showed significant Age C1 interactions with Anger (β = .429, t = 2.302, p = .021) and Disgust (β = .692, t = 3.721, p < .001). Six-month-olds therefore showed reduced overall looking toward the lower face of angry and disgusted expressions compared to older infants.

Table 4. Model fit, estimates and t-values for developmental differences in lower-face looking

	Model Fit			Parameter estimates			
Fixed Effect	-2LL	*ΔD*	*P*	*β*	*SE*	*t*	*P*
Intercept	352802	–	–	–1.039	.078	–13.320	**< .0001**
Time							
Linear	352335	467	**< .0001**	–.188	.210	–.895	.3715
Quadratic	351203	1132	**< .0001**	–.830	.162	–5.128	**< .0001**
Cubic	351046	157	**< .0001**	.856	.130	6.599	**< .0001**
Quartic	350936	110	**< .0001**	.052	.128	.403	.6874
Age	350934	2	.3691				
C1: (9 + 12M) vs 6M				–.272	.168	–1.621	.1071
C2: 12M vs 9M				.146	.190	.767	.4443
Expression	348458	2478	**< .0001**				
Happy				.543	.087	6.283	**< .0001**
Sad				.235	.087	2.721	**.0066**
Surprise				.773	.087	8.933	**< .0001**
Fear				.287	.087	3.323	**.0009**
Anger				.255	.087	2.947	**.0033**
Disgust				.512	.087	5.926	**< .0001**
Age: Time	348421	37	**< .0001**				
Expression: Time	347497	924	**< .0001**				
Age: Expression	347005	492	**< .0001**				
Age C1: Happy				.252	.186	1.356	.1752
Age C1: Sad				.205	.186	1.100	.2714
Age C1: Surprise				–.023	.186	–.123	.9019
Age C1: Fear				.100	.186	.539	.5898
Age C1: Anger				.429	.186	2.302	**.0214**
Age C1: Disgust				.692	.186	3.721	**.0002**
Age C2: Happy				.040	.211	.174	.8617
Age C2: Sad				.242	.211	1.150	.2505
Age C2: Surprise				.053	.211	.251	.8020
Age C2: Fear				.172	.211	.813	.4161
Age C2: Anger				–.049	.211	–.231	.8174
Age C2: Disgust				.196	.211	.931	.3520
Age: Expression: Time	346753	252	**< .0001**				
Age C1: Anger: Linear				–.905	.741	1.221	.2222
Age C1: Anger: Quadratic				.766	.590	1.298	.1946
Age C1: Anger: Cubic				–.490	.496	–.989	.3229
Age C1: Anger: Quartic				–.158	.429	–.369	.7121
Age C1: Disgust: Linear				–.559	.741	–.754	.4510
Age C1: Disgust: Quadratic				.063	.590	.107	.9151
Age C1: Disgust: Cubic				.549	.496	1.108	.2681
Age C1: Disgust: Quartic				.587	.429	1.367	.1719

N: Participants = 101, Trials = 1818, Observations = 92718

Summary

In the above analyses developmental differences in expression *interest* and *scanning* were explored in greater detail by modelling infant looking to the face (interest) and lower face and features (scanning) across time. Firstly, the analysis found no general differences in face interest or scanning across age groups (Figure 2a and Figure 2d). Infants found all expressions of greater interest than neutral, though maintained relatively higher face-looking for surprise, and relatively lower face-looking for anger across time (Figure 2b). This pattern was exaggerated in the six-month age group, with overall higher interest in the surprised face (see Figure 8bi), and a much sharper decline in face interest for anger across time compared to older infants (Figure 2c).

Infants in all age groups showed clear diagnostic looking when scanning expressions (Figure 2e), with increased attention toward the lower-face features for expressions in which these features are highly informative (e.g. surprise, happiness and disgust). Conversely, lower-feature scanning was indistinguishable from neutral within expressions where there is little diagnostic information in the lower features (e.g. fear). Unlike the collapsed AOI analyses, there were also developmental differences in expression scanning. Compared to older infants, six-month-olds maintained interest in the wide mouth of surprise for longer (see Figure 8biii), and there was a transition toward greater lower-feature looking for disgust between nine and twelve months (Figure 2f). Additionally, when considering the lower face as a whole, it was found that six-month-olds show comparatively reduced lower-face looking for angry and disgusted expressions.

Heatmaps

Developmental differences in face scanning for five expression trials (anger, fear, disgust, happiness and sadness; Figures 3-7 respectively) are presented as heatmaps.

Firstly, fixation data (location and duration) collapsed across time were plotted as z-scores for each age group (left panel). A matrix ANOVA was then performed to investigate pixel-wise differences in face-scanning between ages. The *p* values (0 to .1) from this analysis are plotted in the large 'ANOVA map' within each figure. Post-hoc analysis *t*-maps (lower panel) were also used to compare between each of the three age groups. The post-hoc *t*-maps represent mean differences between the fixation age maps within each comparison, and regions of significant differences are outlined in black using a Bonferroni-corrected alpha (.016).

For **anger** (Figure 3), there were only small scattered regions emerging as significantly different across Age (see ANOVA map). Looking at the post-hoc maps, significant differences can be seen only within the twelve and six-month-old comparison. Twelve-month-olds showed much cleaner diagnostic looking toward the eyes and particularly the mouth (red), while six-month-olds fixated a more central region (blue). Similarly for **fear** (Figure 4), older infants showed increased looking toward diagnostic regions such as the whites of the eyes and open mouth. Again there was particularly precise looking in the twelve-month group, while six-month-olds showed significantly greater looking toward a more central facial region (see twelve-six post-hoc map). Figure 5 depicts the developmental differences for **disgust**. Twelve-month-olds showed greater looking toward the mouth-nose diagnostic region compared to nine-month-olds (c.f. Figure 2f). For **happiness** (Figure 6) a section of the mouth-region is fixated more in twelve-month-olds compared to six-month-olds, and for **sadness** (Figure 7) only a small section of the lower lip is fixated more in twelve-month-olds compared to nine-month-olds.

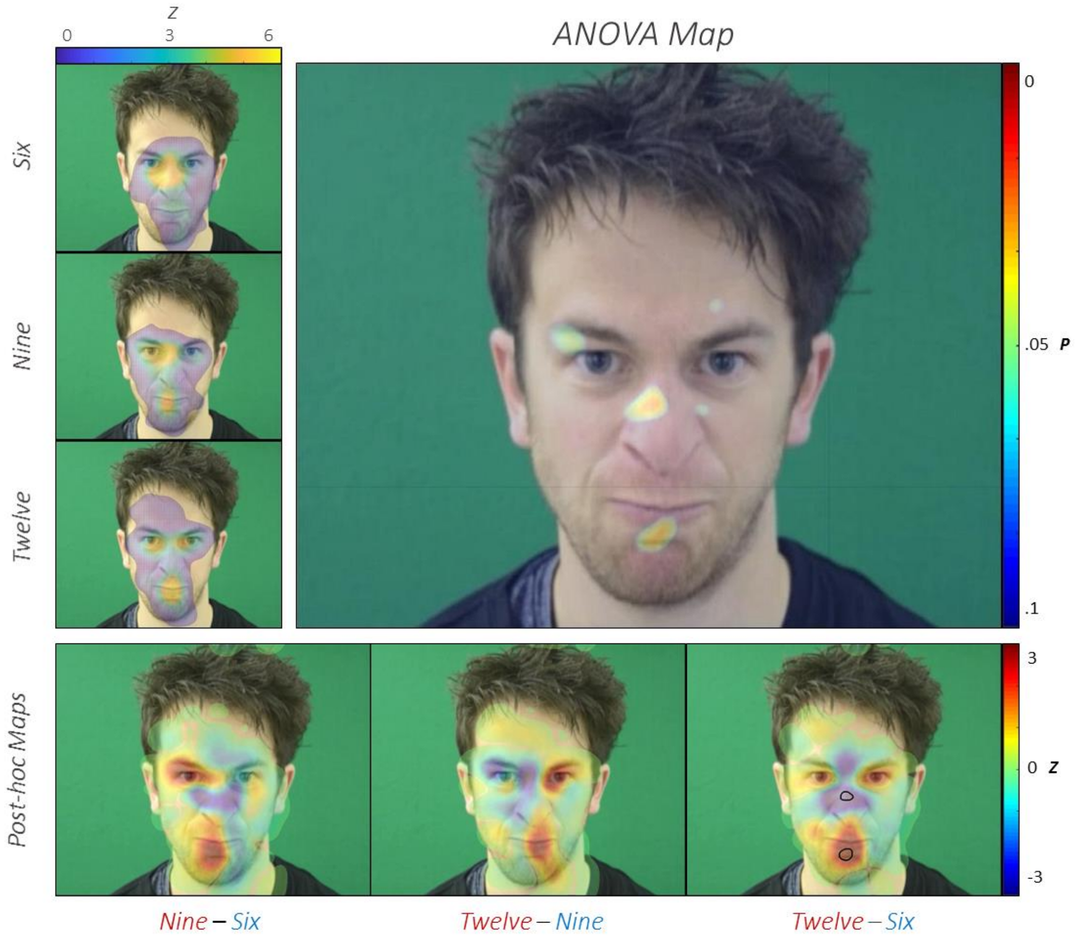

Figure 3. ANOVA heatmap analyses for an angry expression trial.

Fixation location and duration data are plotted as z-scores for each age group (six, nine and twelve months) in the left-side panel (‘Age maps’), with highly-fixated regions appearing orange/yellow. The p values (two-tailed) from 0 (red) to .1 (blue) from a matrix ANOVA across Age are presented in the large central Figure (‘ANOVA map’). In the lower panel, t-maps display the mean differences between post-hoc Age map comparisons (9 – 6 months, 12 – 9 months and 12 – 6 months) as z-scores (‘Post-hoc maps’), with positive (red) values denoting regions of increased looking in the older infants, and negative (blue) values denoting regions of increased looking in younger infants. Clusters of significant differences between Age maps are outlined in black using a Bonferroni-corrected alpha (.016).

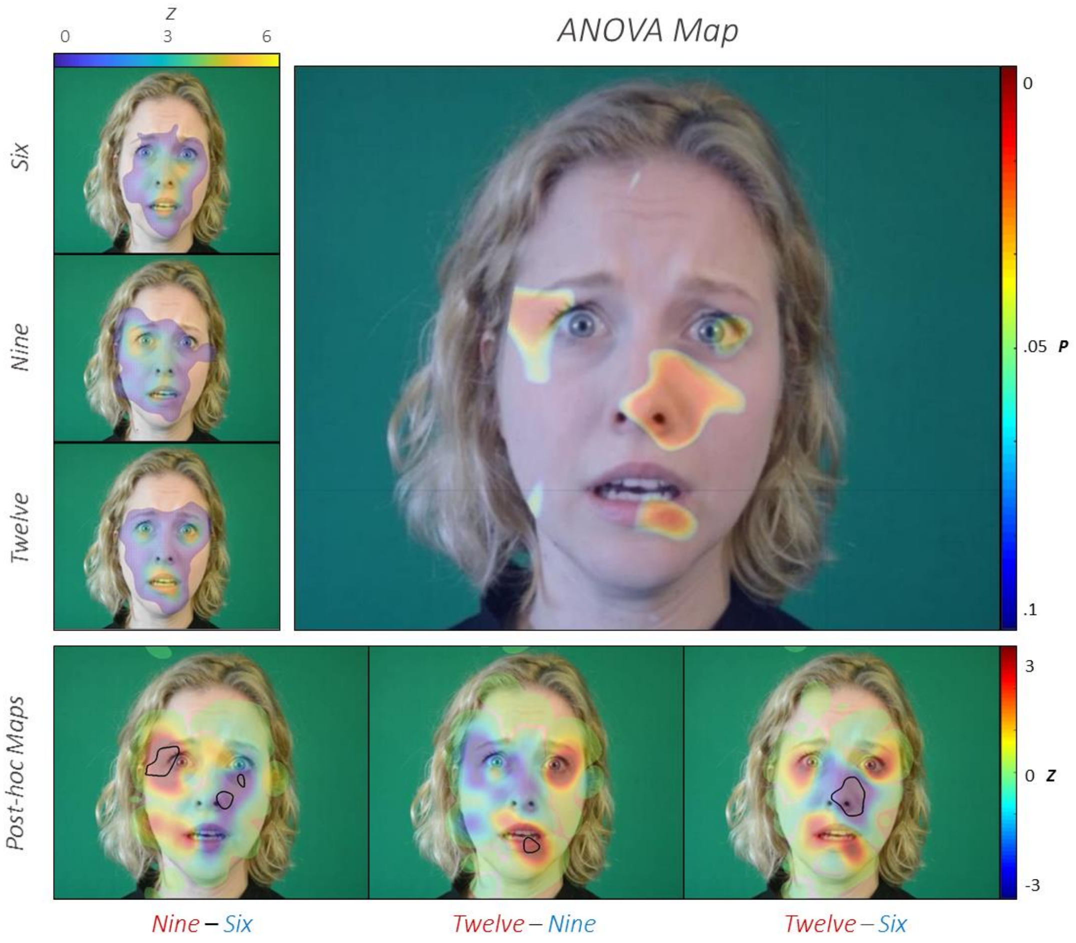

Figure 4. ANOVA heatmap analyses for a fearful expression trial.

See Figure legend 3 for further information.

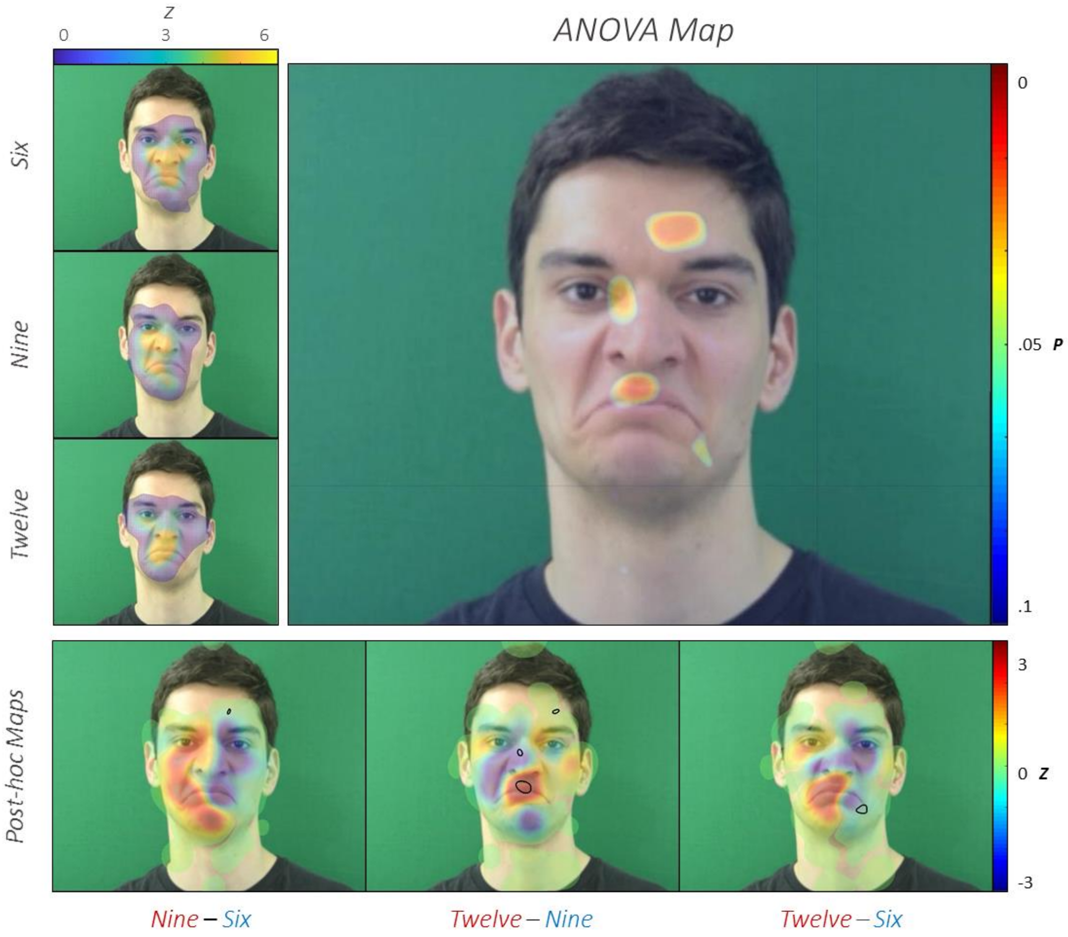

Figure 5. ANOVA heatmap analyses for a disgusted expression trial.

See Figure legend 3 for further information.

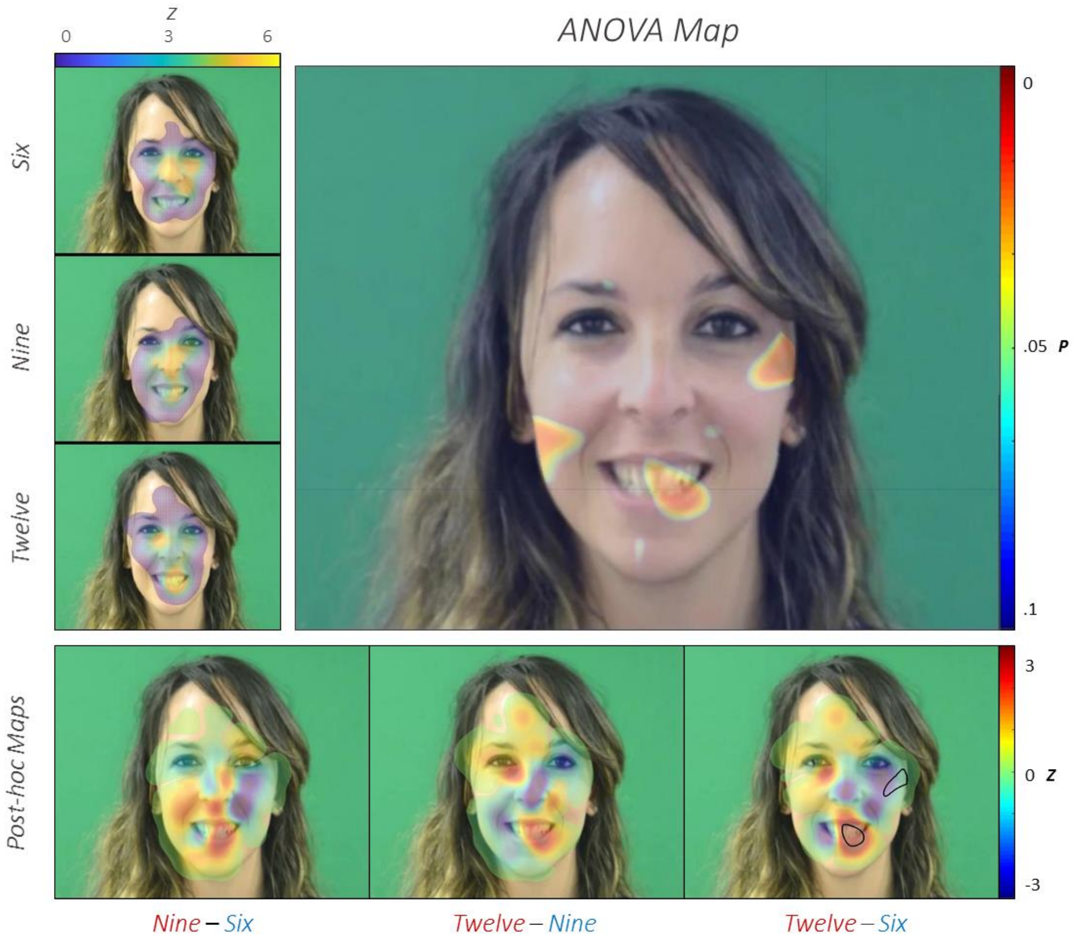

Figure 6. ANOVA heatmap analyses for a happy expression trial.

See Figure legend 3 for further information.

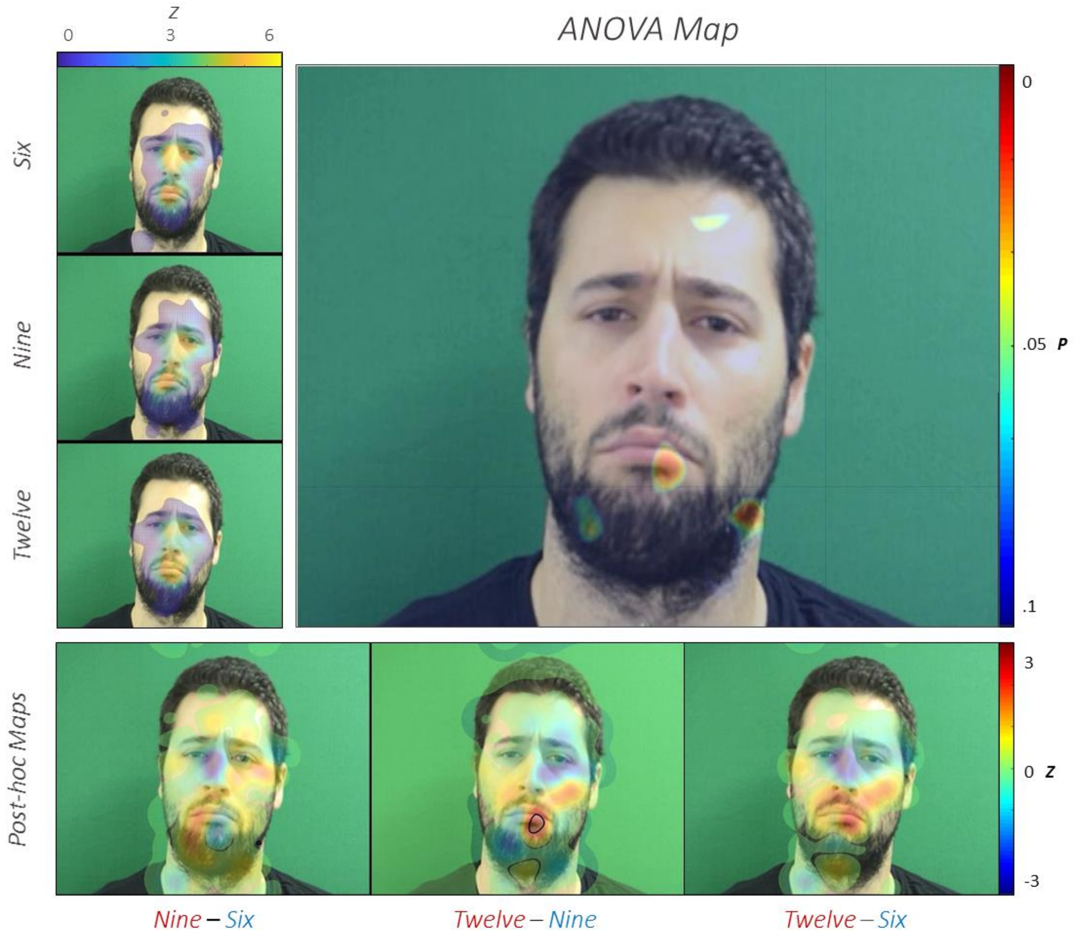

Figure 7. ANOVA heatmap analyses for a sad expression trial.

See Figure legend 3 for further information.

Combining temporal and spatial analyses: Surprise

Comparing six- and twelve-month-olds (Age C1), developmental differences in looking toward the surprised expression were found for face (β = -.413, t = -2.554, p = .011; Table 2; Figure 8bi) and lower-face features (β = -.341, t = -2.061, p = .039; Table 3; Figure 8biii), but not the lower face overall (β = -.023, t = -.123, p = .590; Table 4; Figure 8bii). Figure 8a shows mean looking, ANOVA and post-hoc comparisons for Age differences in scanning for a surprised expression trial. Figure 8b shows Age x Surprise plots for looking toward the face (i), lower face (ii) and lower feature (iii) AOIs. From Figure 8a, we can see that there were few spatial differences in looking between age groups, with infants of all ages demonstrating precise looking toward the diagnostic mouth region. Greater looking toward the mouth in six-month-olds is therefore unlikely to be driven by a greater proportion of looking to that area in comparison to other facial regions, but instead by maintaining their interest there for longer (Lower feature: Age C1 x Surprise x Linear: $\beta = -1.815$, $t = -2.578$, $p = .010$; Age C1 x Surprise x Quartic: $\beta = .896$, $t = 2.198$, $p = .028$). We can see this pattern illustrated in Figure 8b. Six-month-old infants' greater interest in the surprised face in general (Face: Age C1 x Surprise: $\beta = -.413$, $t = -2.554$, $p = .011$) is likely what is underlying these developmental differences in looking to the mouth (Figure 8biii).

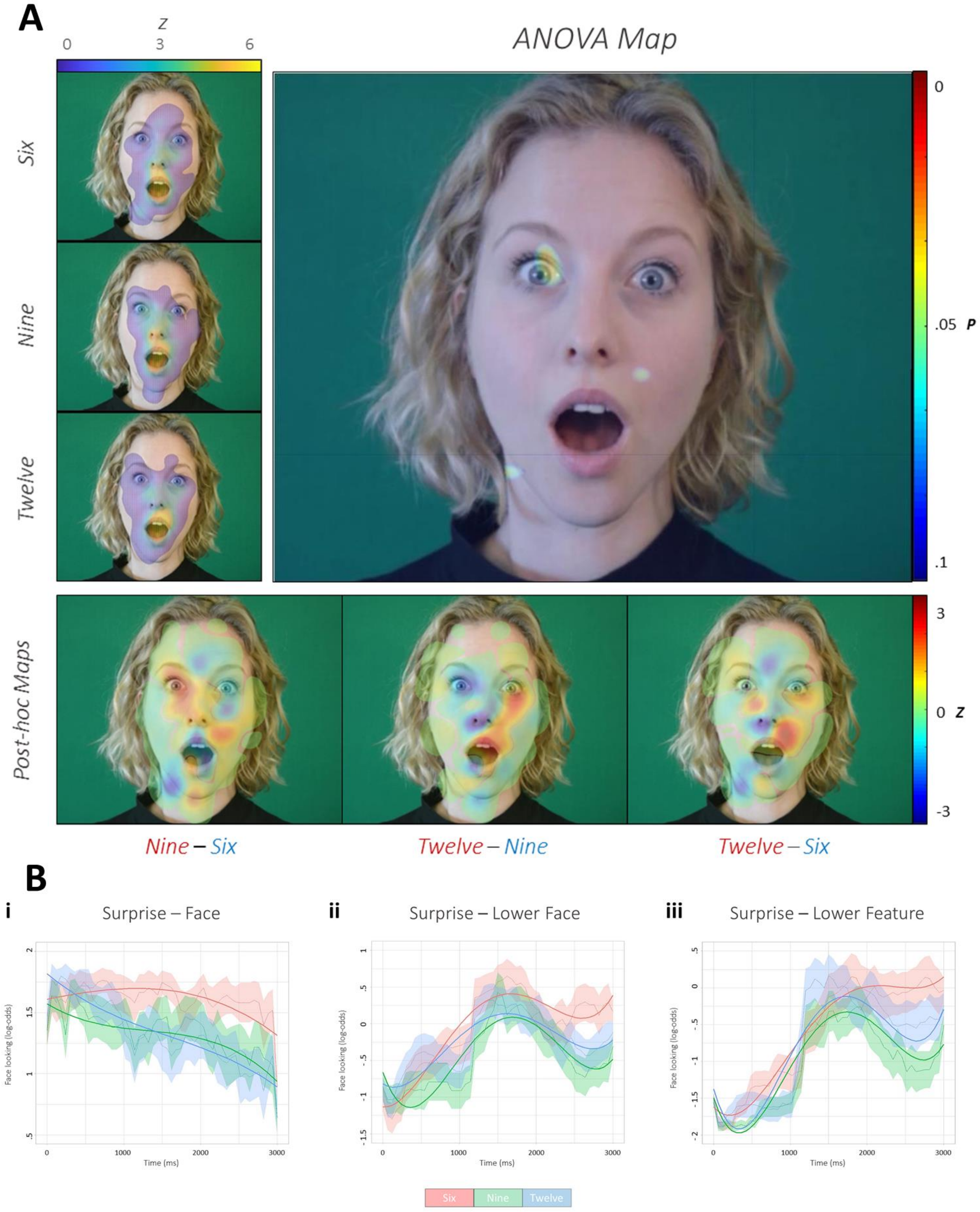

Figure 8. ANOVA heatmap and temporal analysis for a surprised expression trial.

Statistical heatmap analyses (A) for a surprised expression trial, see Figure legend 3 for further information. Developmental differences in face (Bi), lower face (Bii) and lower-face feature (Biii) looking for the surprised expression across time. See Figure legend 2 for more information. Three polynomial terms were used to model face-looking across time (linear, quadratic and cubic), and four for lower-face and lower-feature looking (linear, quadratic, cubic and quartic). The data are divided according to the factor Age: six (red), nine (green) and twelve months (blue).

Discussion

This chapter has explored the development of face interest and scanning for all six basic dynamic expressions (happy, sad, surprise, fear, anger, disgust) across six, nine and twelve months of age. In general, the findings from this chapter are consistent with previous work (see Quinn et al., 2011 for a review), but also make several novel contributions.

Expression interest

Previous work has shown infants prefer happy over neutral and angry expressions (La Barbera et al., 1976; Soussignan et al., 2017), and happy over sad expressions (Montague & Walker-Andrews, 2001; Soken & Pick, 1999). A preference for happy over fearful expressions has also been found in newborns (Farroni et al., 2007), but by five to seven months, infants transition toward showing greater interest in negative expressions, fear in particular (Heck et al., 2016; Peltola, Leppänen, Vogel-Farley, Hietanen, & Nelson, 2009; Vaish et al., 2008). This study found all expressions are looked at longer than neutral and some (surprise, fear) are looked at for significantly longer than anger. Yet there was no particular attentional bias for fearful faces, or any developmental differences in interest toward the fearful face. Interestingly, another study using sequentially-presented dynamic stimuli also failed to find heightened interest for fear compared to happiness, but similarly find preferences for fear over anger and neutral (Soussignan et al., 2017).

Infants looked longer toward surprised faces than any other expression. Previous findings with static stimuli suggest that surprise can be discriminated from other expressions (Ludemann & Nelson, 1988; Serrano, Iglesias, & Loeches, 1992; Young-Browne, Rosenfeld, & Horowitz, 1977), but have not reported that infants attend to this expression more than others. It is conceivable that surprise has increased salience when dynamic. The wide mouth and eyes of surprise may present a greater amplitude of motion compared to dynamic presentations of other expressions such as fear. An expression's perceived intensity is heavily influenced by dynamic information (Biele & Grabowska, 2006; Weyers, Mühlberger, Hefele, & Pauli, 2006) and was not controlled for here. Further work might explore the role of facial motion and affective intensity in dynamic expressions and how this might influence infants' interest in expressions such as surprise.

Consistent with previous literature (La Barbera et al., 1976; Schwartz, Izard, & Ansul, 1985; Soussignan et al., 2017), this study also found anger held infants attention least of all the emotional expressions. Though it is theorised that infants show a general 'negativity bias' (Vaish et al., 2008), this effect may be specific to fear (Peltola, Forssman, Puura, van Ijzendoorn, & Leppänen, 2015), and infants might actually find angry expressions aversive. For example, Schwartz and colleagues (1985) have shown that five-month-old infants demonstrate a marked tendency to disengage from angry facial expression stimuli.

There was also a developmental transition between six and nine months for interest in surprised and angry expressions. Six-month-olds showed increased interest in the surprised expression and a stronger aversion to the angry expression compared to older infants. There was also an indication of heightened interest in disgust at twelve months compared to nine months. It has been noted previously that spontaneous interest in individual emotional expressions may show different developmental trajectories (Quinn et al., 2011). For instance, younger infants' attention might be highly predicted by salient motion cues (Vinter, 1986), and thus demonstrate an increased interest in surprise which presents a greater amplitude of motion compared to other expressions. Older infants, however, may show a *relative* increase in interest toward negative expressions compared to six-month-olds (see Leppänen & Nelson, 2009; Vaish, Grossmann, & Woodward, 2008); first anger by nine months, then disgust by twelve months (Ruba et al., 2017).

Expression scanning

These analyses conducted in this chapter show that infants, like adults (Jack et al., 2014; Scheller, Büchel, & Gamer, 2012; M. L. Smith et al., 2005), attend to diagnostic regions of emotional expressions. For instance, there was a pronounced peak in lower-feature looking as the mouth widened in surprise, but remained predominantly in the eye-region when the eyes widened in fear. Larger peaks in lower-feature looking were also found for other expressions with highly diagnostic and dynamic lower-features (e.g. happiness and disgust), and smaller peaks for expressions with less salient lower regions (e.g. sadness and anger).

For **surprise**, although all age groups showed spatially precise diagnostic looking toward the mouth, this region held the attention of six-month-olds longer than older infants. Consequently, this may not necessarily reflect developmental differences in scanning, but instead reflect six-month-olds' increased interest for the surprised expression in general. For **anger**, twelve-month-olds demonstrated greater overall lower-face looking, and more precise scanning of the eyes and mouth compared to six-month-olds. The previous chapter compared scanning differences between dynamic and static conditions and found increased lower-face looking for dynamic anger. As the lower facial features communicate diagnostic information within dynamic (Jack et al., 2014), but not static (M. L. Smith et al., 2005) angry expressions, it is conceivable that it is this additional information that is directing infant attention more toward the lower face for dynamic stimuli, and what is also being attended to a greater extent in older infants.

For **disgust**, there was a developmental transition in diagnostic looking toward the lower-face features between nine and twelve months of age. Combining the heatmap and temporal analyses, it can be interpreted that there was a stronger quadratic peak in looking toward the diagnostic mouth-nose region across time for twelve-month-olds compared to nine-month-olds. There was also greater overall lower face looking for disgust in twelve-month-olds compared to six-month-olds. Few studies have investigated the development of disgust perception, but recent work has found that dynamic disgust expressions can be discriminated at birth (Addabbo et al., 2018), and categorised by ten months (Ruba et al., 2017), and that twelve-month-olds look more at the nose region of disgust compared to seven-month-olds (Soussignan et al., 2017). This study's results are consistent with these findings and suggest that the diagnostic scanning of dynamic disgust develops from nine to twelve months.

There were no developmental differences in **fearful** face scanning within the AOI-based analyses, but the heatmaps revealed more precise diagnostic looking to the whites of the eyes and open mouth in twelve-month-olds. Similar to anger, the lower features are also diagnostic for dynamic fear (Jack et al., 2014), but not static fear (M. L. Smith et al., 2005), and as such comparatively greater looking toward lower features were found within the dynamic condition (see Chapter 5). However, in contrast to twelve-month-olds, six-month looking clustered much more around the nose and the central regions of the face. These developmental differences were likely missed in the AOI analyses as they cannot be described according to asymmetries in looking along the vertical axis (i.e. upper or lower face bias). Nevertheless, these heatmap findings complement previous literature that note a developmental change in fearful face processing across the first year of life (see Leppänen & Nelson, 2009 for a review).

The AOI-based analyses also did not find any developmental differences in face scanning for **happiness** and **sadness,** but the heatmaps were able to identify highly specific clusters of significant differences *within* AOI regions. For happiness, a section of the lower mouth was fixated more in twelve-month-olds compared to six-month-olds. Similarly for sadness a region of the lower lip was fixated more in twelve-month-olds compared to nine-month-olds. However, in general the findings from this chapter suggest that face scanning strategies for these more familiar facial expressions may be stable by six months of age (c.f. Soussignan et al., 2017).

Methodological tools

Chapters 5 and 6 have used naturalistic video expression stimuli within the GC Social interaction paradigm. The data generated from this research paradigm is rich and complex, yet several methodological tools have been utilised in order to analyse these data whilst preserving important spatial and temporal detail. Dynamic AOIs were created in response to the fundamental challenge of defining where upon a stimulus infants are looking. This method is data-driven, and generates area of interest information automatically for each frame of the video after first setting the initial parameters. This method represents a substantial methodological step forward compared to the current literature (c.f. Hunnius & Geuze, 2004; T. Wilcox, Stubbs, Wheeler, & Alexander, 2013), however there are also some important limitations. This method uses pre-defined colour and luminance for the actor’s skin to differentiate the face from the hair, torso and background, and inversely defines the features of the face. However, various stimulus aspects (e.g. an actor’s facial hair) generate confounds, and some facial features (e.g. the nose) are harder to detect consistently. For many research questions, instead of using particular feature AOIs, larger areas of interest (e.g. entire upper or lower facial regions) might be appropriate (see Chapter 8 for further discussion).

Previous work has identified that arbitrarily carving up the stimulus or presentation time window is unadvisable given that this can distort results or lead us to overlook important patterns (D. J. Barr, 2008; Caldara & Miellet, 2011). For instance, mixed effects modelling (Dink & Ferguson, 2015) was used within this chapter to identify subtle developmental differences in facial expression scanning that were missed when the data were collapsed across time. Nevertheless, large models using multiple polynomial terms can generate a complex pattern of results, can take a long time to fit and can often lead to non-convergence (Baayen et al., 2008; D. J. Barr et al., 2013) and should therefore be approached with caution. Alternatively, Cluster permutation analysis (Maris & Oostenveld, 2007) was used in Chapter 5 to identify time windows of significant differences between dynamic and static conditions. This analysis proved a useful complement to mixed effects modelling when performing simple comparisons (e.g. the temporal differences between two conditions).

Both of these methods, however, require the data to be collapsed across space into AOIs. Chapters 5 and 6 also analysed the data using heatmap analyses. The ANOVA heatmaps used here allowed significant clusters of differences in scanning between age groups to be identified, including some (e.g. fear, see Figure 4) that were not uncovered when the data were collapsed into arbitrary “upper” and “lower” face regions. However, like in the temporal analyses, using heatmaps might identify small and unimportant regions of statistically significant differences which are unlikely to be meaningful, and again require collapsing across one dimension (i.e. time). Yet in combination these methods enable us to “see both sides of the coin”, and allow us to handle the complex data generated when presenting naturalistic stimuli within eye-tracking paradigms.

Conclusion

This chapter has investigated how infants attend to dynamic presentations of facial emotion. By combining analysis methods that have high spatial (statistical heatmaps) and temporal (mixed-effects modelling of looking to dynamic AOIs) resolution, these analyses were able to mitigate the methodological challenges that can arise when collecting eye movement data using highly complex dynamic stimuli. Overall the results indicate that infants by six months already show sophisticated expression processing; preferentially attending biologically-relevant expressions and diagnostic facial regions. However, older infants’ relative increases in interest and scanning precision for negative emotions (e.g. fear, anger and disgust) suggests the perception of facial emotion continues to be adaptively tuned across the first year of life. Notably however, infants’ *scanning* of and *interest* toward facial expressions tells us little about their *understanding* of the meaning conveyed through these facial movements (see Ruba, Meltzoff, & Repacholi, 2019). Future work might consider convergent approaches as recording additional responses from infants (e.g. heart rate, pupil size, facial or manual gestures) would aid interpretations of eye-movement data considerably.

Chapter 7

Using pupillometry as a physiological marker for emotional processing in infants

---

Abstract

Pupil size is a correlate of a subject’s state of interest, attention or arousal, and previous work has demonstrated its utility in measuring infants’ physiological responses to affective stimuli. Pupil size is also automatically recorded alongside gaze-location by most modern eye-tracking cameras, however pupillometry presents substantial methodological challenges in infant research given that pupil size primarily responds to light. This chapter presents two alternative solutions to the confounding effect of luminance. Firstly a paradigm is developed that minimises luminance-based changes in facial expression stimuli, whilst also implementing techniques (e.g. gaze-contingency, video breaks) to ensure infants’ attention to the screen. Secondly the pupil size data recorded from Chapter 5 for naturally-engaging, dynamic facial expressions is analysed, while controlling statistically for the effect of luminance. The results suggest that infants’ only show a clear dilation for dynamic happy expressions, though also show some sensitivity for fear. These results suggest that pupillometry can be a useful convergent methodology to help clarify the meaningfulness of eye-tracking data.

Introduction

While the location of a fixation at any given time point is an important indicator of a subject’s interest in a presented stimulus (Yarbus, 1967), the size of a subject’s pupil across time can also reflect evoked changes in their cognitive or affective processing of that stimulus (Hess & Polt, 1960, 1964; Kahneman & Beatty, 1966; see Hess, 1972; Sirois

& Brisson, 2014 for reviews). This is true for adult participants, but also for infants (Fitzgerald, 1965; Gredebäck & Melinder, 2010; Jackson & Sirois, 2009), and given that modern eye-trackers will automatically record pupil size alongside gaze location, pupillometry is quickly becoming an important method for investigating preverbal populations (see Hepach & Westermann, 2016 for a review).

Pupil size is determined by the tone of the reciprocal dilator and constrictor muscles of the iris, which are centrally controlled by the autonomic nervous system (ANS) via sympathetic and parasympathetic paths respectively. These muscles respond primarily to light, whereby an increase in luminance leads to a sharp pupillary constriction (‘pupillary light reflex’; PLR), which is subsequently followed by a more gradual re-dilation. Under constant illumination, oscillating pupil size represents the homeostasis between activity within sympathetic and parasympathetic systems (Beatty & Lucero-Wagoner, 2000; Loewenfeld & Lowenstein, 1993; Steinhauer, Siegle, Condray, & Pless, 2004). Psychologically-evoked pupillary responses have a strong functional association with the activity of the noradrenergic (NA) system’s locus coeruleus (LC), which can trigger pupillary dilation by promoting sympathetic function via a diffuse network of projections to the spinal cord and autonomic nuclei, including inhibitory connections with parasympathetic nuclei (e.g. Erdinger-Westphal nucleus) which directly innervate the constrictor muscles of the iris (see Aston-Jones & Cohen, 2005; Laeng, Sirois, & Gredebäck, 2012). Thus reflexive pupillary dilations under constant luminance can serve as a key biomarker for the LC-NA system, as well as the cognitive or affective processing with which it is associated.

Facial expressions

As pupillary responses are mediated by the activity of the LC-NA system (Joshi, Li, Kalwani, & Gold, 2016; Murphy, O'Connell, O'Sullivan, Robertson, & Balsters, 2014), dilations are correlates of a subject's state of interest, attention or arousal (Hess & Polt, 1960; Sara, 2009; Sara & Bouret, 2012). Previous work with adults has shown sexually or emotionally arousing stimuli (both visual and auditory) illicit larger dilations relative to neutral stimuli (Attard-Johnson, Bindemann, & Ó Ciardha, 2016; Bradley, Miccoli, Escrig, & Lang, 2008; Henderson, Bradley, & Lang, 2014; Kret, Roelofs, Stekelenburg, & de Gelder, 2013; Partala & Surakka, 2003). Given this, pupillary responses are also a sensitive index of facial expression perception as perceiving emotions in others can trigger an autonomic response in the observer. For instance, adults show larger pupil dilations when perceiving others' negative compared to positive emotions (Yrttiaho, Niehaus, Thomas, & Leppänen, 2017), even when the expressions are presented subliminally (Laeng et al., 2013).

Pupillary correlates of social and affective processing have also been studied in infancy. Paralleling Fantz's classical looking-time research (Fantz, 1963), Fitzgerald (1968) demonstrated that one- and four-month-old infants show greater pupillary dilation for social compared to non-social stimuli. More recent work has found that infants also show dilations in response to the emotions of others. For example, visual and auditory presentations of other infants' emotional displays produced larger dilations compared to neutral stimuli in six- and twelve-month-olds (Geangu, Hauf, Bhardwaj, & Bentz, 2011), and seven-month-old infants showed larger dilations for happy compared to fearful expressions regardless of whether the expression was consciously perceived (Jessen et al., 2016). Pupillometry has also been used to demonstrate that 14-month-olds' reactivity to emotional facial expressions can be modified by contextual factors such as the familiarity or gender of the actor (Gredebäck, Eriksson, Schmitow, Laeng, & Stenberg, 2012) or the congruency of the actor's emotions with their behaviour (Hepach & Westermann, 2013). Differential pupillary responses to emotional expressions have also been found in infants who are at high risk for developing autistic spectrum disorder (ASD) compared to low-risk controls (Wagner et al., 2016), with increased pupil sizes for emotional stimuli at 9 months being predictive of social-communicative functioning at 18 months. Current research therefore suggests that pupillary dilations are a sensitive index of an infant's affective response to the emotional expressions of others.

Nevertheless, previous pupillometry studies have found mixed findings for infant pupil responses to emotional expressions (Geangu et al., 2011; Hepach & Westermann, 2013; Jessen et al., 2016). For instance, Geangu and colleagues (2011) presented 50-second audio-visual clips to infants depicting other infants who were displaying happy, distressed and neutral affect. They reported greater pupillary dilation in response to the videos of distressed infants. Conversely, Jessen and colleagues (2016) presented greyscale images of happy, fearful and neutral faces for one second only, and found increased pupil responses for positive, happy expressions compared to fear. These findings do not neatly align according to affective valence, but given the substantial differences in methodology, these studies are not easily compared. Previous work has also differed greatly in how emotional expressions are portrayed. For example, Hepach and Westermann (2013) presented video clips of seated actors expressing emotions through both their behaviour and their facial expression (e.g. 'thumping' a stuffed animal toy to depict anger), which again is very different to the one-second, static face images in Jessen and colleagues' study. The current study will be the first to explore infant pupillary responses to all six basic emotional expressions, and the first to do so for dynamic and interactive expression stimuli.

Methodological considerations

Unlike traditional eye-tracking, pupil size can be used as a direct correlate of affective processing in infancy, and could therefore provide a valuable complement to methods measuring infant behaviour. However, as pupil size responds primarily to changes in retinal illumination, differences in ambient lighting and stimulus luminance can introduce noise within the data, particularly as pupillary responses to light are typically much larger (over 100%) than psychologically-induced changes in pupil size (approx. 20%; Beatty & Lucero-Wagoner, 2000). The ideal set-up for accurate pupil-size recording would therefore minimise the confounding effect of light. The ambient lighting within the room, and the display monitor's brightness would be kept constant for all participants. The stimuli presented on the screen would be greyscale and matched according to luminance; not varying across time or between presentations. Luminance-matched control stimuli (e.g. random pixels) would also be used within control conditions and presented before the analysis time window to provide accurate non-PLR baseline pupil size measurements. Participants would ideally have a direct, fixed gaze upon the stimulus to minimise the effects of varying luminance across the stimulus region, and any inaccuracies due to the foreshortening effect (Hayes & Petrov, 2016). Pupil data would also be standardised (z-scored and/or baseline-corrected) and time-locked to stimulus onset to remove individual differences in tonic pupil size not attributable to the experimental manipulation (Attard-Johnson, Ó Ciardha, & Bindemann, 2019). Nevertheless, while it is straightforward to implement these practical considerations in adult eye-tracking, they introduce considerable practical challenges when investigating infants (see Hepach & Westermann, 2016 for a review).

A central challenge within developmental eye-tracking is ensuring infant participants remain engaged and attentive to the presentation on the screen. However, while lack of interest might be informative for looking-time research (i.e. as an indication of disinterest or habituation), it is entirely harmful for pupillometry as each look-away and refocus on the screen will trigger light-based pupil size changes and introduce artefacts within the data. As infants look away they also might miss crucial on-screen changes hypothesised to induce a pupillary response. Stimuli designed to minimise the confounding effects of light, such as static, silent, grey images, may not be sufficiently engaging to hold an infant's attention and thus detrimental for collecting good quality data. To mitigate this apparent trade-off, we can either develop methods with highly controlled stimuli that can nevertheless still hold an infant's attention, or alternatively we can present infants with more ecologically-valid and naturally-engaging stimuli whilst compensating statistically for any systematic effects of varying luminance.

The current chapter will implement both of these approaches to investigate infants' event-related pupillary responses to emotional expressions. The first experiment presents infants with highly-controlled stimuli of neutral, angry and fearful eye-regions. Through a series of preliminary experiments, an eye-tracking paradigm is developed which maximises infant attentiveness to these stimuli. The second experiment analyses the pupil data recorded from the facial expression experiment in Chapter 5. The stimuli from the experiment in Chapter 5 were created to maximise ecological validity in lab-based eye-tracking research (via the GC Social Interaction paradigm) and are therefore colourful, dynamic and interactive, and should be sufficient to hold infants' attention. In this experiment, the pupil data from both static and dynamic expressive faces will be analysed whilst also regressing out the confounding effect of stimulus luminance (see Aslin, 2012; Jackson & Sirois, 2009; Yrttiaho et al., 2017).

Method

Experiment 1: Expressive Eyes

To provide a methodological control from which to compare the pupillometry data from the experiment conducted in Chapter 5, a series of pilot studies were conducted. Thirty-two infants (see Chapter 2 for eligibility criteria) were included in this piloting phase, across four preliminary experiments.

Stimuli. The 'Expressive Eyes' stimuli were designed firstly to minimise luminance-dependent changes in pupil size, and secondly to discourage diffuse scanning across the screen. Sixteen grayscale eye-region images (300 x 144 pixels) were used (8 x Neutral, 4 x Fearful and 4 x Angry), with comparable mean luminance between neutral and expressive stimuli (Mean difference = 1.14%, $t(7) = -1.483$, $p = .182$, see Table 1). As expressive stimuli have slightly higher luminance values, any dilations in response to expressive stimuli should not be due to luminance. Images were presented centrally on a neutral grey background (Luminance = 0.8; to give larger pupils, darker backgrounds should be used), and displayed on a Dell 20-inch display monitor (1024 x 768 pixels). The screen luminance was fixed for all participants. As infants during natural scanning fixate different regions of the face, this paradigm attempted to minimise the luminance-based changes in pupil size that might arise from diffuse face scanning by presenting only images of the eye-region. The eye-region is considered diagnostic for both fearful and angry expressions (Hanawalt, 1944; M. L. Smith et al., 2005), and is sufficient for the communication of these emotions (Calder et al., 2000; E. Fox & Damjanovic, 2006). For instance, seven-month-olds can distinguish fearful from non-fearful eyes even when presented below the perceptual threshold (i.e. subliminally; see Jessen & Grossmann, 2014, 2016).

Procedure. The initial experimental procedure consisted of eight trials, presented in a random order. Each trial consisted of three sub-trials: two consecutive presentations of neutral eyes, followed by a single presentation of expressive eyes (24 presentations in total). Each image was displayed for 5 seconds. The relative pupil size differences following the neutral-expressive transition were to be compared to the neutral-neutral transition. Four of the trials included fearful eyes, and four included angry eyes (see Figure 1).

Eye-tracking. Pupil sizes were recorded using an SR Research Desktop-Mount EyeLink 1000+ eye tracker with a 25mm lens operating in remote mode (spatial resolution 0.01°, average gaze position error of 0.25°). The ambient light in the room was kept constant throughout testing and the same low-light conditions were re-created for each participant. A padded target sticker placed centrally on the forehead served as a reference point for recording eye movements and head distance. Prior to the start of each experiment, a five-point calibration procedure was implemented (Experiment Builder, SR Research, Ontario, CA), using custom 'attention grabbers' (animated, noisy circles) to entice looking. These attention grabbers were also used between trials to perform a drift correction. Pupil size was recorded using the EyeLink's scale from 100 to 10000 units, with a precision of 1 unit, with noise levels of 0.2%, corresponding to a resolution of 0.01mm for a 5mm pupil. Pupil data were z-normalised for each participant.

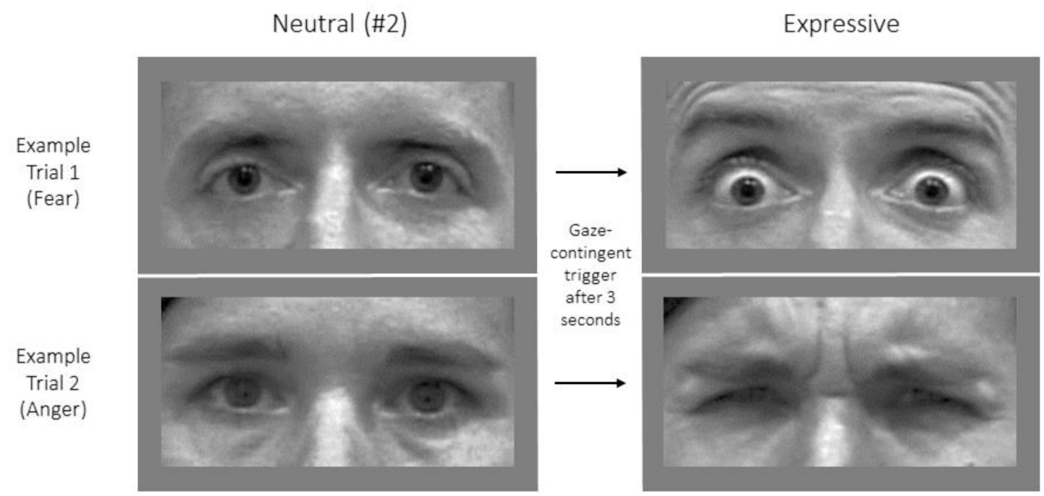

Figure 1. Stimulus examples for the Expressive Eyes experiment.

This experiment included eight trials consisting of three eye-region image presentations each: two neutral and one expressive. This figure depicts the second neutral sub-trial followed by the expressive sub-trial for two example trials (1 x Fear, 1 x Anger). Images were grayscale and matched for luminosity (300 x 144 pixels) and displayed on a uniform grey background. The gaze-contingent transition between images occurred after three seconds, and only if the infant was fixating the stimulus.

Task development. Several problems with this procedure were noticed, and amendments were made in further preliminary experiments. Firstly, images were combined to form a seamless video presentation within trials as the blank screens between image presentations were causing large pupillary constrictions. Secondly, given that the images were small, grey and silent, infants quickly lost interest and looked away from the screen causing considerable data loss. To remedy this, an engaging video was added after the fourth trial (a colourful, musical animation) to refresh infants' attention. Thirdly, the stimulus display time was also reduced to three seconds per image, to reduce the overall time infants were required to attend the screen. Fourthly, in a final adaptation to the experimental procedure, a gaze-contingent transition was added between the images within a trial. After either the first or second eye-region image had been displayed for three seconds (max 5s), the subsequent eye-region image was then displayed *only* if the infant was fixating the stimulus (via a predefined, invisible gaze-contingent boundary placed around the image). This ensured that infants perceived both the neutral-neutral and neutral-expressive transitions, and that the start point for each analysis time window was time-locked to the beginning of the display period (see Hepach & Westermann, 2016 for further practical suggestions).

Table 1. Average stimulus luminance for expressive eyes and face stimuli

	Expression	**Mean**	**Range**	**Difference from Neutral (%)**
E1: Eyes	Neutral	.5029	.0363	--
	Fear	.5124	.0516	0.95
	Anger	.5163	.0430	1.34
E2: Faces	Neutral	.5532	.0686	--
	Happy	.5151	.1037	- 3.81
	Sad	.5584	.0224	0.52
	Surprise	.5121	.0835	- 4.11
	Fear	.5353	.0527	- 1.79
	Anger	.5492	.0189	- 0.40
	Disgust	.5509	.0202	- 0.23

Luminance values were computed on a 0 (full black) to 1 (full white) scale

Experiment 2: Expressive Faces

The second experiment in this chapter will use pupil data collected during the experiment in Chapter 5, which presented dynamic, interactive expressive faces to infants. Further details about participant information (Table 2), stimuli and methodological procedures can be found in Chapter 5. In summary, infants were allocated to either a dynamic or static condition. Within the dynamic condition they were presented with an on-screen actor with neutral affect. Once infants fixated within a pre-defined gaze-contingent boundary placed around the eye-region, the three-second expression animation (either dynamic neutral, happy, sad, surprise, fear, anger or disgust) was triggered. Within the static condition, still frames from the expression videos used in the dynamic condition (peak expressive amplitude, see Figure 1 in Chapter 5) were presented for the entire trial duration (five seconds). In both conditions, pupil data were cut to a defined analysis time-window of three-seconds, following the first eye-region fixation.

Stimulus luminance values from the static condition are given in Table 1. As the images used in the static condition were still frames taken directly from the dynamic videos, these values are representative of stimulus luminance for both conditions. As these images were not grayscale, the values from each of the RGB channels were adjusted to obtain photometric accuracy (i.e. brightness as perceived by a human observer) using this formula:

$$y = R \times a_1 + G \times a_2 + B \times a_3$$

Where the human-corrected luminance y is obtained by multiplying each of the three colour channels R, G and B by a specific adjustment factor a, equal to 0.2126 for red, 0.7152 for green and 0.0722 for blue (Jackson & Sirois, 2009). The average y was then computed for each stimulus, and listed in Table 1. Stimulus luminance was comparable to neutral for sad, angry and disgusted expressions, but noticeably lower for happy and surprise (Mean difference = -3.96%). Overall, luminance for expressive stimuli was only slightly lower than neutral (Neutral *M* = .55, Expressive *M* = .54).

Participants

Thirty-nine infants aged six-, nine- and twelve-months (see Chapter 2 for more information) were included in the final analyses of the ‘Expressive Eyes’ experiment, with a second ‘Expressive Faces’ experiment including pupil data from the 149 infants from both the static and dynamic conditions reported in Chapter 5 (see Table 2). An additional 15 infants from Experiment 1 were excluded for having missing data (zero trials containing sub-trials with > 1s of pupil data), for both experiments, trials with missing data (< 1s of pupil data) were automatically discarded in each analysis.

Table 2. Participant information for both Experiments 1 and 2

	Age (months)	**Age M (SD) (days)**	**N**	
			Male	**Female**
E1: Eyes	6	190.46 (9.95)	7	4
	9	272.53 (13.12)	9	10
	12	360.56 (11.22)	5	4
E2: Static Faces	6	194.67 (26.22)	4	5
	9	276.32 (14.16)	10	9
	12	370.78 (11.47)	4	5
E2: Dynamic Faces	6	192.87 (10.12)	16	17
	9	274.91 (12.49)	16	16
	12	366.43 (13.97)	18	18

Data processing

Both eye-tracking (see Chapter 2 for more information) and pupil size data were processed in MATLAB (Mathworks, R2017a). Raw pupil size values from both experiments (Eyes and Faces) were first converted from EyeLink arbitrary units to standard z-scores. These values were then cleaned by removing extreme values (pupil sizes with a z-score > 3 or < -3) and by removing pupil size values that were recorded during saccades or whilst the infant was not fixating the stimulus. As the gaze-contingent manipulation led to variable trial/sub-trial lengths, sub-trials with greater than one second total fixation duration on the stimulus were then cut to a standard three seconds in length following stimulus onset. Pupil data were averaged across trials for each participant, ignoring missing values, to produce a standard three-second neutral dataset and standard three-second expressive dataset. These data were then baseline-corrected by subtracting the average of the first 50 samples (100ms) from each data point (Attard-Johnson et al., 2019; Reilly, Kelly, Kim, Jett, & Zuckerman, 2019).

Analysis plan

Initially, pupil sizes will be collapsed across the time window and average baseline-corrected pupil sizes will be compared, but following previous work (Geangu et al., 2011; Jackson & Sirois, 2009; Sirois & Brisson, 2014; Sirois & Jackson, 2012) changes in pupil size across time will also be investigated using an approach based on functional data analysis (FDA; Ramsay & Silverman, 1997; see Jackson & Sirois, 2009; Sirois & Brisson, 2014). With this approach, continuous data (e.g. pupil size) can be expressed as a function of another continuous variable (e.g. time). To do this, FDA uses b-splines which are constructed piecewise from cubic polynomial bases to form an overall functional spline of order 4 (see Jackson & Sirois, 2009). In the current analysis these splines will be constructed of 30 bases and fit (through least-squares regression) to each participant's pupil data (per trial). Once a participant dataset of functional splines are created, they can then be used to perform common statistical analyses such as *t*-tests or ANOVA's, with the test values (e.g. means, standard deviations or *t*-scores) also represented as functional curves across time.

Results 1: Expressive Eyes

Mean pupil size

A paired t-test was used to compare relative differences in the change in mean pupil size following a stimulus transition to expressive eyes in comparison to neutral eyes. This test found significantly greater dilations in pupil size for neutral (*M* = -.004, *SD* =

.078) compared to expressive (*M* = -.050, *SD* = .079) stimuli (*t*(37) = -2.664, *p* = .011, *d* = .876). Splitting the data by trial expression, fearful eyes (*t*(35) = -2.269, *p* = .030, *d* = .767), but not angry eyes (*t*(30) = -1.697, *p* = .100, *d* = .620), showed significantly reduced pupil size compared to preceding neutral eyes.

Functional data analysis

In order to assess how pupil size changed from the gaze-contingent onset of the eye-region stimuli till the end of the three-second analysis time-window for neutral and affective stimuli, the difference between the third (expressive) and second (neutral) sub-trial stimuli were compared after first converting the data to curvilinear temporal functions (or ‘b-splines’; Figure 2). Looking firstly at the mean pupil sizes across time plotted in Figure 2 (top row), a pupillary light response (PLR) can be seen in the second neutral stimulus presentation, but not the first (Figure 2, bottom left, 0 – 0.5s), which directly follows the drift correction sequence. After the small initial PLR constrictions, the pupil data indicate a re-dilation in response to neutral eye-region stimuli, while pupil sizes for the expressive stimuli continue to decline. These differences in pupil size evoked by neutral and expressive eyes (Figure 2, bottom central) were significant at several points across the time window (Figure 2, bottom right, 1 – 3s).

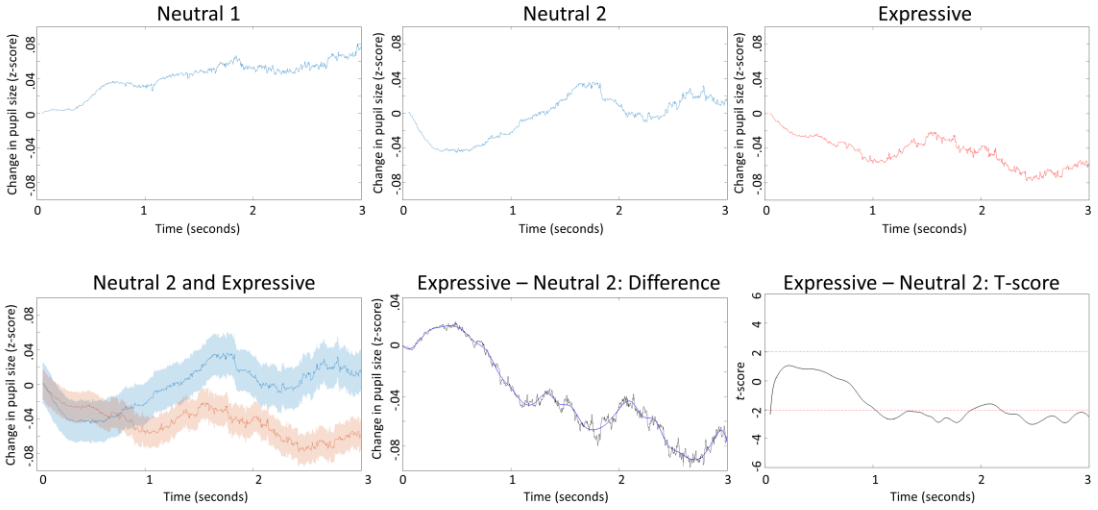

Figure 2. Infants' changes in pupil size in response to greyscale neutral and affective eye-regions.

The mean changes in pupil size are plotted following the onset of the first neutral (control) stimulus (upper left), then the gaze-contingent onset of the second neutral (upper central) and final expressive (either angry or fearful) stimulus (upper right). These data are baseline-corrected and cut to standard three-second segments. The mean change in pupil size for the second neutral (blue) and expressive (red) sub-trials are plotted together (lower left), with a shaded boundary depicting the standard deviation for each dataset. The difference between the change in pupil size for each stimulus is plotted (expressive minus neutral; lower central; black line), and functional data analysis was used to transform the data into b-spline curve functions (lower central, blue line). Curve functions were computed individually for each participant, and used to compute a t-score function, describing the difference between expressive and neutral stimuli across time (lower right). Thresholds for significant difference are plotted in red (critical t = 2.024, p = .05, two-tailed).

Results 2: Expressive Faces

Mean pupil size

A 2 (Condition: Dynamic or Static) x 3 (Age: 6, 9 and 12 months) x 2 (Expression: Neutral or Expressive) mixed ANOVA found no significant difference in mean pupil size between neutral (*M* = -.376) and expressive (*M* = -.361) stimuli ($F(1,143) = .697$, $p = .405$, $\eta_p^2 = .005$). There was a significant main effect of Condition ($F(1,143) = 45.664$, $p < .001$, $\eta_p^2 = .242$), but no significant main effect of Age ($F(2,143) = 1.392$, $p = .252$, $\eta_p^2 = .019$), or significant interactions (all $p > .13$).

While there were no significant differences between mean neutral and expressive pupil sizes overall, comparing differences between expressions via a 2 (Condition) x 3 (Age) x 6 (Expression: Happy, Sad, Surprise, Fear, Anger, Disgust) mixed ANOVA found a significant main effect of Expression ($F(5,545) = 3.075$, $p = .010$, $ŋ_p^2 = .027$). As before, differences between Condition ($F(1,109) = 41.255$, $p < .001$, $ŋ_p^2 = .275$), but not Age ($F(2,109) = .986$, $p = .376$, $ŋ_p^2 = .018$) were significant. There were also no significant interactions (all $p > .28$). Bonferroni-adjusted post-hoc comparisons indicate significantly larger mean pupil sizes for happy ($M = -.152$) compared to fearful ($M = -.299$, $p < .001$) and sad ($M = -.253$, $p = .010$) faces, but within the dynamic condition only (all static comparisons $p > .90$).

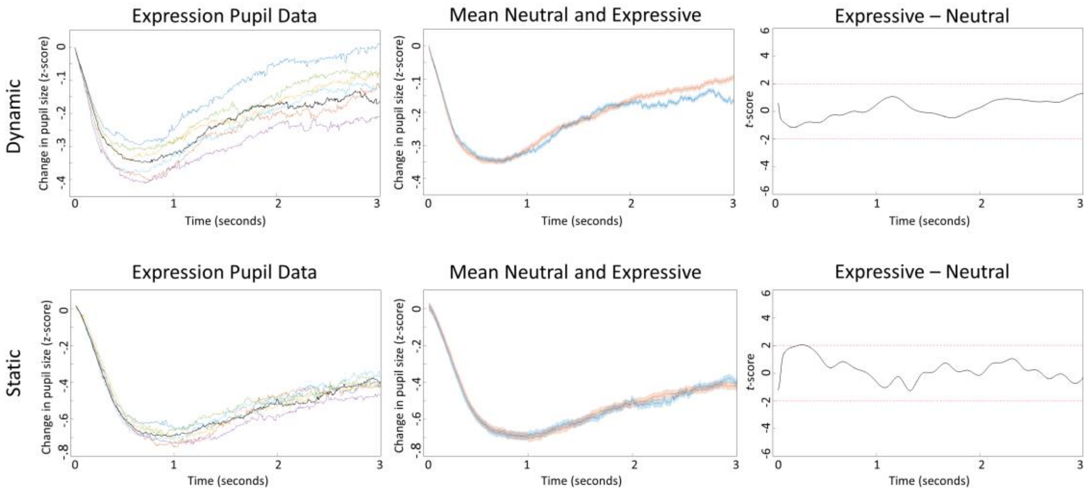

Figure 3. Infants' changes in pupil size in response to dynamic and static expressive faces.

The mean changes in pupil size for all expressions (neutral: black/bold, happy: dark blue, sad: red, surprise: yellow, fear: purple, anger: green, disgust: light blue) collapsed across participant and trial are plotted (left). For both static and dynamic conditions, the three-second analysis time window begins with the first eye-region fixation, but for dynamic this also triggers the three-second expression animation. These data are baseline-corrected and cut to the standard three-seconds. The mean change in pupil size for the neutral (blue) and expressive (red) sub-trials are plotted together (centre), with a shaded boundary depicting the standard error for each dataset. The difference between the change in pupil size for expressive and neutral trials (expressive minus neutral) was converted into b-spline curve functions for each participant and used to compute a t-score function (right). Thresholds for significant difference are plotted in red (dashed line: dynamic critical t = 1.984, static critical t = 2.012, p = .05, two-tailed).

Functional data analysis

As in Experiment 1, a functional approach was used to analyse differences in evoked pupillary responses across time. Figure 3 illustrates the pattern of pupil size change across the three-second time window. Firstly there was a clear PLR at the beginning of all trials, though this pupillary constriction was far more pronounced in the static condition (Dynamic M = -.224, Static M = -.545). After approximately 500ms a pattern of re-dilation can be seen across the remainder of the time window. Mirroring the mean pupil size analyses, the FDA analysis shows that although average neutral and expressive pupil sizes do not significantly differ across the time window (Figure 3, see central and right panels), there is much greater variation between expressions, particularly within the dynamic condition (Figure 3, left panels).

The FDA analyses for individual expressions are summarised in Figure 4 (dynamic condition only). The average change in pupil size (z-scores) for each of the six expressions are plotted (red) alongside neutral (blue). Functional *t*-scores are plotted beneath these descriptive graphs (black), with significance thresholds for Bonferroni-corrected and uncorrected alpha values (solid and dashed lines respectively). These plots show clear light-induced pupillary constrictions for all expressions, but that subsequent re-dilations show variation across expressions. For **happiness** there was a significantly larger pupil dilation compared to neutral, but the results for the other expression are less clear. Pupil dilations are larger for **surprise** and **anger** compared to neutral, but these mean differences did not cross the threshold for significance within this time window. Compared to neutral, **sadness** and **disgust** both show greater pupillary constrictions within the first 500ms. These early differences are likely to be driven by variations in PLR, as the stimuli for both these expressions have high average luminance values (see Table 1). **Fear**, however, not only evokes greater pupillary constriction during the first 500ms, but also presents a reduced re-dilation at later points along the time window (e.g. 1504-1850ms). As the fear stimuli contained relatively low average luminance values compared to neutral (fear = 0.535, neutral = 0.553, -1.79%), these results are in contrast to a luminance-based interpretation.

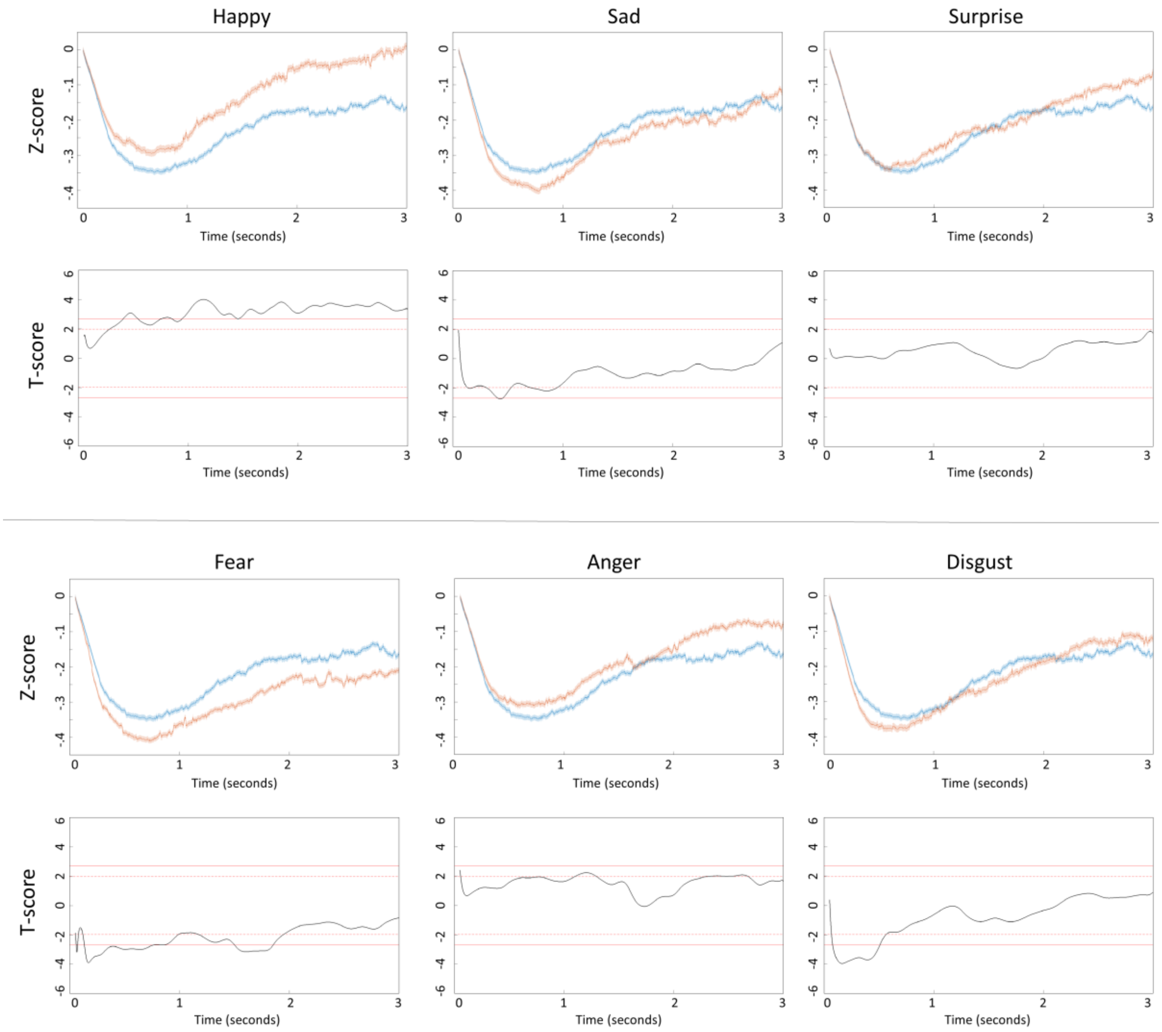

Figure 4. Infants' evoked pupillary responses for dynamic expressive faces.

The mean change in pupil size (z-scores) from the start of the analysis time window are plotted separately for each expression (red) alongside neutral (blue), with the shaded boundary depicting standard error. Beneath these plots, the differences between pupil sizes for expressive compared to neutral faces across time are illustrated using t-score functions (b-splines). Significance thresholds are plotted in red for both uncorrected (dashed line: critical t = 1.984, p = .05, two-tailed) and Bonferroni-corrected (solid line: critical t = 2.693, p = .008, two-tailed, six comparisons) alpha values.

Luminance-adjusted analysis

To statistically assess the contribution of Expression on pupil size during the dynamic condition whilst controlling for differences in stimulus Luminance, an analysis of covariance (ANCOVA) was performed for both *constriction* (0-1000ms) and *re-dilation* (1001-3000ms) phases of each trial using R. The ANCOVA can be described in Wilcoxon notation as:

$$Pupil\ size \sim Luminance + Expression$$

Here the mean pupil size is predicted by the fixed effect of Expression (Happy, Sad, Surprise, Fear, Anger and Disgust) contrasted against neutral, while controlling for the continuous variable of Luminance (min = 0.46, max = 0.57). Firstly, the assumption of homogeneity was checked by including the interaction term (Luminance x Expression) within each model. Luminance did not significantly interact with Expression within either phase (both $p > .32$).

During *constriction*, the ANCOVA yielded a significant effect of the covariate Luminance ($F(1,1515) = 7.993$, $p = .005$, $\eta_p^2 = .005$), but only a marginal effect of Expression after controlling for Luminance ($F(6,1515) = 1.972$, $p = .067$, $\eta_p^2 = .008$). As expected, higher luminance produced lower pupil sizes ($\beta = -.703$). Planned contrasts also revealed significantly greater pupillary constriction for fear compared to neutral ($t(1515) = -2.400$, $p = .016$, $d = .123$; all other $p > .21$). Figure 5 groups the eighteen trials by similar luminance (top left, blue) and by expression (bottom left, red); raw participant pupil size z-scores are plotted with their means and 95% confidence intervals.

During *re-dilation*, the effect of Luminance was small and non-significant ($F(1,1410) = 2.889$, $p = .089$, $\eta_p^2 = .002$). However, the effect of Expression was significant after controlling for Luminance ($F(6,1410) = 2.897$, $p = .008$, $\eta_p^2 = .012$). Planned contrasts indicate that there was a significant pupillary dilation compared to neutral for the happy expression only ($t(1410) = 2.765$, $p = .006$, $d = .147$). A marginally significant dilation was also found for anger ($t(1410) = 1.905$, $p = .057$, $d = .102$), all other contrasts were non-significant (all $p > .12$). The unadjusted means and 95% confidence intervals for the re-dilation phase are illustrated in Figure 5 (right-side).

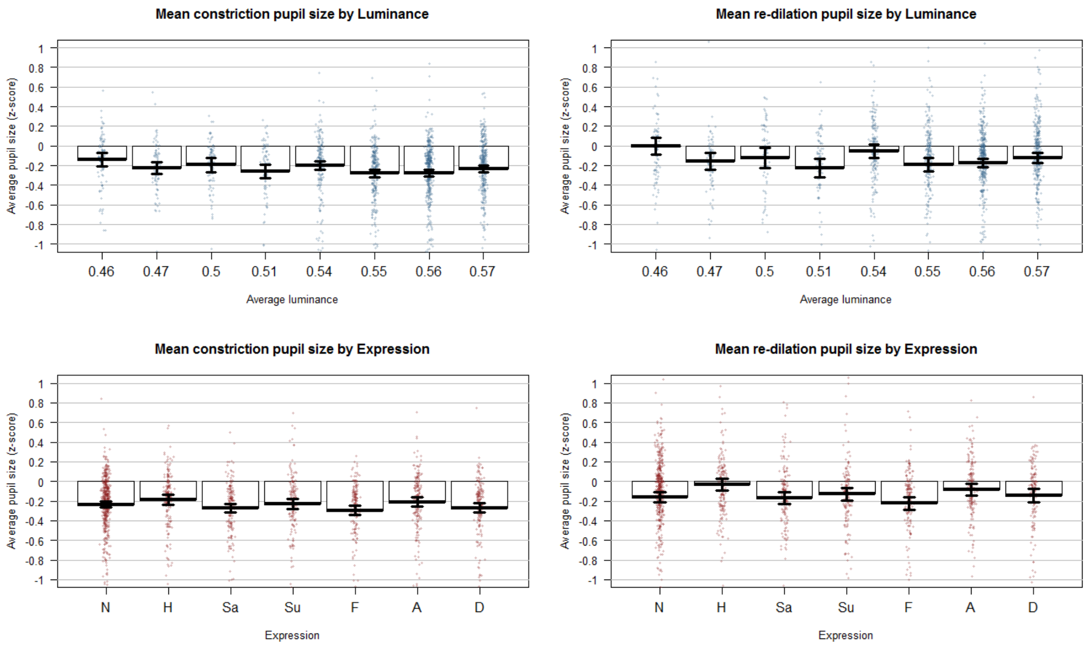

Figure 5. The influence of expression and luminance on pupillary constriction and re-dilation.

The mean change in raw pupil size values (z-scores) from baseline for each trial in the dynamic condition are illustrated. The data are divided by stimulus luminance (top row, blue) and expression (bottom row, red). Luminance was measured on a 0 (100% black) to 1 (100% white) scale. Data from trials (total = 18) with equal luminance and trials depicting the same expressions were collapsed. Trial data are also divided into 'constriction' (0-1000ms, left) and 're-dilation' (1001-3000ms, right) time slots. Bold horizontal lines represent the mean, error bars represent 95% confidence intervals.

Discussion

Evoked pupillary responses are a sensitive index of affective processing in infants (Geangu et al., 2011; Hepach & Westermann, 2013; Jessen et al., 2016), though measuring psychologically-induced changes in pupil size in infant populations presents substantial methodological challenges (Hepach & Westermann, 2016). Infant pupillometry requires a procedure that minimises variations in stimulus luminance whilst simultaneously maintaining an infant's attention to the screen. This chapter implemented two methodological approaches to investigate infants' pupillary responses to emotional expressions. Firstly, in Experiment 1, infants were presented with luminance-controlled stimuli alongside strategies to maintain infant attentiveness. Secondly, in Experiment 2, infants were presented with naturally-engaging, colourful and dynamic stimuli whilst statistically controlling for the effect of varying luminance. Both experiments found that infants' evoked pupillary responses, and thus their sympathetic arousal, varied according to stimulus expression. This indicates that infants can not only discriminate between expressions, but that they also show differential affective responses to the emotional expressions of others.

Experiment 1 found reduced mean pupil sizes for fearful compared to neutral eye-regions. From the FDA analysis comparing neutral and expressive trials as a function of time, it is clear this discrepancy in pupil sizes emerged during the re-dilation period, not during the PLR period where there was also only a minimal response to luminance (see Figure 2). Neutral and expressive stimuli were highly controlled, and did not significantly differ according to mean luminance values. The stimuli themselves were greyscale, unchanging across time, and focused on a particular region (eyes) to discourage diffuse scanning. Pupil data recorded when the infant was not fixating the stimulus (i.e. fixating the background, performing a saccade or looking away from the screen) were discarded prior to analysis, however due to the inclusion of gaze-contingent stimulus transitions and an attention-grabbing video, participant data loss was minimised (15 of an initial 54 participants (27.78%) were excluded from all analyses). These methodological precautions were implemented to ensure that psychologically-evoked, not luminance-evoked, pupillary responses were recorded. Nevertheless, presenting stimuli that are highly-controlled in order to minimise the confounding effects of luminance has substantial disadvantages. Facial expressions do not naturally occur within such contrived and artificial scenarios, so even if adequate ‘clean’ eye-tracking data are gathered, there is a question as to whether these findings will meaningfully generalise to the ‘real world’. A second experiment was therefore conducted to assess whether infant pupillometry is viable with ecologically-valid and naturally-engaging stimuli, whilst statistically controlling for the effect of stimulus luminance.

In the second experiment (using data collected from the experiment in Chapter 5), analysis of evoked pupillary responses for expressive faces collapsed across time found that while pupil sizes for averaged neutral and expressive trials did not significantly differ, pupil size did vary substantially *between* dynamic (but not static) expressive faces. In particular, larger dilations for happy compared to fearful and sad facial expressions were found. These findings were investigated in greater detail by comparing the pupillary responses for each dynamic facial expression against neutral as a function of time using functional data analysis (see Sirois & Brisson, 2014). These plots (see Figure 4) reveal a clear dilatory response compared to neutral for the happy expression only, but for several negative expressions (fear, sadness and disgust) pupil size was reduced compared to neutral during the initial constriction phase (0 to 1000ms). For sadness and disgust, these differences were likely due to a more pronounced pupillary light response as both of these expressions contained trials with high average luminance. However, fear had relatively low average luminance compared to neutral and also showed reduced re-dilation at later points in the time window. This interpretation was investigated by conducting a luminance-adjusted analysis, after which fear (but not sadness or disgust) still showed significantly reduced pupil sizes compared to neutral during the constriction phase (see

Henderson et al., 2014), while happy still showed significantly larger pupil sizes compared to neutral during the re-dilation phase.

Positive social stimuli such as happy expressions are intrinsically rewarding (e.g. O'Doherty et al., 2003), and positively-valenced emotional images can increase sympathetic activity, evoking its concordant pupillometric response (Bradley et al., 2008). Similarly, this study has found that infants show greater sympathetic arousal for happy compared to neutral expressions within simulated dynamic interactions. This finding is consistent with work in children (Sepeta et al., 2012), and recent work with infants (Jessen et al., 2016). However, the current study also found reduced pupil dilatory responses for certain negative emotional expressions such as fear compared to neutral, in both the dynamic expressive face experiment and in the highly-controlled expressive eye-region experiment. Research with adults has found stronger sympathetic responses for negatively-valenced emotional stimuli (Laeng et al., 2013; Yrttiaho et al., 2017), but thus far developmental findings are mixed (Geangu et al., 2011; Hepach & Westermann, 2013; Jessen et al., 2016; Sepeta et al., 2012). It is conceivable therefore that while the familiar and rewarding happy expression reliably triggers a sympathetic response in infants, large emotional reactions to less familiar expressions such as fear may show a more protracted development (see Thomas et al., 2001). Furthermore, for infants a *lack* of an expressive response (i.e. a neutral expression), particularly within a social interaction, may be substantially arousing/distressing as is evident from investigations of the 'still-face' phenomenon (for a review see Adamson & Frick, 2003), and thus may not be a suitable 'baseline' to compare against.

Mixed findings in the literature for infant pupil responses to emotional expressions are likely due to the substantial methodological differences between studies. Previous work has varied in stimulus type (e.g. static and silent vs dynamic and audio-visual), duration (e.g. 2 seconds vs 50 seconds) and content. For instance one study presented static images restricted to facial regions only (Jessen et al., 2016), while another (Hepach & Westermann, 2013) presented video clips of seated actors expressing emotions through behaviours alongside facial expressions (e.g. ‘thumping’ a stuffed animal toy to depict anger). These differences render any comparison of the efficacy of positive and negative expressions to induce emotional arousal in infant observers problematic. The current study also found a clear difference in the magnitude of pupil dilation between dynamic and static displays, with substantially larger ($ŋ_p^2 = .275$) overall pupil sizes and reduced PLR for dynamic expression stimuli (see Henderson et al., 2014). This finding is consistent with previous work using facial electromyography (Sato et al., 2008) and neuroimaging (Kilts et al., 2003; Labar et al., 2003) which has found dynamic expressions illicit stronger responses compared to static expressions. Greater consistency between methods is therefore needed in order to further explore developmental trajectories in the pupillary responses evoked by different facial emotions. Given that findings from the GC Social Interaction facial expression experiment replicated those in the highly-controlled eye-regions experiment (i.e. both experiments found a reduced pupil dilation for fear compared to neutral but a non-significant difference between neutral and angry), and that ‘real-world’ facial expressions occur within dynamic and interactive social contexts, it is recommended that researchers move toward more naturally-engaging and ecologically-valid stimuli, as long as sufficient statistical controls for luminance are implemented.

A variety of statistical methods have been used in this chapter in order to provide a robust and meaningful analysis of infant pupillary responses to facial expression stimuli. The functional data approach (Ramsay & Silverman, 1997; Sirois & Brisson, 2014) is a useful tool to explore pupillometry data without collapsing across time (D. J. Barr, 2008).

Here only simple comparisons were conducted, but future work can use this methodological approach to conduct more sophisticated analyses. For instance, here an average luminance value for each trial was used to represent the entire three-second time window within the luminance-adjusted analysis. However, performing a *functional* ANCOVA which controls for the confounding effect of light using a continuous luminance variable would be a significant improvement.

This chapter has investigated whether pupil size can serve as a useful physiological index for emotional expression processing in infants. While pupil size certainly varies according to psychological factors such as attention or arousal (mediated by the LC-NA system), designing eye-tracking paradigms that can produce good quality data in infants is challenging (see Hepach & Westermann, 2016). This study compared two potential methodological solutions to these challenges. The favoured approach presents infants with naturally-engaging, realistic stimuli alongside robust statistical controls to 'remove' the confounding effect of stimulus luminance. These results suggest that of all the six basic facial expressions (Ekman et al., 1987), infants show a clear dilatory response compared to neutral for dynamic happy expressions only. Given that infants, by seven months, can *perceptually* discriminate and categorise expressions (Ludemann & Nelson, 1988; Nelson et al., 1979), these results question the extent to which perceptual maturity reflects an infant's ability to understand and respond appropriately to the communicated meaning behind facial expressions. This *conceptual* maturity in processing facial expressions likely continues to be refined across development (e.g. Ruba, Meltzoff, & Repacholi, 2019).

Chapter 8

General discussion of methods and findings

Conducting robust and relevant research is uniquely challenging when investigating infant participants given their limited linguistic and behavioural repertoires. However, infant looking behaviours are not random, and as oculomotor precision develops rapidly, the location of an infant's gaze can be a reliable index of their attention and interest (see Buswell, 1935; Yarbus, 1967). This fact was exploited by Fantz (1956, 1963, 1964), whose methodological innovations (centred on recording the duration of infant looking to one exemplar verses another) lead to a 'golden age' of infant experimental research in the 70s and 80s which laid the foundation of our current understanding of infant perceptual abilities (see Braddick & Atkinson, 2011). Preferential looking paradigms were initially designed to answer simple, low-level questions about infant visual perception, but their success lead researchers to apply these methods to increasingly complex and higher-level cognitive questions. These were questions which were concerned about what infants know, not what they can see. Since the 90s and early 2000s infant research has been going through a crisis of conscience; critics have begun to wonder if many infant studies are guilty of an 'overly rich' interpretation of the data (e.g. Haith, 1998), and have begun to question the meaningfulness and validity of looking time measures for investigating infant cognitive abilities (e.g. Aslin, 2007; Heyes, 2014). Meanwhile, a growing sense of uneasiness about the lab-based experimental approach has been simmering beneath the surface since the late 70s (e.g. Bronfenbrenner, 1977; Gibson, 1979). The controlled and tidy stories from laboratory studies did not seem to fit the complex, multi-faceted and messy 'real-world' environments in which infants develop, learn, perceive and act in every day.

Today, developing technologies once again offer the opportunity for methodological innovation. The lasting value of our findings ultimately depends on the validity of our methods. It depends firstly on whether the behaviours we are recording are a true index of an infant's internal world (i.e. can we use eye movements to investigate infant cognitive abilities?), and it depends secondly on whether the stimuli we present are a true index of an infant's external world (i.e. are our stimuli and methods representative of real-life scenarios?). The primary aim of this thesis has therefore been to investigate current topics in infant socio-cognitive research with the experimental rigour of a controlled lab setting, whilst also ensuring our results and subsequent experimental interpretations generalise meaningfully to the real-world. To achieve this, the empirical chapters have implemented several novel and recently-developed methodological tools to investigate infants' sensitivity to other's social behaviours embedded within simulated interactions. The GC Social Interaction paradigm (see Chapter 2) has been designed to replicate the essence of real-world interactions in the lab (e.g. dynamism, contingency and communicativeness). Recent investigations using similar dynamic and contingent paradigms (Keemink, Keshavarzi-Pour, et al., 2019; Vernetti et al., 2018) have noted similar rates of reciprocal responses from infants in comparison to real-world interactions. This suggests that the GC Social Interaction paradigm has a high degree of authenticity and 'believability', and is a much closer representation of everyday interactions compared to the static, unresponsive images often used in studies of infant social cognition.

Infants' preferences for behaviours within social interactions

It is important to embed our experimental stimuli within their natural context as contrived, unrealistic or artificial representations of the phenomena we are purporting to investigate might give different results, and therefore lead us to draw conclusions that do not generalise to real-world environments. Bahrick and colleagues (2002) provide an excellent example. In their study they compared infants' ability to encode and remember identities of others when they were presented either as static images, or within dynamic videos in which they performed repetitive behaviours (e.g. brushing hair or teeth). They found that although infants could discriminate and recognise faces presented as static images, during the dynamic condition, infant attention was diverted away from facial characteristics such that only their behaviours were discriminated and recognised. Naturalistic environments are busy; infants don't often have the opportunity to study still, unresponsive faces presented to them side-by-side with a neutral, direct gaze for an extended period of time (see also Burton, 2013). In Chapter 3 it was noted that much of what we understand about how infants encode race relies on static face presentations, and it was reasoned that isolating race in such a contrived fashion by stripping away all other variables that are always (or nearly always) present might artificially inflate its importance. It was hypothesised that the social behaviours of others would possess greater salience for infants relative to race (c.f. Bahrick et al., 2002) when face stimuli were presented embedded within the naturalistic context of a social interaction (i.e. the GC Social Interaction paradigm).

The first experiment in Chapter 3 replicated and extended the pattern of 'spontaneous' race preferences previously reported in the literature (Fassbender et al., 2016; Liu et al., 2015) using neutral and unresponsive, but dynamic (blinking) faces. Combined with previous work (Kelly et al., 2005) it shows that newborns demonstrate no preferences for race, but younger infants (3- and 6-month-olds) look longer to own-race (familiar) faces, and older infants (9- and 12-month-olds) look longer to other-race (novel) faces. Yet studying 'spontaneous' preferences to blank, blinking faces still tells us little about how infants encode and react to race in social settings. Experiment 2 therefore repeated the visual preference task from Experiment 1, but first allowed infants to familiarise themselves with the on-screen actors via the GC Social Interaction paradigm. Both actors gave identical contingent responses (either smiles or head-turns away depending on the condition). If infant race preferences are robust and generalizable to naturalistic environments, they should still emerge following these brief interactions. The results indicated that following familiarisation there was no significant main effect of race on stimulus looking times ($ŋ_p^2 < .001$), yet when calculating preference as a proportion of total looking within individual age groups, nine-month-old other-race preferences proved resilient (significant but with reduced effect size), while six- and twelve-month preferences were at chance (approximately 50%). This is a curious finding. It does suggest firstly that infant race preferences are different when embedded within social interactions, and that the social behaviours of others, even when identical, are sufficient to remove (6- and 12-month-old) or reduce (9-month-old) race preferences. Yet it also highlights that nine-month-old infants might be particularly sensitive to other-race faces, which may coincide with a more general preference for novelty over familiarity at this age (see Quinn, Lee, & Pascalis, 2019).

It was also noted within Experiment 2 that infants showed greater interest (longer looking) within the condition where the on-screen actors gave smiling, engaging responses compared to the condition where actors turned their head away. Positive, prosocial behaviours such as maintaining mutual gaze and smiling are intrinsically rewarding social stimuli (O'Doherty et al., 2003), and are often preferred by infants (e.g. Kuchuk, Vibbert, & Bornstein, 1986; La Barbera, Izard, Vietze, & Parisi, 1976). Experiment 3 therefore increased the salience of the behaviours within the interactions by presenting infants with actors who responded differently to one another. One actor consistently responded engagingly, the other non-engagingly, and their roles were counter-balanced across conditions. It was hypothesised that infants would encode the behaviours of the actors, and thus would show a visual preference for the actor who responded engagingly, regardless of their race. The results of this experiment showed that no significant preferences for race or behaviour were found across both conditions and all age groups. This was not expected. Infants were highly attentive to these behaviours; analysis of infants' face-scanning showed they consistently looked toward smiles and followed the gaze of the characters on the screen. Why would infants be selectively attending and scanning these behaviours if they were not meaningful? One explanation could be that as infants reliably followed the actor's gaze, they might actually be interpreting the behaviour as a communicative gesture, and not as a symbol of disinterest, and thus would find both responses equally interesting during preference. Secondly, to consider low-level perceptual factors, infants might be sensitive to the higher motion in this head-turn behaviour (see Vinter, 1986), and this could have interfered with preferences for the smile response. A third explanation could be that differences in the race of the actors were confounding and negating preferences for prosocial behaviours. And finally, a fourth explanation might be that due to the gaze-contingent manipulation, infants during the (non-interactive) preference task divided their attention equally between each actor in the hope of 'triggering' a response.

These four potential explanations were investigated across three experiments in Chapter 4. The first experiment was identical to the third experiment in Chapter 3, except that both actors were the same race as the infants (Caucasian). It was thought that if race were a confounding factor, preferences for positive social behaviours would be found once the effect of race was 'removed'. In fact preferences were again at chance, and there were no looking duration biases during the interactions. Experiment 2 then tested two other explanations, that perhaps the non-engaging behaviour (head turn away) was interesting to infants, either due to its communicativeness or its high level of motion. An alternative non-engaging behaviour was therefore chosen in which the actor simply closed their eyes when the infant initiated eye contact. Infants' preferences were again at chance. Finally, in a third experiment, the interactivity of the non-engaging stimulus was removed. Infants were presented with a responsive, smiling social partner alongside an unresponsive actor with neutral affect and random blinking. Infants are highly sensitive to contingency (Bigelow & Birch, 1999), perhaps any response might be interesting to infants as long as it is contingent. Further, the contrast of a non-interactive preference phase following the GC Social Interactions might also cause infants to divide their attention evenly between stimuli in the hope of triggering a response. However, in this third experiment, as in Experiments 1 and 2, infant preferences for the actors were at chance level.

The results from the experiments in Chapter 4 are puzzling. It seems highly unlikely that infants would not show preferences for contingent, dynamic, positive social behaviours over static, neutral and unresponsive behaviours. The sample numbers within these experiments were relatively low due to the fact that it became obvious that the effects were small to non-existent (average *d* = .207), and so these experiments were cut short. To remedy this potential lack of power, a combined analysis was conducted across the three experiments (*N* = 119). This analysis confirmed preferences for behaviours performed during the GC Social Interaction were indeed at chance level (*M* = 49.27%, *d* = .165) and did not vary across age. As infants were highly attentive to these actions it is conceivable that infants' ability to encode facial identity may be disrupted by dynamic behavioural information (c.f. Bahrick et al., 2002), and were therefore unable to show a visual preference during the comparison task. This cannot always be the case, as nine-month-olds were able to demonstrate an other-race preference within Experiment 2 of Chapter 3, but the actors from the experiments in that chapter differed according to race, and thus facial differences were more pronounced making it an easier task to discriminate between them.

Limitations of spontaneous preference paradigms

The results from Chapter 4 also highlight the methodological limitations of visual preference procedures, particularly when investigating spontaneous preferences. Within these forced-choice comparison tasks, if an infant shows significantly longer looking toward one stimulus over another (regardless of side) it is good evidence that they are able to perceptually discriminate between them. Yet the question of *why* they look to one over another is not always easy to answer (see Cohen, 2004). If a standard infant-controlled habituation paradigm is used, where a stimulus is presented repeatedly till infant looking drops by a sufficiently stringent pre-defined criteria (e.g. by 50% compared to the first trial), then any preferential looking in the comparison task will likely be due to novelty. Yet when the question is interested in infants' intrinsic, a priori preferences (i.e. not preferences induced by habituation; 'spontaneous') infant looking can be harder to interpret. If infants do show a significant spontaneous bias for one stimulus over another, this could be attributed to either familiarity or novelty. It could also be attributed to either low-level perceptual or high-level cognitive aspects of the stimulus. Now classic violation of expectation (VoE) tasks (e.g. Baillargeon, 1986; Wynn, 1992; see Chapter 1) attribute increased looking to 'impossible' events as a violation of their cognitive expectations, yet the same results can sometimes be explained via a preference for perceptual familiarity (Cohen, 2004; Haith, 1998). Infants' intrinsic preferences might also differ between familiarity and novelty according to their age (Quinn et al., 2019), as is evident in Experiment 1 of Chapter 3, which might interfere with the interpretation of results if a perceptual bias for novelty is assumed (as in VoE paradigms).

In the present case, novelty and familiarity were easily disambiguated for race. The participants were from Caucasian families in a predominantly Caucasian region of the UK (see Chapter 2 for Kent's demographic information), thus preferences for Caucasian faces could be attributed to familiarity, and African faces to novelty. Preferences for social behaviours are less straightforward however. We might hypothesise that smiling, socially-engaging behaviours carry intrinsic reward value and therefore would be preferred, yet the behaviours of a consistently disengaging adult might appear more novel, both cognitively and perceptually, and thus might also be preferred. Furthermore there is no way of determining for certain which behaviours are familiar or novel for an infant without further information about their home environment. The situation becomes more complicated when no preference is found, and particularly within more ecological paradigms, such as the GC Social Interaction paradigm used here, where there are several variables to consider (e.g. behaviour, contingency, dynamism, encoding), so interpretation of null effects can pose an even greater challenge (see Chapter 4). And while a null preference does not necessarily mean that infants' cannot perceptually discriminate between the two stimuli, it *does* mean that infants' abilities to discriminate, and also their abilities to encode and recall stimulus characteristics (e.g. previous behaviours), can now no longer be assumed.

There is also a deeper question concerning the logic behind spontaneous preferences. Usually, such preference tasks strip away all external stimuli to compete two exemplars along one critical dimension (e.g. race). Yet if infants' show a preference in such a contrived scenario, in the absence of top-down knowledge or any other external input, is this result meaningful? Outside of the lab there will always be complementary and competing sensory information that will almost certainly influence what infants' visually attend to and how those aspects of their environment are processed. In fact, the preferences infants' demonstrate for one characteristic over another within artificial scenarios tell us little about infants' conceptual understanding, nor does it reflect how they might react in a social setting. Thus the most important message from Chapters 3 and 4 is that infant preferences for both race and social behaviours are fundamentally different within simulated interactions. This might reflect a disruption of an infants' ability to encode facial characteristics whilst simultaneously attending dynamic behaviours. Yet whatever the reason, the fact that they are different within a paradigm that is a more faithful representation of an infants' natural environment, suggests that the findings from studies presenting static, non-responsive stimuli might not generalise to real-world environments.

Infants' scanning of dynamic faces

Infants' visual attention *within* a stimulus was also examined in Chapter 3 using pre-defined AOI regions and descriptive heatmaps. Aside from small differences in infants' scanning of own- and other-race faces (particularly in nine-month-olds), infants' visual attention was primarily determined by the behaviour of the actors on the screen. For instance, if they smiled, infants' attention would drop to the mouth region, if they turned away, infants would track the movement of their eyes as they moved to the left or to the right. Most infants' (approx. 90%) exhibited these patterns of social scanning. The fact that these behaviours were dynamic and contingent likely enhanced their salience relative to traditional stimuli. Chapter 5 investigated this possibility explicitly by comparing infants' scanning of dynamic, interactive facial expressions to their scanning of unresponsive, static expression stimuli. The vast majority of research on infants' perception of facial expressions uses static stimuli, despite the fact that expressions are inherently dynamic phenomena (see Krumhuber, Kappas, & Manstead, 2013). Unsurprisingly, recent work using dynamic displays of emotion have produced different results to static stimuli (e.g. Addabbo, Longhi, Marchis, Tagliabue, & Turati, 2018; Heck, Hock, White, Jubran, & Bhatt, 2016; Soussignan et al., 2017), yet no previous work has presented all six basic dynamic facial expressions, nor has both dynamism and contingency been included before within facial expression stimuli. The present work achieves this by investigating all six basic expressions (happy, sad, surprise, fear, anger and disgust) within the GC Social Interaction paradigm.

The results from Chapter 5 indicate that infants attend to static and contingently-triggered dynamic facial expressions differently (contingent/dynamic displays will be henceforth referred to as just 'dynamic'). Infants look longer toward dynamic displays (c.f. Wilcox & Clayton, 1968), and show increased scanning of the lower facial features of dynamic expressions. Scanning patterns were explored in greater detail within the unsegmented analyses (heatmap analyses, mixed-effects models and cluster permutation analysis), highlighting precise spatial and temporal differences between static and dynamic expressions (see Chapter 5 for more information). As noted above, infants do not usually encounter static and unresponsive faces in their day-to-day lives. This is not a trivial point as these findings once again force us to question the ecological validity of current methods and the extent to which results from studies that have used static stimuli can reliably generalise to behaviour outside of the lab.

Infant scanning of facial expressions occurring within simulated interactions is also different in kind to traditional paradigms. Infant scanning of dynamic expressions is *reactive*, with attention being driven toward the lower face as expression-related changes emerge across time. Dynamic stimuli provide additional diagnostic motion cues which are disproportionately communicated through the lower facial features (relative to upper facial features; see Krumhuber et al., 2013 for a review), enhancing the salience of this facial area in comparison to static stimuli (see Jack, Garrod, & Schyns, 2014). Infant face scanning of naturalistic stimuli is also highly sensitive to individual differences in facial morphology and expression production. Recent work has shown that the traditionally held view of a 'default' scanning pattern between the eyes and mouth does not exist (e.g. Arizpe, Walsh, Yovel, & Baker, 2017). The lack of a default 'triangular' scanning pattern is especially true when observing dynamic faces, where motion cues direct attention to salient regions of the face (e.g. Lewkowicz & Hansen-Tift, 2012; Võ et al., 2012). The findings presented in Chapter 5 (but see also Chapter 3) indicate that infants demonstrate 'bespoke' scanning patterns which are dependent not only on the facial expression being observed, but also on the actor's individual facial morphology and idiosyncrasies in their performance of communicative social behaviours such as facial expressions.

Given that facial expressions are scanned differently when embedded within simulated interactions, future work should therefore present ecologically relevant stimuli when investigating infants' perception of facial affect. Building on these findings, Chapter 6 investigated developmental differences in dynamic facial expression interest (i.e. time on face) and scanning (i.e. time looking at particular facial regions) using the dynamic AOIs. Previous work with static stimuli has found that younger infants (i.e. under 7 months) show greater interest in positive expressions, which later transitions to greater interest in negative, threat-related expressions, and particularly in fearful expressions (Peltola, Leppänen, Mäki, et al., 2009; Vaish et al., 2008). The current analysis found, similar to a previous study that presented sequential, dynamic expression stimuli (Soussignan et al., 2017), that infants show no particular looking bias for fearful expressions, though infants did look at fearful expressions longer than neutral and angry expressions. This suggests that there is not a uniform 'negativity bias' for dynamic stimuli; infants show interest in fearful expressions, but show little interest in dynamic anger.

Instead, infants looked at surprise longer than any other expression. Few studies have researched surprise, and those that have presented static stimuli (e.g. Serrano, Iglesias, & Loeches, 1992). Dynamic surprise, with its wide eyes and wide mouth might present a greater amplitude of motion compared to other dynamic expressions (e.g. fear), and thus might be highly engaging for infants, and perceived as having a high intensity (see Biele & Grabowska, 2006). Investigating infant development, this analysis also found *relative* increases in interest toward negative-valence expressions such as dynamic anger by nine months, and toward dynamic disgust by twelve months (c.f. Ruba, Johnson, Harris, & Wilbourn, 2017).

Investigating expression *scanning*, this chapter found that infant looking was again primarily defined by the actor's behaviour (c.f. Chapter 3), and infants, like adults, consistently directed their visual attention toward diagnostic facial regions (i.e. regions most informative for disambiguating expressions; see Gosselin & Schyns, 2001; Jack et al., 2014; Smith, Cottrell, Gosselin, & Schyns, 2005). For example, there was a pronounced peak in lower-feature looking as the mouth widened in surprise, but remained predominantly in the eye-region when the eyes widened in fear. These expression specific differences were large and relatively consistent across age groups, though some 'refinement' of scanning was noted, with older infants' showing increased scanning precision for negative emotions (e.g. fear, anger and disgust). There was also an indication that older infants show an improved ability to attend to regions that are only diagnostic within dynamic expressions (e.g. the mouth region in anger and fear). Yet overall the results indicate that infants' by six months already show sophisticated patterns of expression scanning, particularly for more familiar expressions such as happiness and sadness.

If infants look to facial regions that are most informative for categorising expressions, does this mean they *perceive* and *understand* them in the same way as adults? The answer to this question is likely no. Current work suggests that the understanding of emotional expressions shows a protracted development (even into the teen years) and requires a combination of different contextual cues (emotion labels, situational causes, behavioural consequences etc.) before the specific affective meaning is learned (see Widen, 2013 for a review). From the previous chapters it is clear that infants' are highly sensitive to motion information within dynamic presentations, yet regions of high motion usually correspond to regions that are most useful for perceiving expressions (e.g. think of the scrunched nose in disgust, the widening smile in happy, the opening mouth in surprise etc.). Therefore infants might be attending these regions based on their low-level perceptual information (e.g. motion) and not because they are extracting the expression's conceptual meaning. Recently, work has begun to distinguish *perceptual* categories from *conceptual* categories of emotion (Ruba et al., 2019). Perceptual categories (e.g. visually distinguishing a sub-section of similar expressions with scrunched noses, i.e. disgust) are thought to develop from seven to twelve months (e.g. Ludemann & Nelson, 1988; Ruba et al., 2017), before undergoing a process of 'enrichment' through language and experience in which they transform into conceptual categories (i.e. perceptual categories that are associated with their affective meaning; see Ruba et al., 2019). Does this mean therefore that infants in their first year derive no conceptual meaning from facial expressions whatsoever? It is thought that infants might possess 'broad' conceptual categories (e.g. positive vs negative valence; high vs low arousal) which are then later refined in childhood (Barrett, 2017; Widen, 2013). For now though the literature remains in disagreement, and some researchers believe specific conceptual categories might also begin to develop during later stages of infancy (e.g. Ruba et al., 2019; Walker-Andrews, 1997).

Convergent methods: Pupillometry and eye movements

Since the 60s pupil size has been a known correlate of infant cognitive or affective arousal (Fitzgerald, 1965), and has shown potential to become a complementary measure alongside eye movements. Yet recording pupil size at that time was an arduous task, requiring the experimenter to measure the pupil for each image individually (at a sampling rate equivalent to 2Hz) and by hand. Today, modern eye trackers automatically record pupil size alongside gaze direction at a high sampling rate (e.g. 500Hz), and recent work has suggested pupillary dilations are a sensitive index of an infant's affective response to the emotional expressions of others (e.g. Jessen, Altvater-Mackensen, & Grossmann, 2016; Wagner, Luyster, Tager-Flusberg, & Nelson, 2016). Yet pupillometry still poses substantial methodological challenges for infancy research (see Hepach & Westermann, 2016). Pupillary responses to light are typically much larger (over 100%) than psychologically-induced changes in pupil size (approx. 20%; Beatty & Lucero-Wagoner, 2000), therefore accurate pupil size recording requires that the confounding effect of light is minimised (see Chapter 4). However, stimuli designed to minimise the confounding effects of light, such as static, silent, grey images, may not be sufficiently engaging to hold an infant's attention and are thus detrimental for collecting good quality pupil data. Chapter 7 therefore attempted to develop an emotional expression paradigm that methodologically controlled for the confounding effect of luminance (e.g. using greyscale, luminance-matched, static eye-region images), that was also sufficiently engaging to minimise data loss (e.g. using gaze-contingent stimulus transitions, and video breaks). However, the chapter also analysed pupil data from the 'expressive faces' GC Social Interaction paradigm conducted in Chapter 5, whilst attempting to *statistically* control for the confounding effect of luminance. The results from both approaches were then compared.

The results from the first highly-controlled ‘expressive eyes’ experiment found reduced pupillary re-dilation for fearful eyes relative to neutral eyes (only fearful and angry expressive eyes were included). Results from the second GC Social Interaction ‘expressive faces’ task, after controlling for the effect of luminance, also found reduced pupil sizes for fearful faces compared to neutral (during constriction), but additionally found significantly larger pupil sizes for happy compared to neutral (during re-dilation). This replicates the findings of Jessen and colleagues (2016), who found that infants produce larger pupillary dilations for static happy compared to static fearful expressions, and suggests that infants show differential sympathetic arousal responses dependent on the facial expressions of others.

Using the neutral expression as a baseline, only the happy expression evoked a clear dilatory response, indicating that observing happy and engaging emotional faces within dynamic and contingent interactions can increase sympathetic activity in infants. This finding is consistent with pupillometry studies in both adults (Bradley et al., 2008) and older children (Sepeta et al., 2012) and suggests that findings with adult participants that indicate smiling expressions are intrinsically rewarding social stimuli (e.g. O’Doherty et al., 2003) extend to infancy. This is an important finding as it also helps to clarify the results from eye tracking experiments in earlier chapters. For instance, Chapter 4 found infants did not show a preference for socially-engaging over non-engaging behaviours. These results indicate that this was likely not due to any lack of affective response in infants, but instead due to the disruption of identity encoding (c.f. Bahrick et al., 2002). Further, Chapter 6 found that infants directed their visual attention toward diagnostic regions of all six basic emotional expressions, yet the question was still open as to whether they possessed any conceptual understanding of the affective meaning behind these expressions (see Ruba et al., 2019). As infants only showed a reciprocal physiological reaction for the dynamic happy expression, this suggests that it was only when viewing this expression that any affective meaning was received by infants. Other current work from our lab using the GC Social Interaction paradigm (Keemink, Prunty, Wood, & Kelly, 2019) suggests that of the six basic expressions, infants also only show behavioural responses (e.g. smiles, vocalisations) to dynamic happy faces, not other expressions. Considering this alongside the finding that infants' scanning strategies for happy expressions appear stable by six months of age, infants in their first year might therefore be already developing a conceptual representation for the familiar, positive expression of happiness, but they might not yet possess specific conceptual understanding for other facial expressions. For less familiar emotions such as disgust, developmental differences in scanning (see Chapter 6) suggest that infants are likely still forming and refining their perceptual categories of these emotions toward the end of their first year.

Previous commentators have theorised that during their first year, infants might possess 'broad' conceptual categories of emotion (e.g. positive vs negative valence) which are later refined (Barrett, 2017; Widen, 2013). While the only positive expression (happy) was distinguished from the other emotions, the findings from Chapter 7 (and also Chapter 6) do not support the existence of a homogenous 'negative' conceptual category. For instance, infant pupillary responses to angry and fearful expressions were quite different; dynamic anger produced marginal pupil dilations (though see Experiment 1 of Chapter 7), whilst fear produced significantly reduced pupil sizes in comparison to neutral. Previous work in adults has shown that both positive and negative-valence emotional stimuli produce dilatory responses compared to neutral (Bradley et al., 2008), with stronger dilations for negative emotions (e.g. Yrttiaho, Niehaus, Thomas, & Leppänen, 2017). Given that pupil sizes were significantly reduced for fearful expressions compared to neutral in both Experiment 1 and 2, despite considerable difference between the stimuli presented, it suggests that infants show sympathetic sensitivity to fearful expressions, though the meaning of a reduced dilatory response in infancy is not easy to interpret, and pupillometry researchers (e.g. Bradley et al., 2008; Partala & Surakka, 2003) caution against interpreting bi-directional effects of emotion on pupil size (c.f. Hess & Polt, 1960). Similar to visual preference procedures, where a significant effect can be interpreted as either a novelty or a familiarity preference, the underlying cause of pupillary responses can also be ambiguous. It is also conceivable that for infants a lack of an expressive response (i.e. a neutral expression), particularly within a social interaction, may be arousing and distressing as is evident from investigations of the ‘still-face’ phenomenon (for a review see Adamson & Frick, 2003), and thus may not be a suitable ‘baseline’ to compare against.

The second experiment of Chapter 7 also compared pupil sizes for dynamic and static facial expressions (the pupil data were recorded alongside the eye-tracking data in Chapter 5). Chapter 5 found significantly longer looking toward the dynamic and contingently-animated stimuli compared to the unresponsive, static stimuli. Chapter 7 also found infants responded to the same stimuli with substantially larger pupil sizes for dynamic, compared to static, expressions ($ŋ_p^2 = .275$). This convergent evidence indicates that dynamic expressions are both more interesting (looking duration) and more emotionally affecting (pupil size) for infants (see LaBar, Crupain, Voyvodic, & McCarthy, 2003; Sato, Fujimura, & Suzuki, 2008). Evoked sympathetic responses were also similar across static expressions, suggesting infants’ reactions showed less differentiation between static stimuli. Given the large variation in methods used within infant pupillometry thus far (see Chapter 7), greater consistency across methods will be necessary to compare results effectively. The findings in this chapter suggest that as the results from the GC Social Interaction paradigm were substantially different (yet the findings were still consistent with the luminance-controlled experiment), and that facial expressions naturally occur within dynamic and interactive social contexts, it is once again recommended that researchers move toward more naturalistic stimuli and methods. There is also clear merit in using a convergent approach where the complementary measures of eye movements and evoked pupillary responses are recorded simultaneously.

Methodological tools

It is clear from the empirical research surveyed above that using naturalistic stimuli produces different results, and as these methods are a closer representation of infant's day-to-day experiences, the findings generated within paradigms such as the GC Social Interaction likely possess greater *relevance* for real-world environments. However, these findings also demonstrate that deviating from static stimuli and simple analyses which collapse data across time and space (c.f. Wilcox & Clayton, 1968) provide a much more complex picture of facial expression scanning. It may be tempting therefore to continue to reduce social stimuli within artificial and over-simplified formats. However, this approach greatly sacrifices the meaningfulness and generalisability of our data. Instead, this thesis has proposed that we must strive to develop more advanced methodological and analytical techniques that enable us to represent these social phenomena as faithfully as possible. In this thesis, several methodological tools have been presented that enable us to analyse the rich and complex data generated within social interaction eye-tracking paradigms. Specifically, the empirical chapters have introduced novel techniques such as dynamic AOIs and statistical heatmaps and also applied existing techniques such as growth-curve analysis, cluster permutation analyses (Dink & Ferguson, 2015), and functional data analysis (Sirois & Brisson, 2014) to infant research. The appropriateness of these methods will now be assessed using empirical examples from experiments within this thesis.

Dynamic AOIs

In response to the fundamental challenge of defining precisely *where* an infant is looking upon a dynamic stimulus a tool has been developed by the author that automatically generates AOI regions using information inherent within each frame of a video stimulus (see Chapter 5; c.f. Hessels, Benjamins, Cornelissen, & Hooge, 2018). A major advantage of this method is that it is data-driven. It therefore avoids the substantial issues that arise from using fixed and arbitrarily-defined interest regions (see Caldara & Miellet, 2011), and automatically adapts to temporal changes in the stimulus. However, although this method represents a significant step forward in the literature (c.f. Hunnius & Geuze, 2004; T. Wilcox, Stubbs, Wheeler, & Alexander, 2013), it is not without its limitations. The dynamic AOIs use pre-defined colour and luminance information for the actor's skin to determine the 'face AOI'. This accurately defined the outer contour of the face and neck, differentiating it from hair, clothing and background. However, various stimulus aspects (e.g. an actor's facial hair) generated confounds that required further custom scripting to overcome. Facial features were also defined in contrast to the colour and luminance of the skin, such that visually distinct features were defined with greater accuracy and consistency (e.g. the nostrils were detected more readily than the bridge of the nose; see Figure 1).

Figure 1 compares the efficacy of fixed and dynamic AOIs in defining the location of fixations on facial features. Fixations captured by both fixed and dynamic AOIs are displayed in black and although they show considerable overlap, there are also marked differences. Fixations captured by fixed AOIs but not dynamic AOIs (blue) represent the imprecision of fixed interest areas as fixations are included across a much broader region. However, fixations captured by the dynamic AOIs but not the fixed AOIs (red) reflect looking to an automatically defined nose AOI and greater accuracy across time within dynamic AOIs (the fixed AOIs were defined for the expressive frame, and are therefore inaccurate in the earlier neutral frame, see Figure 1). Depending on the spatial precision of eye-tracking and the rate of data loss, using wider AOIs such as entire upper and lower facial regions may be more appropriate (see Chapters 5 and 6 for examples). Indeed, in the absence of a specific and directed hypothesis, it is important to question the wisdom of ever arbitrarily carving a face stimulus into multiple discrete units. This approach is not only haphazard, but from a statistical standpoint it also dramatically increases the likelihood of generating a main effect of AOI within an ANOVA as certain areas of the face (e.g., forehead, cheeks, and chin) are rarely, if ever fixated (see Hills & Pake, 2013, for an 8 AOI example). Fixed AOIs may also be sufficient if interest areas do not change substantially across time. In its current form, the dynamic AOIs use a fixed and pre-defined midpoint on the vertical axis to differentiate upper and lower regions. Future revisions of this method might consider an automatically generated midpoint, as well as a midpoint on the horizontal axis to compare differences in looking to the left and right of the face where relevant.

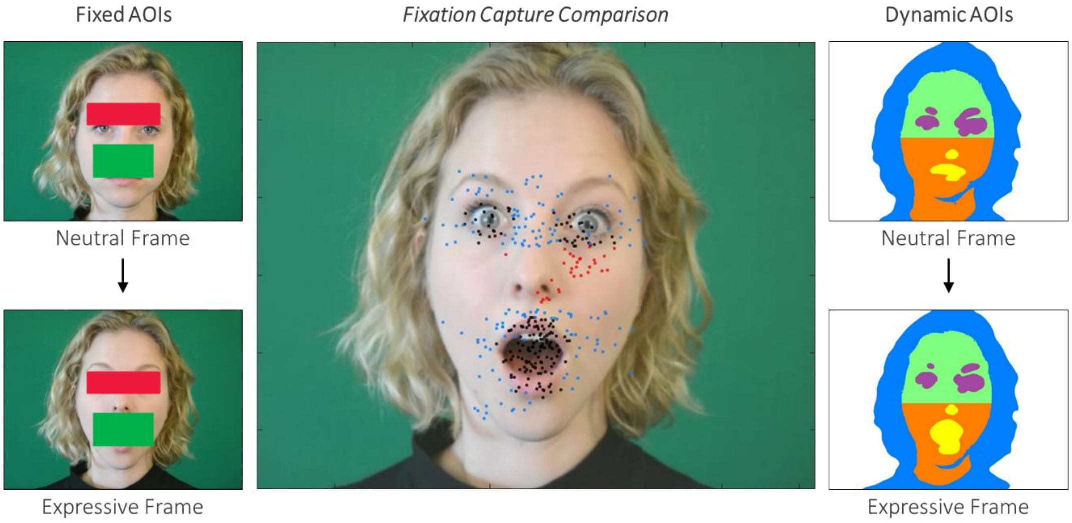

Figure 1. The efficacy of fixed and dynamic AOIs

An illustration of the efficacy of fixed and dynamic AOIs in ‘capturing’ fixations upon internal facial regions of video facial expression stimuli. Fixed mouth (green) and eye (red) region AOIs defined for an expressive frame of a surprised face stimulus are not accurately located in an earlier neutral frame (left panel). Dynamic upper (purple) and lower (yellow) AOIs automatically adapt to stimulus changes across time and identify facial regions with greater precision (right panel). The central figure displays the location of fixations from a surprised trial within the dynamic condition of the experiment in Chapter 5 (*N* = 101). Fixations defined as being within a facial feature by both fixed and dynamic AOIs are coloured black. Fixations captured by fixed but not dynamic AOIs are coloured blue and fixations captured by dynamic but not fixed AOIs are coloured red.

Temporal analysis

When investigating dynamic social behaviours such as facial expressions which unfold over time, it becomes critical to observe how the infant responds to a constantly changing stimulus. Therefore fully collapsing or partially segmenting data across time can also obscure important patterns (D. J. Barr, 2008). Chapter 6 explored developmental differences in expression scanning, first by collapsing across time and conducting an ANOVA, and secondly by representing changes across time by using orthogonal power polynomials within mixed-effects modelling (via the eyetrackingR package; Dink &

Ferguson, 2015). The collapsed analysis found significant differences in scanning across expressions, but not across age groups. The temporal analysis, however, found similar diagnostic looking in all age groups, but also found subtle developmental differences in facial expression scanning, such as a transition toward greater diagnostic lower-feature looking for disgust between nine and twelve months, presenting a more pronounced 'inverted-U' curve over time (Figure 2a). These more fine-grained patterns were unidentified in the collapsed analyses and complement previous work (e.g. Ruba et al., 2017). These models are powerful statistical tools, and can incorporate multiple variables whilst also controlling for random effects (e.g. trial and participant variance). However, large models can take a long time to fit and maximal random effects structures can often lead to non-convergence (Baayen et al., 2008; D. J. Barr et al., 2013). Successful models also generate a complex pattern of findings, particularly if multiple polynomial terms are used to define time (see Chapter 6 tables 2 – 4 for 'highly condensed' summaries of findings). Mixed-effects modelling should therefore be approached with caution, and may not be suitable for all research questions.

Cluster permutation analysis (Maris & Oostenveld, 2007) provides a useful complement to mixed-effects models. It was used in Chapter 5 to identify specific time-windows in which static and dynamic conditions diverged. Two expressions that found significant effects within the collapsed AOI analysis (i.e. happy and angry) were analysed in more detail using mixed-effects modelling, but for the other expressions only cluster permutations analysis was used. For several of these expressions (e.g. sadness, surprise, fear) average lower-face looking for the trial was similar for static and dynamic expressions, yet scanning patterns did differ across time (e.g. a later 'peak' in lower-face looking for dynamic coinciding with the emergence of the expression; Figure 2b). Cluster permutation analysis was therefore useful for identifying informative temporal differences when performing simple comparisons (e.g. investigating differences between two conditions).

Functional data analysis (Ramsay & Silverman, 1997), is a statistical method that has been recommended for analysing pupillometry data (Jackson & Sirois, 2009; Sirois & Brisson, 2014). This method uses b-splines to express continuous data as a function of time and is able to elegantly summarise complex temporal patterns. This method was used in Chapter 7 to compare evoked pupil sizes for expressive faces and eyes against a neutral baseline (Figure 2c). Statistical tests were conducted on the functional curves themselves, producing values (e.g. mean, standard deviations, standard error, *t*-scores etc.) that are themselves represented as functional curves across time. Using this method it is easy to visualise temporal patterns and to identify when significant differences emerged. The flexibility of this method means that, unlike cluster permutation analysis, more advanced statistical tests can be performed (e.g. using the functional splines within an ANOVA). And although FDA analysis has been implemented successfully in infant pupillometry research (e.g. Geangu, Hauf, Bhardwaj, & Bentz, 2011; Jackson & Sirois, 2009), given that it succinctly visualises and analyses temporal data, it might also be considered for use in infant eye-tracking research.

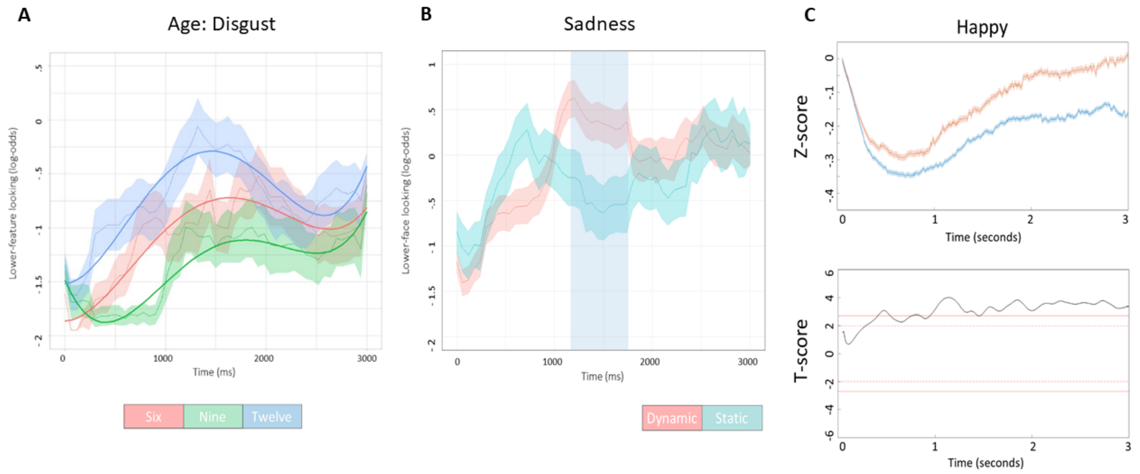

Figure 2. Examples of temporal analysis methods

Collapsing eye tracking and pupillometry data across time can obscure important patterns. Several methods have been used in this thesis to represent changes across time, and a selection of examples are given here. Firstly a depiction of mixed- effects modelling where polynomial growth curves were used to represent the curve of the data across time. In this example (A; Chapter 6, Figure 2) a significant difference was found between nine and twelve-month-old infants for looking toward the lower features of disgust, which negatively interacted with the quadratic time term, indicating a more pronounced ‘inverted U’ shape across time for 12-month-olds. Second cluster permutation analysis was used to identify significant time-windows of divergence between conditions (static vs dynamic). In this example (B; Chapter 5, Figure 6), mean looking to the lower face was not significantly different when collapsed across time, but a significant time window (shaded grey) was identified across time due to a later peak in lower looking within the dynamic condition. Thirdly, functional data analysis was used to analyse pupil size across time. In this example (C; Chapter 7, Figure 4), raw z-scores representing the change in pupil size across time for happy and neutral expressions (top) were converted to b-splines and a functional t-test was performed (Bonferroni corrected) between the expressions (happy minus neutral). The resulting functional t-score curve crosses the threshold for significance (1-3 seconds).

Heatmaps

Whilst collapsing across time can obscure important temporal patterns, collapsing across space (i.e. the stimulus region) or arbitrarily segmenting the stimulus can also obscure findings and distort the data (Caldara & Miellet, 2011; Hessels et al., 2016). Even when AOIs are generated using a data-driven, automatic approach, any segmentation of fixation patterns can increase the risk of error in our analyses. The dynamic AOIs used in the previous analyses contain a pre-defined midpoint for each face to differentiate upper- and lower-face looking. However, differences in face scanning do not always align with such rigid *a-priori* distinctions. Statistical heatmaps can therefore be a useful 'bottom-up' tool to analyse fixation data without collapsing into particular interest regions. Figure 3 (Figure 4 from Chapter 6) presents data from a matrix ANOVA that was performed to investigate differences in pixel-wise face-scanning between ages for a fearful expression trial. The ANOVA map shows that for the fearful trial, older infants show increased looking to diagnostic regions such as the whites of the eyes and open mouth. There was more precise looking toward these regions in the twelve-month age group, with six-month-olds showing significantly greater looking toward a more central facial region (see twelve-month versus six-month post-hoc map). Developmental differences in diagnostic looking toward both eyes and mouth facial features were not identified in any AOI analysis as they cannot be described according to asymmetries in looking along the vertical axis. Nevertheless, these heatmap findings complement previous literature that note a developmental change in fearful face processing across the first year of life (see Leppänen & Nelson, 2009 for a review). Like the temporal analyses, analysing data without collapsing across space can identify patterns that were missed using AOIs, and visualising the data in this way can also help to clarify and interpret AOI findings (e.g. following gaze at a 'lower-spatial trajectory', outside of the pre-defined AOI region; see Chapter 3 and see also Yu, Yurovsky, & Xu, 2012).

Yet similarly to the temporal analyses, heatmap analyses can also identify highly specified regions of significant differences (e.g. developmental differences in happy and sad expressions; see Figures 6 and 7 of Chapter 6). Such small and particular differences between groups within heatmap analyses are likely not to be meaningful, especially considering the variation in face scanning between trials within the same expression (see Chapter 5, Figure 6). It is therefore recommended that researchers consider heatmap findings alongside complementary AOI analyses and across multiple trials before making strong theoretical claims.

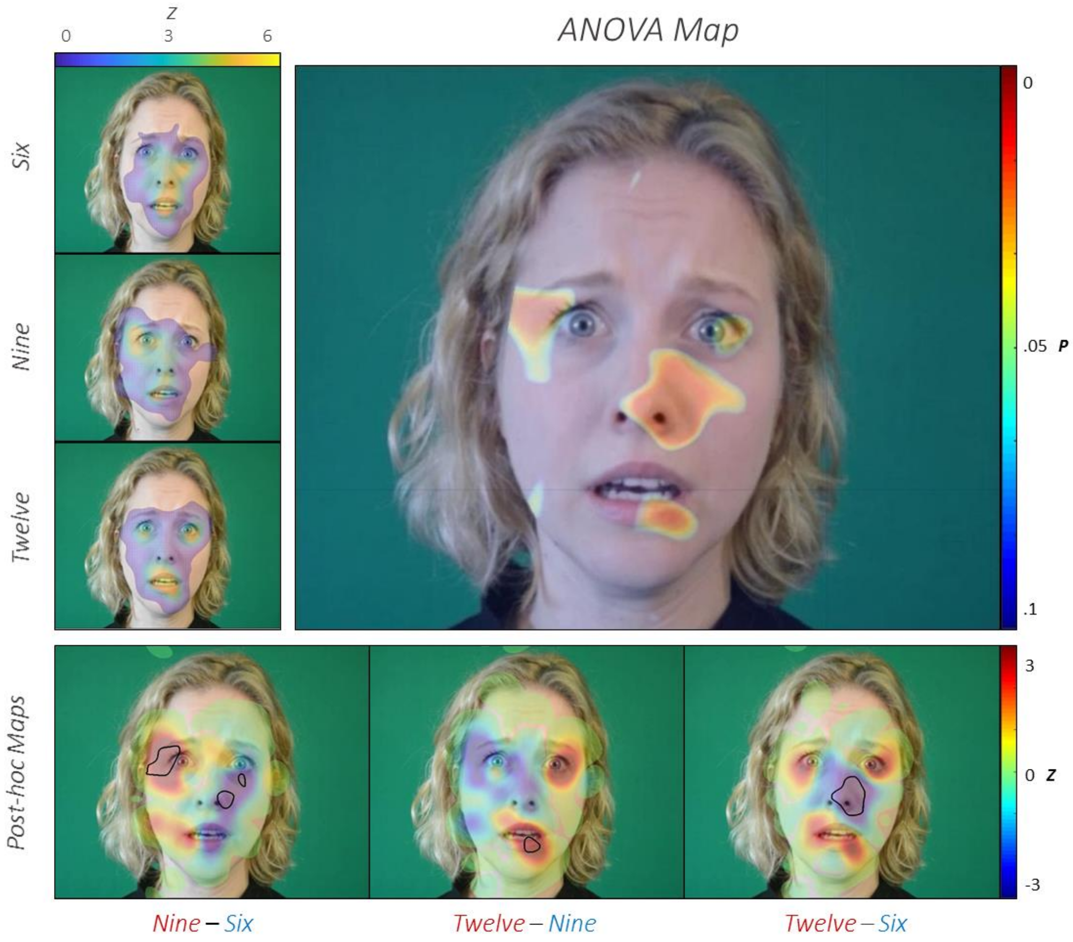

Figure 3. Developmental differences in fearful face scanning

A reproduction of the ANOVA heatmap analysis for a fearful trial. This example was reproduced from Chapter 6, Figure 4.

Implications for future research

The collection of methodological and analytical tools that have been demonstrated within this thesis have the potential to aid researchers in resolving many of the methodological and theoretical debates within infant research. Combining dynamic, naturalistic stimuli and gaze-contingent eye-tracking allows researchers to improve both the internal and external validity of infant experimental paradigms. We can get closer to representing what infants experience in their everyday environments, whilst also providing infants with an *active* behavioural role.

The findings from the empirical work in this thesis illustrate this difference. When infants are actively engaging in a social interaction with an adult, socio-perceptual differences such as race become less salient, and infants' attention is instead driven by the adult's behaviour. This has strong theoretical relevance to the growing field surrounding infant race preferences and race category formation (Quinn et al., 2019), and questions the extent to which perceptual asymmetries within preference tasks are a reliable indicator of infants' *conceptual* understanding of race. Similarly, the chapters focusing on facial expressions show clear differences in scanning for static and dynamic expressions, and also show that infants dynamically target their attention toward facial regions most salient for each expression. As in the earlier chapters, infants scanning patterns imply sophisticated processing of social behaviours. However, a critical contribution of this thesis to this literature is provided in Chapter 7 through the convergent application of pupillometry alongside the eye-tracking study from Chapters 5 and 6. This chapter demonstrates that although infants might show sophisticated *scanning* of facial expressions, they only show a clear physiological response for happiness. It thus emphasises that perceptual ability does not equate to conceptual understanding. This point has relevance for infant eye-tracking in general, and is reminiscent of key theoretical debates regarding the meaningfulness of recording infant looking to assess how infants think, feel or understand.

While the application of gaze-contingent eye-tracking and naturalistic stimuli can provide new information relevant to our theoretical understanding of infant development, it does also provide substantial analytical challenges given the volume and complexity of the data generated. A key contribution of this thesis for future research has been to demonstrate the utility of various analysis tools that allow researchers to preserve vital temporal and spatial patterns intrinsic to social phenomena. Furthermore, the creation of ‘Dynamic AOIs’, which are interest regions generated using information inherent within a video stimulus, should also prove useful for future eye-tracking research.

Conclusions

This thesis began with a historical survey of infant behavioural methods which identified two potential pitfalls that emerge in infancy research which correspond generally to the internal and external validity of our methods. Firstly, the stimuli we present to infants are often insufficient representations of real-world phenomena leading us to question the relevance of our findings. Secondly, recording infant looking as a means to investigate higher-level, cognitive processes has also lead us to question if our ‘rich’ interpretations are really supported by the data. Through a series of empirical examples, this thesis has introduced a new methodological template; the ‘GC Social Interaction paradigm’. This paradigm is designed to reproduce the essential characteristics of naturalistic interactions in a lab setting, and thus offers a substantial improvement in ecological validity compared to the methods currently used to investigate infant socio-cognitive development. And although the data generated from this paradigm are rich and complex, we should not be deterred by the processing challenges that they present. Within the empirical chapters, several analysis tools (e.g. Dynamic AOIs, temporal analyses and heatmaps) were demonstrated which will hopefully aid researchers in handling such datasets. Also, to ensure the accurate and meaningful interpretation of our results, a convergent methodological approach is recommended. Collecting several complementary measures of infant responses (pupillometry is used here, but there are other options including heart-rate or behavioural responses) provides additional constraints which can support and clarify findings. In conclusion, technological advances have the potential to enable researchers to investigate infant socio-cognitive development with greater *rigour* and *relevance* than ever before. Future research should therefore take advantage of these opportunities.

References

Adamson, L. B., & Frick, J. E. (2003). The Still Face: A History of a Shared Experimental Paradigm. *Infancy, 4*(4), 451–473. https://doi.org/10.1207/S15327078IN0404_01

Addabbo, M., Longhi, E., Marchis, I. C., Tagliabue, P., & Turati, C. (2018). Dynamic facial expressions of emotions are discriminated at birth. *PloS One, 13*(3), e0193868. https://doi.org/10.1371/journal.pone.0193868

Adolph, K. E. (2020). Ecological Validity: Mistaking the Lab for Real Life. In R. J. Sternberg (Ed.), *My Biggest Research Mistake: Adventures and Misadventures in Psychological Research* (pp. 187–190). Thousand Oakes, CA: SAGE Publications.

Anzures, G., Wheeler, A., Quinn, P. C., Pascalis, O., Slater, A. M., Heron-Delaney, M., … Lee, K. (2012). Brief daily exposures to Asian females reverses perceptual narrowing for Asian faces in Caucasian infants. *Journal of Experimental Child Psychology, 112*(4), 484–495. https://doi.org/10.1016/j.jecp.2012.04.005

Arizpe, J., Walsh, V., Yovel, G., & Baker, C. I. (2017). The categories, frequencies, and stability of idiosyncratic eye-movement patterns to faces. *Vision Research, 141*, 191–203. https://doi.org/10.1016/j.visres.2016.10.013

Aslin, R. N. (2000). Why take the cog out of infant cognition? *Infancy, 1*(4), 463–470. https://doi.org/10.1207/S15327078IN0104_6

Aslin, R. N. (2007). What's in a look? *Developmental Science, 10*(1), 48–53. https://doi.org/10.1111/j.1467-7687.2007.00563.x

Aslin, R. N. (2012). Infant eyes: A window on cognitive development. *Infancy,* *17*(1), 126–140. https://doi.org/10.1111/j.1532-7078.2011.00097.x

Aslin, R. N., McMurray, B., & Aslin RN. (2004). Automated corneal-reflection eye tracking in infancy: Methodological developments and applications to cognition. *Infancy*, *6*(2), 155–163. https://doi.org/10.1207/s15327078in0602_1

Aslin, R. N., & Salapatek, P. (1975). Saccadic localization of visual targets by the very young human infant. *Perception & Psychophysics*, *17*(3), 293–302. https://doi.org/10.3758/BF03203214

Aslin, R. N., & Smith, L. B. (1988). Perceptual development. *Annual Review of Psychology*, *39*(1), 435–473. https://doi.org/10.1146/annurev.ps.39.020188.002251

Aston-Jones, G., & Cohen, J. D. (2005). An Integrative Theory of Locus Coeruleus-Norepinepherine Function: Adaptive Gain and Optimal Performance. *Annual Review of Neuroscience*, *28*(1), 403–450. https://doi.org/10.1146/annurev.neuro.28.061604.135709

Atkinson, J., Braddick, O., & Braddick, F. (1974). Acuity and contrast sensivity of infant vision. *Nature*, *247*(5440), 403–404. https://doi.org/10.1038/247403a0

Attard-Johnson, J., Bindemann, M., & Ó Ciardha, C. (2016). Pupillary Response as an Age-Specific Measure of Sexual Interest. *Archives of Sexual Behavior*, *45*(4), 855–870. https://doi.org/10.1007/s10508-015-0681-3

Attard-Johnson, J., Ó Ciardha, C., & Bindemann, M. (2019). Comparing methods for the analysis of pupillary response. *Behavior Research Methods*, *51*(1), 83–95. https://doi.org/10.3758/s13428-018-1108-6

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Bahrick, L. E., Gogate, L. J., & Ruiz, I. (2002). Attention and memory for faces and actions in infancy: The salience of actions over faces in dynamic events. *Child Development*, *73*(6), 1629–1643. https://doi.org/10.1111/1467-8624.00495

Baillargeon, R. (1986). Representing the existence and the location of hidden objects: Object permanence in 6-and 8-month-old infants. *Cognition*, *23*(1), 21–41. https://doi.org/10.1016/0010-0277(86)90052-1

Baillargeon, R. (1987). Object permanence in 3 1/2-and 4 1/2-month-old infants. *Developmental Psychology*, *23*(5), 655. https://doi.org/10.1037/0012-1649.23.5.655

Baillargeon, R. (1999). Young infants' expectations about hidden objects: A reply to three challenges. *Developmental Science*, *2*(2), 115–132. https://doi.org/10.1111/1467-7687.00061

Baillargeon, R., Spelke, E. S., & Wasserman, S. (1985). Object permanence in five-month-old infants. *Cognition*, *20*(3), 191–208. https://doi.org/10.1016/0010-0277(85)90008-3

Bar-Haim, Y., Ziv, T., Lamy, D., & Hodes, R. M. (2006). Nature and Nurture in Own- Race Face Processing. *Psychological Science*, *17*(2), 159–163. https://doi.org/10.1111/j.1467-9280.2006.01679.x

Barr, D. J. (2008). Analyzing "visual world" eyetracking data using multilevel logistic regression. *Journal of Memory and Language*, *59*(4), 457–474. https://doi.org/10.1016/j.jml.2007.09.002

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Barr, R., & Hayne, H. (1999). Developmental changes in imitation from television during infancy. *Child Development*, *70*(5), 1067–1081. https://doi.org/10.1111/1467-8624.00079

Barrett, L. F. (2017). *How emotions are made: The secret life of the brain*. New York, NY: Houghton Mifflin Harcourt.

Bates, D., Maechler, M., & Bolker, B. (2014). Fitting Linear Mixed-Effects Models using lme4. *ArXiv E-Prints,* (1406.5823).

Bayet, L., Quinn, P. C., Laboissière, R., Caldara, R., Lee, K., Pascalis, O., … Lee, K. (2017). Fearful but not happy expressions boost face detection in human infants. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1862), 20171054. https://doi.org/10.1098/rspb.2017.1054

Bazhenova, O. V., Stroganova, T. A., Doussard-Roosevelt, J. A., Posikera, I. A., & Porges, S. W. (2007). Physiological responses of 5-month-old infants to smiling and blank faces. *International Journal of Psychophysiology, 63*(1), 64–76. https://doi.org/10.1016/j.ijpsycho.2006.08.008

Beatty, J., & Lucero-Wagoner, B. (2000). The pupillary system. In *Handbook of psychophysiology* (2nd ed., pp. 142–162). Cambridge, UK: Cambridge University Press.

Berkson, G., & Fitz-Gerald, F. L. (1963). Eye fixation aspect of attention to visual stimuli in infant chimpanzees. *Science*, *139*(3555), 586–587. https://doi.org/10.1126/science.139.3555.586

Bertin, E., & Striano, T. (2006). The still-face response in newborn, 1.5-, and 3-month-old infants. *Infant Behavior and Development*, *29*(2), 294–297. https://doi.org/10.1016/j.infbeh.2005.12.003

Biele, C., & Grabowska, A. (2006). Sex differences in perception of emotion intensity in dynamic and static facial expressions. *Experimental Brain*

Research, *171*(1), 1–6. https://doi.org/10.1007/s00221-005-0254-0

Bigelow, A. E., & Birch, S. A. J. (1999). The effects of contingency in previous interactions on infants' preference for social partners. *Infant Behavior and Development*, *22*(3), 367–382. https://doi.org/10.1016/S0163-6383(99)00016-8

Birch, E. E., Gwiazda, J., & Held, R. (1982). Stereoacuity development for crossed and uncrossed disparities in human infants. *Vision Research*, *22*(5), 507–513. https://doi.org/10.1016/0042-6989(82)90108-0

Bogartz, R. S., Shinskey, J. L., & Speaker, C. J. (1997). Interpreting infant looking: The event set X event set design. *Developmental Psychology*, *33*(3), 408. https://doi.org/10.1037/0012-1649.33.3.408

Bornstein, M. H. (1975). Qualities of color vision in infancy. *Journal of Experimental Child Psychology*, *19*(3), 401–419. https://doi.org/10.1016/0022-0965(75)90070-3

Bornstein, M. H., & Arterberry, M. E. (2003). Recognition, discrimination and categorisation of smiling by 5-month-old infants. *Developmental Science*, *5*(5), 585–599. https://doi.org/10.1111/1467-7687.00314

Bower, T. G. R. (1966). The visual world of infants. *Scientific American*. https://doi.org/10.1038/scientificamerican1266-80

Braddick, O., & Atkinson, J. (2011). Development of human visual function. *Vision Research*, *51*(13), 1588–1609. https://doi.org/10.1016/j.visres.2011.02.018

Bradley, M. B., Miccoli, L. M., Escrig, M. a, & Lang, P. J. (2008). The pupil as a measure of emotional arousal and automatic activation (Author Manuscript). *Psychophysiology*, *45*(4), 602. https://doi.org/10.1111/j.1469-8986.2008.00654.x.The

Broadbent, D. E. (1958). *Perception and communication*. Elmsford, NY: Pergamon

Press.

Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist*, *32*(7), 513. https://doi.org/10.1037/0003-066X.32.7.513

Bronfenbrenner, U. (1979). *The ecology of human development*. Cambridge, MA: Harvard University Press.

Brooks, R., & Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental Science*, *8*(6), 535–543. https://doi.org/10.1111/j.1467-7687.2005.00445.x

Brunswik, E. (1943). Organismic achievement and environmental probability. *Psychological Review*, *50*(3), 255. https://doi.org/10.1037/h0060889

Bühler, K. (1930). *The mental development of the child*. New York: Harcourt Brace.

Burge, T. (2018). Do infants and nonhuman animals attribute mental states? *Psychological Review*, *125*(3), 409. https://doi.org/10.1037/rev0000091

Burton, A. M. (2013). Why has research in face recognition progressed so slowly? The importance of variability. *Quarterly Journal of Experimental Psychology*, *66*(8), 1467–1485. https://doi.org/10.1080/17470218.2013.800125

Buswell, G. T. (1935). *How people look at pictures: a study of the psychology and perception in art.* Oxford, England: Univ. Chicago Press.

Caldara, R., & Miellet, S. (2011). iMap: a novel method for statistical fixation mapping of eye movement data. *Behavior Research Methods*, *43*(3), 864–878. https://doi.org/10.3758/s13428-011-0092-x

Calder, A. J., Young, A. W., Keane, J., Dean, M., Young, A. W., & Dean, M. (2000). Configural information in facial expression perception. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(2), 527–

551. https://doi.org/10.1037/0096-1523.26.2.527

Caron, R. F., Caron, A. J., & Myers, R. S. (1985). Do infants see emotional expressions in static faces? *Child Development*, *56*, 1552–1560. https://doi.org/10.2307/1130474

Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, *59*, i–174. https://doi.org/10.2307/1166214

Cashon, C. H., & Cohen, L. B. (2000). Eight-month-old infants' perception of possible and impossible events. *Infancy*, *1*(4), 429–446. https://doi.org/10.1207/S15327078IN0104_4

Cesana-Arlotti, N., Martin, A., Téglás, E., Vorobyova, L., Cetnarski, R., & Bonatti, L. L. (2018). Precursors of logical reasoning in preverbal human infants. *Science*, *359*(6381), 1263–1266. https://doi.org/10.1126/science.aao3539

Chawarska, K., & Shic, F. (2009). Looking but not seeing: Atypical visual scanning and recognition of faces in 2 and 4-Year-old children with Autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *39*(12), 1663–1672. https://doi.org/10.1007/s10803-009-0803-7

Cohen, L. B. (2004). Uses and misuses of habituation and related preference paradigms. *Infant and Child Development: An International Journal of Research and Practice*, *13*(4), 349–352. https://doi.org/10.1002/icd.355

Colombo, J., & Mitchell, D. W. (2009). Infant visual habituation. *Neurobiology of Learning and Memory*, *92*(2), 225–234. https://doi.org/10.1016/j.nlm.2008.06.002

Cooper, R. P., & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development*, *61*(5), 1584–1595. https://doi.org/10.1111/j.1467-8624.1990.tb02885.x

Cosmides, L., Tooby, J., & Kurzban, R. (2003). Perceptions of race. *Trends in Cognitive Sciences*, *7*(4), 173–179. https://doi.org/10.1016/S1364-6613(03)00057-3

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, *13*(4), 148–153. https://doi.org/10.1016/j.tics.2009.01.005

D'Entremont, B., Hains, S. M. J., & Muir, D. W. (1997). A demonstration of gaze following in 3- to 6-month-olds. *Infant Behav Dev*, *20*(4), 569–572. https://doi.org/10.1016/S0163-6383(97)90048-5

Darwin, C. R. (1872). *The expression of emotion in animals and man* (Vol. 2). London: John Murray.

Darwin, C. R. (1877). A biographical sketch of an infant. *Mind*, *2*(7), 285–294.

Deligianni, F., Senju, A., Gergely, G., & Csibra, G. (2011). Automated gaze-contingent objects elicit orientation following in 8-month-old infants. *Developmental Psychology*, *47*(6), 1499–1503. https://doi.org/10.1037/a0025659

Deutsch, J. A., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological Review*, *70*(1), 80. https://doi.org/10.1037/h0039515

Dink, J., & Ferguson, B. (2015). *eyetrackingR:* Retrieved from http://www.eyetrackingr.com

Dobson, V., & Teller, D. Y. (1978). Visual acuity in human infants: a review and comparison of behavioral and electrophysiological studies. *Vision Research*, *18*(11), 1469–1483.

Duchowski, A. T. (2017). Eye tracking methodology. In *Theory and practice* (3rd ed.). Gewerbestrasse, Switzerland: Springer.

Duchowski, A. T., Cournia, N., & Murphy, H. (2004). Gaze-contingent displays: A review. *CyberPsychology & Behavior*, *7*(6), 621–634. https://doi.org/10.1089/cpb.2004.7.621

Dunbar, R. I. M., & Shultz, S. (2007). Evolution in the social brain. *Science*, *317*(5843), 1344–1347. https://doi.org/10.1126/science.1145463

Dunn, J., & Kendrick, C. (1980). The arrival of a sibling: changes in patterns of interaction between mother and first-born child. *Journal of Child Psychology and Psychiatry*, *21*(2), 119–132. https://doi.org/10.1111/j.1469-7610.1980.tb00024.x

Eckstein, M. K., Guerra-Carrillo, B., Miller Singley, A. T., & Bunge, S. A. (2017). Beyond eye gaze: What else can eyetracking reveal about cognition and cognitive development? *Developmental Cognitive Neuroscience*, *25*, 69–91. https://doi.org/10.1016/j.dcn.2016.11.001

Ekman, P. (1993). Facial expression and emotion. *The American Psychologist*, *48*(4), 384–392. https://doi.org/10.1037/0003-066X.48.4.384

Ekman, P., & Friesen, W. V. (1976). *Pictures of facial affect*. Palo Alto, CA: Consulting Psychologists Press.

Ekman, P., Friesen, W. V., O'Sullivan, M., Chan, A., Diacoyanni-Tarlatzis, I., Heider, K., … Tomita, M. (1987). Universals and cultural differences in the judgments of facial expressions of emotion. *Journal of Personality and Social Psychology*, *53*(4), 712–717. https://doi.org/10.1037/0022-3514.53.4.712

Elsabbagh, M., Gliga, T., Pickles, A., Hudry, K., Charman, T., & Johnson, M. H. (2013). The development of face orienting mechanisms in infants at-risk for autism. *Behavioural Brain Research*, *251*, 147–154.

https://doi.org/10.1016/j.bbr.2012.07.030

Falck-Ytter, T., Gredebäck, G., & von Hofsten, C. (2006). Infants predict other people’s action goals. *Nature Neuroscience*, *9*(7), 878. https://doi.org/10.1038/nn1729

Fantz, R. L. (1956). A method for studying early visual development. *Perceptual and Motor Skills*, *6*(1), 13–15. https://doi.org/10.2466/pms.1956.6.g.13

Fantz, R. L. (1961). The origin of form perception. *Scientific American*, *204*(5), 66–73. https://doi.org/10.1038/scientificamerican0561-66

Fantz, R. L. (1963). Pattern vision in newborn infants. *Science*, *140*(3564), 296–297. https://doi.org/10.1126/science.140.3564.296

Fantz, R. L. (1964). Visual experience in infants: Decreased attention to familiar patterns relative to novel ones. *Science*, *146*(3644), 668–670. https://doi.org/10.1126/science.146.3644.668

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences*, *99*(14), 9602–9605. https://doi.org/10.1073/pnas.152159999

Farroni, T., Massaccesi, S., Pividori, D., & Johnson, M. H. (2004). Gaze following in newborns. *Infancy*, *5*(1), 39–60. https://doi.org/10.1207/s15327078in0501_2

Farroni, T., Menon, E., & Johnson, M. H. (2006). Factors influencing newborns’ preference for faces with eye contact. *Journal of Experimental Child Psychology*, *95*(4), 298–308. https://doi.org/10.1016/j.jecp.2006.08.001

Farroni, T., Menon, E., Rigato, S., & Johnson, M. H. (2007). The perception of facial expressions in newborns. *European Journal of Developmental Psychology*, *4*(1), 2–13. https://doi.org/10.1080/17405620601046832

Fassbender, I., Teubert, M., & Lohaus, A. (2016). The development of preferences for own-race versus other-race faces in 3-, 6-and 9-month-old Caucasian infants. *European Journal of Developmental Psychology*, *13*(1), 152–165. https://doi.org/10.1080/17405629.2015.1073585

Feng, G. (2011). Eye Tracking: A brief guide for developmental researchers. *Journal of Cognition and Development*, *12*(1), 1–11. https://doi.org/10.1080/15248372.2011.547447

Field, T. M., Cohen, D., Garcia, R., & Greenberg, R. (1984). Mother-stranger face discrimination by the newborn. *Infant Behavior and Development*, *7*(1), 19–25. https://doi.org/10.1016/S0163-6383(84)80019-3

Field, T. M., Goldstein, S., Vega-Lahr, N., & Porter, K. (1986). Changes in imitative behavior during early infancy. *Infant Behavior and Development*, *9*(4), 415–421. https://doi.org/10.1016/0163-6383(86)90015-9

Field, T. M., Woodson, R., Greenberg, R., & Cohen, D. (1982). Discrimination and imitation of facial expressions by neonates. *Science*, *218*(4568), 179–181. https://doi.org/10.1126/science.7123230

Fitzgerald, E. (1965). Autonomic Pupillary Reflex Activity during Early Infancy and Its Relation to Social and Nonsocial Visual. *Journal of Experimental Child Psychology*, *6*(3), 470–482. https://doi.org/10.1016/0022-0965(68)90127-6

Flom, R., & Pick, A. D. (2005). Experimenter affective expression and gaze following in 7-month-olds. *Infancy*, *7*(December), 207–218. https://doi.org/10.1207/s15327078in0702_5

Fox, E., & Damjanovic, L. (2006). The eyes are sufficient to produce a threat superiority effect. *Emotion*, *6*(3), 534–539. https://doi.org/10.1037/1528-3542.6.3.534

Fox, R., & McDaniel, C. (1982). The perception of biological motion by human infants. *Science*, *218*(4571), 486–487. https://doi.org/10.1126/science.7123249

Franchak, J. M., Kretch, K. S., Soska, K. C., & Adolph, K. E. (2011). Head-mounted eye tracking: A new method to describe infant looking. *Child Development*, *82*(6), 1738–1750. https://doi.org/10.1111/j.1467-8624.2011.01670.x

Franchak, J. M., Kretch, K. S., Soska, K. C., Babcock, J. S., & Adolph, K. E. (2010). Head-mounted eye-tracking of infants' natural interactions: a new method. *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications*, 21–27. https://doi.org/10.1145/1743666.1743671

Francois, J., Coufal, K., & Chaparro, B. (2018). The Use of Eye-Tracking in the Investigation of Prelinguistic Infants: A Review. *Ann Cogn Sci*, *2*(1), 69–77.

Friedman, S. (1972). Habituation and recovery of visual response in the alert human newborn. *Journal of Experimental Child Psychology*, *13*(2), 339–349. https://doi.org/10.1016/0022-0965(72)90095-1

Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *358*(1431), 459–473. https://doi.org/10.1098/rstb.2002.1218

Gaither, S. E., Pauker, K., & Johnson, S. P. (2012). Biracial and monoracial infant own-race face perception: an eye tracking study. *Developmental Science*, *15*(6), 775–782. https://doi.org/10.1111/j.1467-7687.2012.01170.x

Geangu, E., Hauf, P., Bhardwaj, R., & Bentz, W. (2011). Infant pupil diameter changes in response to others' positive and negative emotions. *PLoS ONE*, *6*(11), e27132. https://doi.org/10.1371/journal.pone.0027132

Geangu, E., Ichikawa, H., Lao, J., Kanazawa, S., Yamaguchi, M. K., Caldara, R., & Turati, C. (2016). Culture shapes 7-month-olds' perceptual strategies in discriminating facial expressions of emotion. *Current Biology*, *26*(14), R663–

R664. https://doi.org/10.1016/j.cub.2016.05.072

Gergely, G., Knadasdy, Z., Csibra, G., & Biro, S. (1995). Taking the International Stance at 12 Months of Age. *Cognition*, *56*(95), 165–193. https://doi.org/10.1016/0010-0277(95)00661-H

Gibbs, J. C. (1979). The meaning of ecologically oriented inquiry in contemporary psychology. *American Psychologist*, *34*(2), 127. https://doi.org/10.1037/0003-066X.34.2.127

Gibson, E. J. (1969). *Principles of perceptual learning and development.* East Norwalk, CT: Appleton-Century-Crofts.

Gibson, E. J., & Walk, R. D. (1960). The "visual cliff." *Scientific American*, *202*, 64–71. https://doi.org/10.1038/scientificamerican0460-64

Gibson, J. J. (1966). *The senses considered as perceptual systems.* Oxford, England: Houghton Mifflin.

Gibson, J. J. (1979). *The ecological approach to visual perception.* Boston, MA, US: Houghton, Mifflin and Company.

Godard, O., Baudouin, J.-Y. Y., Schaal, B., & Durand, K. (2016). Affective matching of odors and facial expressions in infants: Shifting patterns between 3 and 7 months. *Developmental Science*, *19*(1), 155–163. https://doi.org/10.1111/desc.12292

Goren, C. C., Sarty, M., & Wu, P. Y. K. (1975). Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics*, *56*(4), 544–549.

Gosselin, F., & Schyns, P. G. (2001). Bubbles: a technique to reveal the use of information in recognition tasks. *Vision Research*, *41*(17), 2261–2271. https://doi.org/10.1016/S0042-6989(01)00097-9

Gratier, M., Devouche, E., Guellai, B., Infanti, R., Yilmaz, E., & Parlato-Oliveira, E. (2015). Early development of turn-taking in vocal interaction between mothers and infants. *Front. Psychol*, *6*(1167), 10–3389. https://doi.org/10.3389/fpsyg.2015.01167

Gredebäck, G., Eriksson, M., Schmitow, C., Laeng, B., & Stenberg, G. (2012). Individual differences in face processing: Infants' scanning patterns and pupil dilations are influenced by the distribution of parental leave. *Infancy*, *17*(1), 79–101. https://doi.org/10.1111/j.1532-7078.2011.00091.x

Gredebäck, G., Fikke, L., & Melinder, A. (2010). The development of joint visual attention: a longitudinal study of gaze following during interactions with mothers and strangers. *Developmental Science*, *13*(6), 839–848. https://doi.org/10.1111/j.1467-7687.2009.00945.x

Gredebäck, G., Johnson, S. P., & Von Hofsten, C. (2010). Eye tracking in infancy research. *Developmental Neuropsychology*, *35*(1), 1–19. https://doi.org/10.1080/87565640903325758

Gredebäck, G., & Melinder, A. (2010). Infants' understanding of everyday social interactions: A dual process account. *Cognition*, *114*(2), 197–206. https://doi.org/10.1016/j.cognition.2009.09.004

Grossmann, T., & Johnson, M. H. (2007). The development of the social brain in human infancy. *European Journal of Neuroscience*, *25*(January), 909–919. https://doi.org/10.1111/j.1460-9568.2007.05379.x

Hainline, L. (1978). Developmental changes in visual scanning of face and nonface patterns by infants. *Journal of Experimental Child Psychology*, *25*(1), 90–115. https://doi.org/10.1016/0022-0965(78)90041-3

Hains, S. M. J., & Muir, D. W. (1996). Infant sensitivity to adult eye direction. *Child*

Development, *67*(5), 1940–1951. https://doi.org/10.1111/j.1467-8624.1996.tb01836.x

Haith, M. M. (1969). Infrared television recording and measurement of ocular behavior in the human infant. *American Psychologist*, *24*(3), 279. https://doi.org/10.1037/h0028419

Haith, M. M. (1998). Who put the cog in infant cognition? Is rich interpretation too costly? *Infant Behavior and Development*, *21*(2), 167–179. https://doi.org/10.1016/S0163-6383(98)90001-7

Haith, M. M., Bergman, T., & Moore, M. J. (1977). Eye Contact and Face Scanning in Early Infancy. *Science*, *198*(4319), 853–855. https://doi.org/10.1126/science.918670

Hanawalt, N. G. (1944). The role of the upper and the lower parts of the face as a basis for judging facial expressions: II. In posed expressions and “candid-camera” pictures. *The Journal of General Psychology, 31*(1), 23–36. https://doi.org/10.1080/00221309.1944.10545217

Hari, R., & Kujala, M. V. (2009). Brain basis of human social interaction: from concepts to brain imaging. *Physiological Reviews*, *89*(2), 453–479. https://doi.org/10.1152/physrev.00041.2007

Haviland, J. M., & Lelwica, M. (1987). The induced affect response: 10-week-old infants’ responses to three emotion expressions. *Developmental Psychology*, *23*(1), 97–104. https://doi.org/10.1037/0012-1649.23.1.97

Hayes, T. R., & Petrov, A. A. (2016). Mapping and correcting the influence of gaze position on pupil size measurements. *Behavior Research Methods*, *48*(2), 510–527. https://doi.org/10.3758/s13428-015-0588-x

Hayne, H., Herbert, J., & Simcock, G. (2003). Imitation from television by 24-and

30-month-olds. *Developmental Science, 6*(3), 254–261. https://doi.org/10.1111/1467-7687.00281

Heck, A., Hock, A., White, H., Jubran, R., & Bhatt, R. S. (2016). The development of attention to dynamic facial emotions. *Journal of Experimental Child Psychology*, *147*, 100–110. https://doi.org/10.1016/j.jecp.2016.03.005

Heck, A., Hock, A., White, H., Jubran, R., & Bhatt, R. S. (2017). Further evidence of early development of attention to dynamic facial emotions: Reply to Grossmann and Jessen. *Journal of Experimental Child Psychology*, *153*, 155–162. https://doi.org/10.1016/j.jecp.2016.08.006

Henderson, R. R., Bradley, M. M., & Lang, P. J. (2014). Modulation of the initial light reflex during affective picture viewing. *Psychophysiology*, *51*(9), 815–818. https://doi.org/10.1111/psyp.12236

Hepach, R., Vaish, A., & Tomasello, M. (2015). Novel paradigms to measure variability of behavior in early childhood: posture, gaze, and pupil dilation. *Frontiers in Psychology*, *6*(July), 1–12. https://doi.org/10.3389/fpsyg.2015.00858

Hepach, R., & Westermann, G. (2013). Infants’ sensitivity to the congruence of others’ emotions and actions. *Journal of Experimental Child Psychology*, *115*(1), 16–29. https://doi.org/10.1016/j.jecp.2012.12.013

Hepach, R., & Westermann, G. (2016). Pupillometry in Infancy Research. *Journal of Cognition and Development*, *17*(3), 359–377. https://doi.org/10.1080/15248372.2015.1135801

Heron-Delaney, M., Anzures, G., Herbert, J. S., Quinn, P. C., Slater, A. M., Tanaka, J. W., … Pascalis, O. (2011). Perceptual training prevents the emergence of the other race effect during infancy. *PLoS ONE*, *6*(5), e19858.

https://doi.org/10.1371/journal.pone.0019858

Herrnstein, R. J. (1977). Doing what comes naturally: A reply to Professor Skinner. *American Psychologist*, *32*(12), 1013–1016. https://doi.org/10.1037/0003-066X.32.12.1013

Hess, E. H. (1972). Pupillometrics: A method of studying mental, emotional and sensory processes. In *Handbook of Psychophysiology* (pp. 491–531). New York, NY: Holt, Rinehart & Winston.

Hess, E. H., & Polt, J. M. (1960). Pupil size as related to interest value of visual stimuli. *Science (New York, N.Y.)*, *132*(3423), 349–350. https://doi.org/10.1126/science.132.3423.349

Hess, E. H., & Polt, J. M. (1964). Pupil Size in Relation to Mental Activity during Simple Problem-Solving. *Science*, *143*(3611), 1190–1192. https://doi.org/10.1126/science.143.3611.1190

Hessels, R. S., Benjamins, J. S., Cornelissen, T. H. W., & Hooge, I. T. C. (2018). A validation of automatically-generated Areas-of-Interest in videos of a face for eye-tracking research. *Frontiers in Psychology*, *9*, 1367. https://doi.org/10.3389/fpsyg.2018.01367

Hessels, R. S., Kemner, C., van den Boomen, C., & Hooge, I. T. C. (2016). The area-of-interest problem in eyetracking research: A noise-robust solution for face and sparse stimuli. *Behavior Research Methods*, *48*(4), 1694–1712. https://doi.org/10.3758/s13428-015-0676-y

Heyes, C. M. (2014). False belief in infancy: a fresh look. *Developmental Science*, *17*(5), 647–659. https://doi.org/10.1111/desc.12148

Hills, P. J., & Pake, J. M. (2013). Eye-tracking the own-race bias in face recognition: Revealing the perceptual and socio-cognitive mechanisms. *Cognition*, *129*(3), 586–597. https://doi.org/10.1016/j.cognition.2013.08.012

Hoehl, S., & Markova, G. (2018). Moving developmental social neuroscience toward a second-person approach. *PLoS Biology, 16*(12), e3000055. https://doi.org/10.1371/journal.pbio.3000055

Holmqvist, K., Nyström, M., & Mulvey, F. (2012). Eye tracker data quality: what it is and how to measure it. *Proceedings of the Symposium on Eye Tracking Research and Applications*, 45–52. https://doi.org/10.1145/2168556.2168563

Hood, B. M., Willen, J. D., Driver, J., Hood, B. M., & Willen, J. D. (1998). Adult's eyes trigger shifts of visual attention in human infants. *Psychological Science, 9*(2), 131–134. https://doi.org/10.1111/1467-9280.00024

Houston-Price, C., & Nakai, S. (2004). Distinguishing novelty and familiarity effects in infant preference procedures. *Infant and Child Development: An International Journal of Research and Practice*, *13*(4), 341–348. https://doi.org/10.1002/icd.364

Hugenberg, K., Young, S. G., Bernstein, M. J., & Sacco, D. F. (2010). The categorization-individuation model: an integrative account of the other-race recognition deficit. *Psychological Review*, *117*(4), 1168–1187. https://doi.org/10.1037/a0020463

Hunnius, S., & Geuze, R. H. (2004). Developmental changes in visual scanning of dynamic faces and abstract stimuli in infants: A longitudinal study. *Infancy*, *6*(2), 231–255. https://doi.org/10.1207/s15327078in0602_5

Hunter, M. A., & Ames, E. W. (1988). A multifactor model of infant preferences for novel and familiar stimuli. In C. Rovee-Collier & L. P. Lipsitt (Eds.), *Advances in infancy research* (pp. 69–95). Westport, CT: Ablex Publishing.

Ichikawa, H., Kanazawa, S., & Yamaguchi, M. K. (2011). The movement of internal facial features elicits 7 to 8-month-old infants' preference for face patterns. *Infant and Child Development*, *20*(5), 464–474. https://doi.org/10.1002/icd.724

Itier, R. J., & Batty, M. (2009). Neural basis of eye and gaze processing: The core of social cognition. *Neurosci Biobehav Rev*, *33*(6), 843–863. https://doi.org/10.1016/j.neubiorev.2009.02.004

Itier, R. J., Villate, C., & Ryan, J. D. (2007). Eyes always attract attention but gaze orienting is task-dependent: Evidence from eye movement monitoring. *Neuropsychologia*, *45*(5), 1019–1028. https://doi.org/10.1016/j.neuropsychologia.2006.09.004

Jack, R. E., Garrod, O. G. B. B., & Schyns, P. G. (2014). Dynamic facial expressions of emotion transmit an evolving hierarchy of signals over time. *Current Biology*, *24*(2), 187–192. https://doi.org/10.1016/j.cub.2013.11.064

Jack, R. E., Sun, W., Delis, I., Garrod, O. G. B., & Schyns, P. G. (2016). Four not six: Revealing culturally common facial expressions of emotion. *Journal of Experimental Psychology: General*, *145*(6), 708–730. https://doi.org/10.1037/xge0000162

Jackson, I. R., & Sirois, S. (2009). Infant cognition: Going full factorial with pupil dilation. *Developmental Science*, *12*(4), 670–679. https://doi.org/10.1111/j.1467-7687.2008.00805.x

James, W., Burkhardt, F., Bowers, F., & Skrupskelis, I. K. (1890). *The principles of psychology* (Vol. 1). London: Macmillan.

Jenkins, J. J. (1974). Remember that old theory of memory? Well, forget it. *American Psychologist*, *29*(11), 785. https://doi.org/10.1037/h0037399

Jessen, S., Altvater-Mackensen, N., & Grossmann, T. (2016). Pupillary responses reveal infants' discrimination of facial emotions independent of conscious perception. *Cognition*, *150*, 163–169. https://doi.org/10.1016/j.cognition.2016.02.010

Jessen, S., & Grossmann, T. (2014). Unconscious discrimination of social cues from eye whites in infants. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(45), 16208–16213. https://doi.org/10.1073/pnas.1411333111

Jessen, S., & Grossmann, T. (2016). The developmental emergence of unconscious fear processing from eyes during infancy. *Journal of Experimental Child Psychology*, *142*, 334–343. https://doi.org/10.1016/j.jecp.2015.09.009

Johnson, S. P., Amso, D., & Slemmer, J. A. (2003). Development of object concepts in infancy: Evidence for early learning in an eye-tracking paradigm. *Proceedings of the National Academy of Sciences*, *100*(18), 10568–10573. https://doi.org/10.1073/pnas.1630655100

Johnson, S. P., & Aslin, R. N. (1995). Perception of object unity in 2-month-old infants. *Developmental Psychology*, *31*(5), 739. https://doi.org/10.1037/0012-1649.31.5.739

Jones, W., & Klin, A. (2013). Attention to eyes is present but in decline in 2-6-month-old infants later diagnosed with autism. *Nature*, *504*(7480), 427–431. https://doi.org/10.1038/nature12715

Joshi, S., Li, Y., Kalwani, R. M., & Gold, J. I. (2016). Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate Cortex. *Neuron*, *89*(1), 1–14. https://doi.org/10.1016/j.neuron.2015.11.028

Kahneman, D., & Beatty, J. (1966). Pupil Diameter and Load on Memory. *Science*, *154*(3756), 1583–1585. https://doi.org/10.1126/science.154.3756.1583

Kanan, C., Bseiso, D. N. F., Ray, N. A., Hsiao, J. H., & Cottrell, G. W. (2015). Humans have idiosyncratic and task-specific scanpaths for judging faces. *Vision Research*, *108*, 67–76. https://doi.org/10.1016/j.visres.2015.01.013

Kaye, H. (1967). Infant sucking behavior and its modification. In *Advances in child development and behavior* (Vol. 3, pp. 1–52). Elsevier.

Keemink, J. R., Keshavarzi-Pour, M. J., & Kelly, D. J. (2019). Infants' responses to interactive gaze-contingent faces in a novel and naturalistic eye-tracking paradigm. *Developmental Psychology*, *55*(7), 1362–1371. https://doi.org/10.1037/dev0000736

Keemink, J. R., Prunty, J. E., Wood, N., & Kelly, D. J. (2019). Eye Movements and Behavioural Responses to Interactive Gaze-Contingent Faces. *Poster Presented at the Lancaster Conference on Infant and Early Child Development*. Lancaster, UK.

Kellman, P. J., & Spelke, E. S. (1983). Perception of partly occluded objects in infancy. *Cognitive Psychology*, *15*(4), 483–524. https://doi.org/10.1016/0010-0285(83)90017-8

Kelly, D. J., Duarte, S., Meary, D., Bindemann, M., & Pascalis, O. (2019). Infants rapidly detect human faces in complex naturalistic visual scenes. *Developmental Science*, e12829. https://doi.org/10.1111/desc.12829

Kelly, D. J., Liu, S., Lee, K., Quinn, P. C., Pascalis, O., Slater, A. M., & Ge, L. (2009). Development of the other-race effect during infancy: Evidence toward universality? *Journal of Experimental Child Psychology*, *104*(1), 105–114. https://doi.org/10.1016/j.jecp.2009.01.006

Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Ge, L., & Pascalis, O. (2007). The other-race effect develops during infancy: Evidence of perceptual narrowing.

Psychological Science, *18*(12), 1084–1089. https://doi.org/10.1111/j.1467-9280.2007.02029.x

Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Gibson, A., Smith, M., … Ge, L. (2005). Three-month-olds, but not newborns, prefer own-race faces. *Developmental Science*, *8*(6), 31–36. https://doi.org/10.1111/j.1467-7687.2005.0434a.x

Kestenbaum, R., & Nelson, C. A. (1990). The recognition and categorization of upright and inverted emotional expressions by 7-month-old infants. *Infant Behavior and Development*, *13*(4), 497–511. https://doi.org/10.1016/0163-6383(90)90019-5

Kilts, C. D., Egan, G., Gideon, D. A., Ely, T. D., & Hoffman, J. M. (2003). Dissociable neural pathways are involved in the recognition of emotion in static and dynamic facial expressions. *Neuroimage*, *18*(1), 156–168. https://doi.org/10.1006/nimg.2002.1323

Kinzler, K. D., Shutts, K., DeJesus, J., & Spelke, E. S. (2009). Accent trumps race in guiding children’s social preferences. *Social Cognition*, *27*(4), 623–634. https://doi.org/10.1521/soco.2009.27.4.623

Kinzler, K. D., & Spelke, E. S. (2011). Do infants show social preferences for people differing in race? *Cognition*, *119*(1), 1–9. https://doi.org/10.1016/j.cognition.2010.10.019

Kochukhova, O., & Gredebäck, G. (2010). Preverbal infants anticipate that food will be brought to the mouth: An eye tracking study of manual feeding and flying spoons. *Child Development*, *81*(6), 1729–1738. https://doi.org/10.1111/j.1467-8624.2010.01506.x

Kolkur, S., Kalbande, D., Shimpi, P., Bapat, C., & Jatakia, J. (2017). Human skin detection using RGB, HSV and YCbCr color models. *ArXiv E-Prints*, (1708.02694).

Kotsoni, E., Haan, M. D. E., & Johnson, M. H. (2001). Categorical perception of facial expressions by 7-month-old infants. *Perception*, *30*(9), 1115–1125. https://doi.org/10.1068/p3155

Kowler, E. (2011). Eye movements: the past 25 years. *Vision Research*, *51*(13), 1457–1483. https://doi.org/10.1016/j.visres.2010.12.014

Kret, M. E., Roelofs, K., Stekelenburg, J. J., & de Gelder, B. (2013). Emotional signals from faces, bodies and scenes influence observers' face expressions, fixations and pupil-size. *Frontiers in Human Neuroscience*, *7*, 810. https://doi.org/10.3389/fnhum.2013.00810

Kretch, K. S., & Adolph, K. E. (2015). Active vision in passive locomotion: Real-world free viewing in infants and adults. *Developmental Science*, *18*(5), 736–750. https://doi.org/10.1111/desc.12251

Kretch, K. S., Franchak, J. M., & Adolph, K. E. (2014). Crawling and walking infants see the world differently. *Child Development*, *85*(4), 1503–1518. https://doi.org/10.1111/cdev.12206

Krumhuber, E. G., Kappas, A., & Manstead, A. S. R. (2013). Effects of dynamic aspects of facial expressions: A review. *Emotion Review*, *5*(1), 41–46. https://doi.org/10.1177/1754073912451349

Kuchuk, A., Vibbert, M., & Bornstein, M. H. (1986). The perception of smiling and its experiential correlates in three-month-old infants. *Child Development*, *57*(4), 1054–1061. https://doi.org/http://dx.doi.org/10.2307/1130379

Kuhn, D. (1978). Mechanisms of cognitive and social development: One psychology or two? *Human Development*, *21*(2), 92–118.

https://doi.org/10.1159/000271577

Kurzban, R., Tooby, J., & Cosmides, L. (2001). Can race be erased? Coalitional computation and social categorization. *Proceedings of the National Academy of Sciences*, *98*(26), 15387–15392. https://doi.org/10.1073/pnas.251541498

La Barbera, J. D., Izard, C. E., Vietze, P., & Parisi, S. A. (1976). Four- and six-month-old infants' visual responses to joy, anger, and neutral expressions. *Child Development*, *47,* 535–538. https://doi.org/10.2307/1128816

Labar, K. S., Crupain, M. J., Voyvodic, J. T., & McCarthy, G. (2003). Dynamic perception of facial affect and facial identity in the human brain. *Cerebral Cortex*, *13*(10), 1023–1033. https://doi.org/10.1093/cercor/13.10.1023

Laeng, B., Sæther, L., Holmlund, T., Wang, C. E. A., Waterloo, K., Eisemann, M., & Halvorsen, M. (2013). Invisible emotional expressions influence social judgments and pupillary responses of both depressed and non-depressed individuals. *Frontiers in Psychology, 4,* 291. https://doi.org/10.3389/fpsyg.2013.00291

Laeng, B., Sirois, S., & Gredebäck, G. (2012). Pupillometry: A window to the preconscious? *Perspectives on Psychological Science*, *7*(1), 18–27. https://doi.org/10.1177/1745691611427305

Legerstee, M., Markova, G., & Fisher, T. (2007). The role of maternal affect attunement in dyadic and triadic communication. *Infant Behavior and Development*, *30*(2), 296–306. https://doi.org/10.1016/j.infbeh.2006.10.003

Leigh, R. J., & Zee, D. S. (2006). *The neurology of eye movements* (4th ed.). Oxford: Oxford University Press.

Leppänen, J. M., & Nelson, C. A. (2009). Tuning the developing brain to social signals of emotions. *Nature Reviews Neuroscience*, *10*(1), 37–47.

https://doi.org/10.1038/nrn2554

Leslie, A. M. (1982). The perception of causality in infants. *Perception*, *11*(2), 173–186. https://doi.org/10.1068/p110173

Leslie, A. M. (1984). Spatiotemporal continuity and the perception of causality in infants. *Perception*, *13*(3), 287–305. https://doi.org/10.1068/p130287

Lewin, K. (1943). Defining the “field at a given time.” *Psychological Review*, *50*(3), 292. https://doi.org/10.1037/h0062738

Lewis, M. B., Kagan, J., Campbell, H., & Kalafat, J. (1966). The cardiac response as a correlate of attention in infants. *Child Development*, 63–71. https://doi.org/10.2307/1126429

Lewkowicz, D. J. (2001). The Concept of Ecological Validity: What Are Its Limitations and Is It Bad to Be Invalid? *Infancy*, *2*(4), 437–450. https://doi.org/10.1207/S15327078IN0204_03

Lewkowicz, D. J., & Hansen-Tift, A. M. (2012). Infants deploy selective attention to the mouth of a talking face when learning speech. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(5), 1431–1436. https://doi.org/10.1073/pnas.1114783109

Lickliter, R., & Bahrick, L. E. (2001). The salience of multimodal sensory stimulation in early development: Implications for the issue of ecological validity. *Infancy*, *2*(4), 451–467. https://doi.org/10.1207/S15327078IN0204_04

Liu, S., Quinn, P. C., Wheeler, A., Xiao, N. G., Ge, L., & Lee, K. (2011). Similarity and difference in the processing of same- and other-race faces as revealed by eye tracking in 4- to 9-month-olds. *Journal of Experimental Child Psychology*, *108*(1), 180–189. https://doi.org/10.1016/J.JECP.2010.06.008

Liu, S., Xiao, W. S., Xiao, N. G., Quinn, P. C., Zhang, Y., Chen, H., … Lee, K.

(2015). Development of visual preference for own-versus other-race faces in infancy. *Developmental Psychology*, *51*(4), 500–511. https://doi.org/10.1037/a0038835

Loewenfeld, I. E., & Lowenstein, O. (1993). *The pupil: Anatomy, physiology, and clinical applications* (Vol. 2). Wiley-Blackwell.

Ludemann, P. M., & Nelson, C. A. (1988). Categorical representation of facial expressions by 7-month-old infants. *Developmental Psychology*, *24*(4), 492. https://doi.org/10.1037/0012-1649.24.4.492

Macchi Cassia, V., Bulf, H., Quadrelli, E., & Proietti, V. (2014). Age-related face processing bias in infancy: Evidence of perceptual narrowing for adult faces. *Developmental Psychobiology*, *56*(2), 238–248. https://doi.org/10.1002/dev.21191

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024

Maurer, D., & Salapatek, P. (1976). Developmental changes in the scanning of faces by young infants. *Child Development*, 523–527. https://doi.org/10.2307/1128813

McCall, R. B. (1977). Challenges to a science of developmental psychology. *Child Development*, *48*(2), 333–344. https://doi.org/10.2307/1128626

McMurray, B., & Aslin, R. N. (2004). Anticipatory eye movements reveal infants' auditory and visual categories. *Infancy*, *6*(2), 203–229. https://doi.org/10.1207/s15327078in0602_4

Mcquaid, N. E., Bibok, M. B., & Carpendale, J. I. M. (2009). Relation Between Maternal Contingent Responsiveness and Infant Social Expectations. *Infancy*, *14*(3), 390–401. https://doi.org/10.1080/15250000902839955

Mehoudar, E., Arizpe, J., Baker, C. I., & Yovel, G. (2014). Faces in the eye of the beholder: Unique and stable eye scanning patterns of individual observers. *Journal of Vision*, *14*(7), 6. https://doi.org/10.1167/14.7.6

Meltzoff, A. N. (1988). Imitation of televised models by infants. *Child Development*, *59*(5), 1221–1229. https://doi.org/10.2307/1130485

Meltzoff, A. N., & Moore, M. K. (1998). Object representation, identity, and the paradox of early permanence: Steps toward a new framework. *Infant Behavior and Development*, *21*(2), 201–235. https://doi.org/10.1016/S0163-6383(98)90003-0

Minar, N. J., & Lewkowicz, D. J. (2017). Overcoming the other-race effect in infancy with multisensory redundancy: 10--12-month-olds discriminate dynamic other-race faces producing speech. *Developmental Science*, e12604. https://doi.org/10.1111/desc.12604

Mirman, D., Dixon, J. A., & Magnuson, J. S. (2008). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of Memory and Language*, *59*(4), 475–494. https://doi.org/10.1016/j.jml.2007.11.006

Montague, D. P. F., & Walker-Andrews, A. S. (2001). Peekaboo: a new look at infants’ perception of emotion expressions. *Developmental Psychology*, *37*(6), 826–838. https://doi.org/10.1037/0012-1649.37.6.826

Munakata, Y. (2000). Challenges to the violation-of-expectation paradigm: Throwing the conceptual baby out with the perceptual processing bathwater? *Infancy*, *1*(4), 471–477. https://doi.org/10.1207/S15327078IN0104_7

Munakata, Y., McClelland, J. L., Johnson, M. H., & Siegler, R. S. (1997).

Rethinking infant knowledge: Toward an adaptive process account of successes and failures in object permanence tasks. *Psychological Review*, *104*(4), 686. https://doi.org/10.1037/0033-295X.104.4.686

Murphy, P. R., O’Connell, R. G., O’Sullivan, M., Robertson, I. H., & Balsters, J. H. (2014). Pupil diameter covaries with BOLD activity in human locus coeruleus: Pupil Diameter and Locus Coeruleus Activity. *Human Brain Mapping*, *35*(8), 4140–4154. https://doi.org/10.1002/hbm.22466

Murray, L., & Trevarthen, C. (1986). The infant’s role in mother–infant communications. *Journal of Child Language*, *13*(01), 15–29. https://doi.org/10.1017/S0305000900000271

Nadel, J., Carchon, I., Kervella, C., Marcelli, D., & Râ, D. (1999). Expectancies for social contingency in 2-month-olds. *Developmental Science*, *2*(2), 164–173. https://doi.org/10.1111/1467-7687.00065

Neisser, U. (1976). *Cognition and reality: Principles and implications of cognitive psychology*. New York, NY: Henry Holt & Co.

Nelson, C. A., Morse, P., & Leavitt, L. (1979). Recognition of facial expressions by seven-month-old infants. *Child Development*, *50*(4), 1239–1242. https://doi.org/10.2307/1129358

Nielsen, M., Simcock, G., & Jenkins, L. (2008). The effect of social engagement on 24-month-olds’ imitation from live and televised models. *Developmental Science*, *11*(5), 722–731. https://doi.org/10.1111/j.1467-7687.2008.00722.x

Noton, D., & Stark, L. (1971a). Eye movements and visual perception. *Scientific American*, *224*(6), 34–43.

Noton, D., & Stark, L. (1971b). Scanpaths in saccadic eye movements while viewing and recognizing patterns. *Vision Research*, *11*(9), 929--IN8.

https://doi.org/10.1016/0042-6989(71)90213-6

Nyström, P., Bölte, S., Falck-Ytter, T., & The EASE Team. (2017). Responding to Other People's Direct Gaze: Alterations in Gaze Behavior in Infants at Risk for Autism Occur on Very Short Timescales. *Journal of Autism and Developmental Disorders*, *47*(11), 3498–3509. https://doi.org/10.1007/s10803-017-3253-7

O'Doherty, J., Winston, J., Critchley, H., Perrett, D., Burt, D. M., Dolan, R. J., … Dolan, R. J. (2003). Beauty in a smile: the role of medial orbitofrontal cortex in facial attractiveness. *Neuropsychologia*, *41*(2), 147–155. https://doi.org/10.1016/S0028-3932(02)00145-8

Oakes, L. M. (2012). Advances in eye tracking in infancy research. *Infancy*, *17*(1), 1–8. https://doi.org/10.1111/j.1532-7078.2011.00101.x

Office for National Statistics. (2011). *2011 Census: Cultural diversity in Kent*. https://doi.org/10.5257/census/aggregate-2011-1

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, *308*(5719), 255–258. https://doi.org/10.1126/science.1107621

Partala, T., & Surakka, V. (2003). Pupil size variation as an indication of affective processing. *International Journal of Human Computer Studies*, *59*(1–2), 185–198. https://doi.org/10.1016/S1071-5819(03)00017-X

Peltola, M. J., Forssman, L., Puura, K., van Ijzendoorn, M. H., & Leppänen, J. M. (2015). Attention to Faces Expressing Negative Emotion at 7 Months Predicts Attachment Security at 14 Months. *Child Development*, *86*(5), 1321–1332. https://doi.org/10.1111/cdev.12380

Peltola, M. J., Hietanen, J. K., Forssman, L., & Leppänen, J. M. (2013). The emergence and stability of the attentional bias to fearful faces in infancy. *Infancy*, *18*(6), 905–926. https://doi.org/10.1111/infa.12013

Peltola, M. J., Leppänen, J. M., Mäki, S., & Hietanen, J. K. (2009). Emergence of enhanced attention to fearful faces between 5 and 7 months of age. *Social Cognitive and Affective Neuroscience*, *4*(2), 134–142. https://doi.org/10.1093/scan/nsn046

Peltola, M. J., Leppänen, J. M., Vogel-Farley, V. K., Hietanen, J. K., & Nelson, C. A. (2009). Fearful faces but not fearful eyes alone delay attention disengagement in 7-month-old infants. *Emotion (Washington, D.C.)*, *9*(4), 560–565. https://doi.org/10.1037/a0015806

Piaget, J. (1929). *The child's conception of the world*. Savage, Maryland: Littlefield Adams.

Piaget, J. (1954). *The construction of reality in the child*. Basic Books.

Piaget, J. (1964). Part I: Cognitive development in children: Piaget development and learning. *Journal of Research in Science Teaching*, *2*(3), 176–186. https://doi.org/10.1002/tea.3660020306

Pickron, C. B., Fava, E., & Scott, L. S. (2017). Follow My Gaze: Face Race and Sex Influence Gaze-Cued Attention in Infancy. *Infancy*, *22*(5), 626–644. https://doi.org/10.1111/infa.12180

Powell, G., Wass, S. V., Erichsen, J. T., & Leekam, S. R. (2016). First evidence of the feasibility of gaze-contingent attention training for school children with autism. *Autism*, *20*(8), 927–937. https://doi.org/10.1177/1362361315617880

Quinn, P. C., Anzures, G., Izard, C. E., Lee, K., Pascalis, O., Slater, A. M., & Tanaka, J. W. (2011). Looking across domains to understand infant representation of emotion. *Emotion Review*, *3*(2), 197–206. https://doi.org/10.1177/1754073910387941

Quinn, P. C., Lee, K., & Pascalis, O. (2019). Face Processing in Infancy and

Beyond: The Case of Social Categories. *Annual Review of Psychology*, *70*, 165–189. https://doi.org/10.1146/annurev-psych-010418-102753

Quinn, P. C., Yahr, J., Kuhn, A., Slater, A. M., & Pascalis, O. (2002). Representation of the gender of human faces by infants: A preference for female. *Perception*, *31*(9), 1109–1121. https://doi.org/10.1068/p3331

Ramsay, J., & Silverman, B. (1997). *Functional data analysis*. New York: Springer-Verlag.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372–422. https://doi.org/10.1037/0033-2909.124.3.372

Reid, V. M., & Striano, T. (2005). Adult gaze influences infant attention and object processing: Implications for cognitive neuroscience. *European Journal of Neuroscience*, *21*(6), 1763–1766. https://doi.org/10.1111/j.1460-9568.2005.03986.x

Reilly, J., Kelly, A., Kim, S. H., Jett, S., & Zuckerman, B. (2019). The human task-evoked pupillary response function is linear: Implications for baseline response scaling in pupillometry. *Behavior Research Methods*, *51*(2), 865–878. https://doi.org/10.3758/s13428-018-1134-4

Richoz, A.-R., Lao, J., Pascalis, O., & Caldara, R. (2018). Tracking the recognition of static and dynamic facial expressions of emotion across the life span. *Journal of Vision*, *18*(9), 5. https://doi.org/10.1167/18.9.5

Rivera, S. M., Wakeley, A., & Langer, J. (1999). The drawbridge phenomenon: Representational reasoning or perceptual preference? *Developmental Psychology*, *35*(2), 427. https://doi.org/10.1037/0012-1649.35.2.427

Ruba, A. L., Johnson, K. M., Harris, L. T., & Wilbourn, M. P. (2017).

Developmental changes in infants’ categorization of anger and disgust facial expressions. *Developmental Psychology*, *53*(10), 1826–1832. https://doi.org/10.1037/dev0000381

Ruba, A. L., Meltzoff, A. N., & Repacholi, B. M. (2019). How do you feel? Preverbal infants match negative emotions to events. *Developmental Psychology*, *55*(6), 1138–1149. https://doi.org/10.1037/dev0000711

Salapatek, P., & Kessen, W. (1966). Visual scanning of triangles by the human newborn. *Journal of Experimental Child Psychology*, *3*(2), 155–167. https://doi.org/10.1016/0022-0965(66)90090-7

Sangrigoli, S., & De Schonen, S. (2004). Recognition of own-race and other-race faces by three-month-old infants. *Journal of Child Psychology and Psychiatry*, *45*(7), 1219–1227. https://doi.org/10.1111/j.1469-7610.2004.00319.x

Sangrigoli, S., Pallier, C., Argenti, A.-M., Ventureyra, V. A. G., & de Schonen, S. (2005). Reversibility of the Other-Race Effect in Face Recognition During Childhood. *Psychological Science*, *16*(6), 440–444. https://doi.org/10.1111/j.0956-7976.2005.01554.x

Sara, S. J. (2009). The locus coeruleus and noradrenergic modulation of cognition. *Nature Reviews. Neuroscience*, *10*(3), 211–223. https://doi.org/10.1038/nrn2573

Sara, S. J., & Bouret, S. (2012). Orienting and Reorienting: The Locus Coeruleus Mediates Cognition through Arousal. *Neuron*, *76*(1), 130–141. https://doi.org/10.1016/j.neuron.2012.09.011

Sato, W., Fujimura, T., & Suzuki, N. (2008). Enhanced facial EMG activity in response to dynamic facial expressions. *International Journal of Psychophysiology*, *70*(1), 70–74. https://doi.org/10.1016/j.ijpsycho.2008.06.001

Scaife, M., & Bruner, J. S. (1975). The capacity for joint visual attention in the infant. *Nature*, *253*(5489), 265–266. https://doi.org/10.1038/253265a0

Scheller, E., Büchel, C., & Gamer, M. (2012). Diagnostic features of emotional expressions are processed preferentially. *PloS One*, *7*(7), e41792. https://doi.org/10.1371/journal.pone.0041792

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., & Vogeley, K. (2013). Toward a second-person neuroscience 1. *Behavioral and Brain Sciences*, *36*(4), 393–414. https://doi.org/10.1017/S0140525X12000660

Schilling, T. H. (2000). Infants' looking at possible and impossible screen rotations: The role of familiarization. *Infancy*, *1*(4), 389–402. https://doi.org/10.1207/S15327078IN0104_2

Schmuckler, M. A. (2001). What is ecological validity? A dimensional analysis. *Infancy*, *2*(4), 419–436. https://doi.org/10.1207/S15327078IN0204_02

Schütz, A. C., Braun, D. I., & Gegenfurtner, K. R. (2011). Eye movements and perception: A selective review. *Journal of Vision*, *11*(5), 9. https://doi.org/10.1167/11.5.9

Schwartz, G. M., Izard, C. E., & Ansul, S. E. (1985). The 5-month-old's ability to discriminate facial expressions of emotion. *Infant Behavior and Development*, *8*(1), 65–77. https://doi.org/10.1016/S0163-6383(85)80017-5

Schyns, P. G., Petro, L. S., & Smith, M. L. (2007). Dynamics of Visual Information Integration in the Brain for Categorizing Facial Expressions. *Current Biology*, *17*(18), 1580–1585. https://doi.org/10.1016/j.cub.2007.08.048

Senju, A., & Csibra, G. (2008). Gaze Following in Human Infants Depends on Communicative Signals. *Current Biology*, *18*(9), 668–671. https://doi.org/10.1016/j.cub.2008.03.059

Senju, A., Csibra, G., & Johnson, M. H. (2008). Understanding the referential nature of looking: Infants' preference for object-directed gaze. *Cognition*, *108*(2), 303–319. https://doi.org/10.1016/j.cognition.2008.02.009

Sepeta, L., Tsuchiya, N., Davies, M. S., Sigman, M., Bookheimer, S. Y., & Dapretto, M. (2012). Abnormal social reward processing in autism as indexed by pupillary responses to happy faces. *Journal of Neurodevelopmental Disorders*, *4*(1), 17. https://doi.org/10.1186/1866-1955-4-17

Serrano, J. M., Iglesias, J., & Loeches, A. (1992). Visual discrimination and recognition of facial expressions of anger, fear, and surprise in 4-to 6-month-old infants. *Developmental Psychobiology*, *25*(6), 411–425. https://doi.org/10.1002/dev.420250603

Singarajah, A., Chanley, J., Gutierrez, Y., Cordon, Y., Nguyen, B., Burakowski, L., & Johnson, S. P. (2017). Infant attention to same- and other-race faces. *Cognition*, *159*, 76–84. https://doi.org/10.1016/j.cognition.2016.11.006

Sirois, S., & Brisson, J. (2014). Pupillometry. *Wiley Interdisciplinary Reviews: Cognitive Science*, *5*(6), 679–692. https://doi.org/10.1002/wcs.1323

Sirois, S., & Jackson, I. R. (2012). Pupil dilation and object permanence in infants. *Infancy*, *17*(1), 61–78. https://doi.org/10.1111/j.1532-7078.2011.00096.x

Skinner, B. F. (1935). The generic nature of the concepts of stimulus and response. *The Journal of General Psychology*, *12*(1), 40–65. https://doi.org/10.1080/00221309.1935.9920087

Slater, A. M., Von der Schulenburg, C., Brown, E., Badenoch, M., Butterworth, G., Parsons, S., & Samuels, C. (1998). Newborn infants prefer attractive faces. *Infant Behavior and Development*, *21*(2), 345–354. https://doi.org/10.1016/S0163-6383(98)90011-X

Slater, A., & Morison, V. (1985). Shape constancy and slant perception at birth. *Perception*, *14*(3), 337–344. https://doi.org/10.1068/p140337

Smith, L. B. (1999). Do infants possess innate knowledge structures? The con side. *Developmental Science*, *2*(2), 133–144. https://doi.org/10.1111/1467-7687.00062

Smith, M. L., Cottrell, G. W., Gosselin, F., & Schyns, P. G. (2005). Transmitting and decoding facial expressions. *Psychological Science*, *16*(3), 184–189. https://doi.org/10.1111/j.0956-7976.2005.00801.x

Soken, N. H., & Pick, A. D. (1999). Infants' Perception of Dynamic Affective Expressions: Do Infants Distinguish Specific Expressions? *Child Development*, *70*(6), 1275. https://doi.org/10.1111/1467-8624.00093

Sokolov, E. N. (1963). *Perception and the conditioned reflex*. New York, NY: Macmillan.

Soussignan, R., Dollion, N., Schaal, B., Durand, K., Reissland, N., & Baudouin, J.-Y. (2017). Mimicking emotions: how 3--12-month-old infants use the facial expressions and eyes of a model. *Cognition and Emotion*, 1–16. https://doi.org/10.1080/02699931.2017.1359015

Soussignan, R., Nadel, J., Canet, P., & Gerardin, P. (2006). Sensitivity to Social Contingency and Positive Emotion in 2-Month-Olds. *Infancy*, *10*(2), 123–144. https://doi.org/10.1207/s15327078in1002_2

Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science*, *18*(7), 587–592. https://doi.org/10.1111/j.1467-9280.2007.01944.x

Spelke, E. S. (1976). Infants' intermodal perception of events. *Cognitive Psychology*, *8*(4), 553–560. https://doi.org/10.1016/0010-0285(76)90018-9

Spelke, E. S. (1998). Nativism, empiricism, and the origins of knowledge. *Infant Behavior and Development*, *21*(2), 181–200. https://doi.org/10.1016/S0163-6383(98)90002-9

Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review*, *99*(4), 605. https://doi.org/10.1037/0033-295X.99.4.605

Spelke, E. S., Kestenbaum, R., Simons, D. J., & Wein, D. (1995). Spatiotemporal continuity, smoothness of motion and object identity in infancy. *British Journal of Developmental Psychology*, *13*(2), 113–142. https://doi.org/10.1111/j.2044-835X.1995.tb00669.x

SR Research. (2010). *EyeLink 1000 User Manual* (v1.5.2). Mississauga, ON, Canada: SR Research Ltd.

Steinhauer, S. R., Siegle, G. J., Condray, R., & Pless, M. (2004). Sympathetic and parasympathetic innervation of pupillary dilation during sustained processing. *International Journal of Psychophysiology*, *52*(1), 77–86. https://doi.org/10.1016/j.ijpsycho.2003.12.005

Stern, D. N., Hofer, L., Haft, W., & Dore, J. (1985). Affect attunement: The sharing of feeling states between mother and infant by means of inter-modal fluency. *Social Perception in Infants*, 249–268.

Stern, W. (1924). *Psychology of early childhood*. New York: H. Colt & Co.

Striano, T., Henning, A., & Stahl, D. (2005). Sensitivity To Social Contingencies Between 1 and 3 Months of Age. *Developmental Science*, *8*(6), 509–518. https://doi.org/10.1111/j.1467-7687.2005.00442.x

Striano, T., Henning, A., & Stahl, D. (2006). Sensitivity to interpersonal timing at 3 and 6 months of age. *Interaction Studies*, *7*(2), 251–271.

https://doi.org/10.1075/is.7.2.08str

Striano, T., Henning, A., & Vaish, A. (2006). Selective looking by 12-month-olds to a temporally contingent partner. *Interaction Studies*, *7*(2), 233–250. https://doi.org/10.1075/is.7.2.07str

Striano, T., & Reid, V. M. (2006). Social cognition in the first year. *Trends in Cognitive Sciences*, *10*(10), 471–476. https://doi.org/10.1016/j.tics.2006.08.006

Striano, T., & Stahl, D. (2005). Sensitivity to triadic attention in early infancy. *Developmental Science*, *8*(4), 333–343. https://doi.org/10.1111/j.1467-7687.2005.00421.x

Sugden, N. A., & Marquis, A. R. (2017). Meta-Analytic Review of the Development of Face Discrimination in Infancy: Face Race, Face Gender, Infant Age, and Methodology Moderate Face Discrimination. *Psychological Bulletin*, *143*(11), 1201–1244. https://doi.org/10.1037/bul0000116

Tamis-LeMonda, C. S., Kuchirko, Y., Luo, R., Escobar, K., & Bornstein, M. H. (2017). Power in methods: Language to infants in structured and naturalistic contexts. *Developmental Science*, *20*(6), e12456. https://doi.org/10.1111/desc.12456

Teller, D. Y. (1979). The forced-choice preferential looking procedure: A psychophysical technique for use with human infants. *Infant Behavior and Development*, *2,* 135–153. https://doi.org/10.1016/S0163-6383(79)80016-8

Teller, D. Y., & Movshon, J. A. (1986). Visual development. *Vision Research*, *26*(9), 1483–1506. https://doi.org/10.1016/0042-6989(86)90169-0

Tenenbaum, E. J., Shah, R. J., Sobel, D. M., Malle, B. F., & Morgan, J. L. (2013). Increased Focus on the Mouth Among Infants in the First Year of Life: A Longitudinal Eye-Tracking Study. *Infancy*, *18*(4), 534–553.

https://doi.org/10.1111/j.1532-7078.2012.00135.x

Thomas, K. M., Drevets, W. C., Whalen, P. J., Eccard, C. H., Dahl, R. E., Ryan, N. D., & Casey, B. J. (2001). Amygdala response to facial expressions in children and adults. *Biological Psychiatry*, *49*(4), 309–316. https://doi.org/10.1016/S0006-3223(00)01066-0

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: the origins of cultural cognition. *The Behavioral and Brain Sciences*, *28*(5), 675–691; discussion 691-735. https://doi.org/10.1017/S0140525X05000129

Trevarthen, C. (1979). Communication and cooperation in early infancy: A description of primary intersubjectivity. In M. Bullowa (Ed.), *Before speech: The beginning of interpersonal communication* (Vol. 1, pp. 530–571). Cambridge, UK: Cambridge University Press.

Tronick, E. Z., Als, H., Adamson, L. B., Wise, S., & Brazelton, T. B. (1978). The infant’s response to entrapment between contradictory messages in face-to-face interaction. *Journal of the American Academy of Child Psychiatry*, *17*(1), 1–13. https://doi.org/http://dx.doi.org/10.1016/S0002-7138(09)62273-1

Vaish, A., Grossmann, T., & Woodward, A. L. (2008). Not all emotions are created equal: the negativity bias in social-emotional development. *Psychological Bulletin*, *134*(3), 383–403. https://doi.org/10.1037/0033-2909.134.3.383

Vernetti, A., Senju, A., Charman, T., Johnson, M. H., Gliga, T., Team, B., & others. (2018). Simulating interaction: Using gaze-contingent eye-tracking to measure the reward value of social signals in toddlers with and without autism. *Developmental Cognitive Neuroscience*, *29*, 21–29. https://doi.org/10.1016/j.dcn.2017.08.004

Vernetti, A., Smith, T. J., & Senju, A. (2017). Gaze-contingent reinforcement learning reveals incentive value of social signals in young children and adults. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1850), 20162747. https://doi.org/10.1098/rspb.2016.2747

Vinter, A. (1986). The role of movement in eliciting early imitations. *Child Development*, *57*(Feb., 1986), 66–71. https://doi.org/10.2307/1130638

Võ, M. L. H., Smith, T. J., Mital, P. K., & Henderson, J. M. (2012). Do the eyes really have it? Dynamic allocation of attention when viewing moving faces. *Journal of Vision*, *12*(13), 3–3. https://doi.org/10.1167/12.13.3

Von Helmholtz, H. (1925). Physiological optics. *Optical Society of America*, *3,* 318.

Von Hofsten, C., Kochukhova, O., & Rosander, K. (2007). Predictive tracking over occlusions by 4-month-old infants. *Developmental Science*, *10*(5), 625–640. https://doi.org/10.1111/j.1467-7687.2007.00604.x

Vygotsky, L. S. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.

Wagner, J. B., Luyster, R. J., Tager-Flusberg, H., & Nelson, C. A. (2016). Greater Pupil Size in Response to Emotional Faces as an Early Marker of Social-Communicative Difficulties in Infants at High Risk for Autism. *Infancy*, *21*(5), 560–581. https://doi.org/10.1111/infa.12128

Walden, T. A., & Ogan, T. A. (1988). The development of social referencing. *Child Development*, *59*(5), 1230–1240. https://doi.org/10.2307/1130486

Walker-Andrews, A. S. (1997). Infants' perception of expressive behaviors: Differentiation of multimodal information. *Psychological Bulletin*, *121*(3), 437–456. https://doi.org/10.1037/0033-2909.121.3.437

Walker-Andrews, A. S., & Bahrick, L. E. (2001). Perceiving the real world: Infants' detection of and memory for social information. *Infancy, 2*(4), 469–481.

https://doi.org/10.1207/S15327078IN0204_05

Wallace, D. B., Franklin, M. B., & Keegan, R. T. (1994). The observing eye: A century of baby diaries. *Human Development*, *37*(1), 1–29. https://doi.org/10.1159/000278233

Wang, Q., Bolhuis, J., Rothkopf, C. A., Kolling, T., Knopf, M., & Triesch, J. (2012). Infants in control: Rapid anticipation of action outcomes in a gaze-contingent paradigm. *PLoS ONE*, *7*(2), 1–6. https://doi.org/10.1371/journal.pone.0030884

Wass, S. V., Forssman, L., & Leppänen, J. (2014). Robustness and precision: How data quality may influence key dependent variables in infant eye-tracker analyses. *Infancy*, *19*(5), 427–460. https://doi.org/10.1111/infa.12055

Wass, S. V., Noreika, V., Georgieva, S., Clackson, K., Brightman, L., Nutbrown, R., … Leong, V. (2018). Parental neural responsivity to infants' visual attention: how mature brains influence immature brains during social interaction. *PLoS Biology, 16*(12), e2006328. https://doi.org/10.1371/journal.pbio.2006328

Wass, S. V., Porayska-Pomsta, K., & Johnson, M. H. (2011). Training attentional control in infancy. *Current Biology*, *21*(18), 1543–1547. https://doi.org/10.1016/j.cub.2011.08.004

Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychological Review*, *20*(2), 158–177. https://doi.org/10.1037/h0074428

Watson, J. B. (1928). *The ways of behaviorism*. Oxford, England: Harper.

Watson, J. B., & McDougall, W. (1929). *The battle of behaviorism: An exposition and an exposure*. WW Norton & Company.

Watson, J. B., & Rayner, R. (1920). Conditioned emotional reactions. *Journal of Experimental Psychology*, *3*(1), 1–14. https://doi.org/10.1037/h0069608

Wendt, D., Brand, T., & Kollmeier, B. (2014). An eye-tracking paradigm for analyzing the processing time of sentences with different linguistic complexities. *PLoS One*, *9*(6), e100186. https://doi.org/10.1371/journal.pone.0100186

Werner, H. (1926). *Comparative psychology of mental development*. New York: International Universities Press.

Weyers, P., Mühlberger, A., Hefele, C., & Pauli, P. (2006). Electromyographic responses to static and dynamic avatar emotional facial expressions. *Psychophysiology*, *43*(5), 450–453. https://doi.org/10.1111/j.1469-8986.2006.00451.x

Wheeler, A., Anzures, G., Quinn, P. C., Pascalis, O., Omrin, D. S., & Lee, K. (2011). Caucasian infants scan own-and other-race faces differently. *PloS One*, *6*(4), e18621. https://doi.org/10.1371/journal.pone.0018621

Widen, S. C. (2013). Children's interpretation of facial expressions: The long path from valence-based to specific discrete categories. *Emotion Review*, *5*(1), 72–77. https://doi.org/10.1177/1754073912451492

Widen, S. C., & Russell, J. A. (2015). Do Dynamic Facial Expressions Convey Emotions to Children Better Than Do Static Ones? *Journal of Cognition and Development*, *16*(5), 802–811. https://doi.org/10.1080/15248372.2014.916295

Wilcox, B. M., & Clayton, F. L. (1968). Infant visual fixation on motion pictures of the human face. *Journal of Experimental Child Psychology*, *6*(1), 22–32. https://doi.org/10.1016/0022-0965(68)90068-4

Wilcox, T., Stubbs, J. A., Wheeler, L., & Alexander, G. M. (2013). Infants' scanning of dynamic faces during the first year. *Infant Behavior and Development*, *36*(4), 513–516. https://doi.org/10.1016/j.infbeh.2013.05.001

Wilms, M., Schilbach, L., Pfeiffer, U., Bente, G., Fink, G. R., & Vogeley, K. (2010).

It’s in your eyes-using gaze-contingent stimuli to create truly interactive paradigms for social cognitive and affective neuroscience. *Social Cognitive and Affective Neuroscience, 5*(1), 98–107. https://doi.org/10.1093/scan/nsq024

Wolff, P. H. (1987). *The development of behavioral states and the expression of emotions in early infancy*. University of Chicago Press.

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor’s reach. *Cognition*, *69*(1), 1–34. https://doi.org/10.1016/S0010-0277(98)00058-4

Wundt, W. M. (1909). *Völkerpsychologie* (Vol. 3). W. Engelmann.

Wynn, K. (1992). Addition and subtraction by human infants. *Nature*, *358*(6389), 749–750. https://doi.org/10.1038/358749a0

Xiao, N. G., Quinn, P. C., Liu, S., Ge, L., Pascalis, O., & Lee, K. (2015). Eye tracking reveals a crucial role for facial motion in recognition of faces by infants. *Developmental Psychology*, *51*(6), 744–757. https://doi.org/10.1037/dev0000019

Xiao, N. G., Wu, R., Quinn, P. C., Liu, S., Tummeltshammer, K. S., Kirkham, N. Z., … Lee, K. (2018). Infants Rely More on Gaze Cues From Own-Race Than Other-Race Adults for Learning Under Uncertainty. *Child Development*, *89*(3), e229–e244. https://doi.org/10.1111/cdev.12798

Xiao, W. S., Quinn, P. C., Pascalis, O., & Lee, K. (2014). Own-and other-race face scanning in infants: Implications for perceptual narrowing. *Developmental Psychobiology*, *56*(2), 262–273. https://doi.org/10.1002/dev.21196

Xiao, W. S., Xiao, N. G., Quinn, P. C., Anzures, G., & Lee, K. (2013). Development of face scanning for own-and other-race faces in infancy. *International Journal of Behavioral Development*, *37*(2), 100–105. https://doi.org/10.1177/0165025412467584

Yarbus, A. L. (1967). Eye Movements During Perception of Complex Objects. *Eye Movements and Vision*, 171–211. https://doi.org/10.1007/978-1-4899-5379-7_8

Young-Browne, G., Rosenfeld, H. M., & Horowitz, F. D. (1977). Infant discrimination of facial expressions. *Child Development, 48*(2), 555–562. https://doi.org/10.2307/1128653

Yrttiaho, S., Niehaus, D., Thomas, E., & Leppänen, J. M. (2017). Mothers' pupillary responses to infant facial expressions. *Behavioral and Brain Functions, 13*(1), 2. https://doi.org/10.1186/s12993-017-0120-9

Yu, C., & Smith, L. B. (2013). Joint attention without gaze following: Human infants and their parents coordinate visual attention to objects through eye-hand coordination. *PloS One, 8*(11), e79659. https://doi.org/10.1371/journal.pone.0079659

Yu, C., Yurovsky, D., & Xu, T. (2012). Visual data mining: An exploratory approach to analyzing temporal patterns of eye movements. *Infancy, 17*(1), 33–60. https://doi.org/10.1111/j.1532-7078.2011.00095.x