

Understanding the Naturalistic Toddler Visual Environment of Faces, Bodies, and Emotional
Expressions: How Individual Differences Relate to this Socioemotional Input

Emma Jayne Jackson

PhD

University of York

Psychology

September 2024

Abstract

Toddlerhood is a period of rapid socioemotional development. A key group of socioemotional skills relate to extracting information from faces and bodies. Knowing the role of environmental visual input in driving this development is crucial, particularly in toddlerhood, and especially given that developmental trajectories in this domain are very diverse. The first step in understanding what toddlers could learn from the environment is characterising what information is in their view. Prior research is limited but would suggest that faces specifically are infrequent in the toddler view. This thesis focuses on the degree to which toddlers see faces, bodies and facial emotional expressions (FEE) as they navigate their naturalistic home environments, as well as factors that shape individual differences in this input. Chapters 1 and 2 focus on faces and bodies, while chapters 4 and 5 focus on FEE. To contextualise toddler input, we must understand elements of mature, adult processing and viewing of faces, bodies and FEE. So, there is some discussion of adult processing, and chapter 3 focuses on adult processing of real FEE. Not only this, but we characterise the corresponding visual input of adult caregivers with respect to faces, bodies and FEE in chapters 1 and 4. Overall, we address one question: what are the natural statistics of socioemotional stimuli in the toddler visual input? We find that, as compared to caregivers, toddlers infrequently see faces, and less often see bodies. Real-world FEE are highly ambiguous, especially threat-related FEE. Both toddlers and caregivers rarely see threat-related FEE, with most faces in view being happy or neutral. Temperament and the presence of child siblings are related to individual differences in the frequency of FEE in the toddler input but not faces or bodies. This raises fundamental questions on how toddlers may learn from scarce visual input.

Declaration

I declare that this thesis is a presentation of original work, and I am the sole author.

Where there are multiple contributors to the work forming the basis for an empirical chapter, this is stated. This work has not previously been presented for a degree or other qualification at this University or elsewhere. All sources are acknowledged as references

ABSTRACT	2
DECLARATION	3
ACKNOWLEDGEMENTS	12
LITERATURE REVIEW	13
KNOWN CHARACTERISTICS OF FACE, FACIAL EMOTIONAL EXPRESSION AND BODY PROCESSING	13
<i>Face Processing</i>	14
Typical Adult Face Processing.	14
The Importance of Understanding Face Processing Development in Infancy and Toddlerhood.	17
Characterising The Emergence of Face Processing in Infancy and Toddlerhood.	20
Face Processing at Birth.	21
Face Processing Development Between 3- and 6-months.	22
Face Processing Development Between 6- and 9-months.	24
Face Processing Development Between 9- and 12-months.	25
Face Processing Development Between 12- and 24-months.	26
Face Processing Development in Toddlers.	27
Face Processing Development in Childhood and Beyond.	28
Theories of Face Processing Development.	29
<i>Facial Emotional Expression Processing</i>	31
Typical Adult Facial Expression Processing.	34
Issues and Controversies in Defining and Conceptualising Facial Emotional Expressions.	36
Basic Emotion Theories.	36
Modern Challenges to the Idea of Basic Emotions.	38
The Specific Role of Context in Judging Facial Expressions.	40
Characterising The Emergence of Facial Emotional Expression Processing.	41
Facial Expression Processing in Neonates.	41
Facial Expression Processing at 1- to 3-Months.	44
Facial Expression Processing at 4- to 6-Months.	46
Facial Expression Processing at 7- to 9-Months.	49

Facial Expression Processing at 10- to 14-Months.	53
Facial Expression Processing at 15- to 24-Months.	57
Facial Expression Processing in Toddlers and Children.	58
Summary of Facial Expression Processing Development.	60
<i>Body Processing</i>	61
Typical Adult Body Processing.	61
Adult Recognition of Bodies.	62
Processing of Motion and Posture Information from Bodies.	62
Characterising The Emergence of Body Processing in Infancy and Toddlerhood.	64
Body Processing at Birth to 6-Months.	64
Body Processing at 7- to 12-Months.	65
Body Processing at 13- to 24-Months.	67
TEMPERAMENTAL DIFFERENCES AND SOCIOEMOTIONAL DEVELOPMENT	68
<i>Associations Between Early Temperament and Childhood Social Development</i>	68
<i>Potential Explanations for the Relationship Between Temperament and Socioemotional Development</i>	69
PROCESSES OF DEVELOPMENT: MATURATION VERSUS ENVIRONMENTAL INPUT	75
KNOWN CHARACTERISTICS OF THE VISUAL SOCIOEMOTIONAL ENVIRONMENT	79
<i>Known Characteristics of the Adult Visual Environment</i>	79
<i>Known Characteristic of the Infant and Toddler Visual Environments</i>	80
OVERALL RESEARCH AIMS	84
<i>The Projects at Hand</i>	85
Study 1.	85
Study 2.	86
Study 3.	87
Study 4.	87
Study 5.	88
GENERAL METHODS	89

PARTICIPANTS -----	90
APPARATUS -----	90
PROCEDURE -----	93
DATA EXCLUSION -----	94
ALGORITHMIC CODING -----	95
<i>Manual Coding for Face and Body Algorithm Validation</i> -----	97
<i>Validity of the Computer Vision Models for Face and Body Detection</i> -----	98
<i>Validity of the Computer Vision Models for Facial Emotional Expression Detection & Categorisation</i>	
-----	99
PARTICIPANT VARIABLES -----	100
<i>Temperament Assessments</i> -----	101
QUANTIFYING THE FREQUENCY OF FACES AND BODIES IN THE EGOCENTRIC VIEW OF TODDLERS ----	102
QUANTIFYING THE FREQUENCY OF FACES AND BODIES IN THE EGOCENTRIC VIEW OF TODDLERS -----	104
<i>Face and Body Processing Development</i> -----	106
<i>The Importance of Environmental Input</i> -----	109
<i>Age, Motor Ability, Motivation and the Environmental Input</i> -----	112
<i>Prior Quantification of the Input</i> -----	115
<i>Understanding the Face Decline in the Context of Adult Visual Input</i> -----	118
<i>Aims and Hypotheses</i> -----	120
METHODS -----	123
<i>Data Analysis Strategy</i> -----	123
RESULTS -----	123
<i>Incidence of Faces and Bodies in the Toddler View</i> -----	123
<i>Testing the Adequacy of Coarser Sampling Rates for Face Frequency</i> -----	124
<i>Comparison with Caregiver-View Frequency</i> -----	126
<i>The Frequency of Faces in the Caregiver Perspective</i> -----	128

DISCUSSION -----	129
<i>Variability and Future Research Directions</i> -----	134
<i>Conclusions</i> -----	136
FACTORS AFFECTING THE FREQUENCY OF FACES AND BODIES IN THE EGOCENTRIC VIEW OF TODDLERS	
-----	137
FACTORS AFFECTING THE FREQUENCY OF FACES AND BODIES IN THE EGOCENTRIC VIEW OF TODDLERS-----	139
<i>Prior Research on the Socioemotional Visual Input</i> -----	141
<i>Face and Body Processing Development and Variation in Ability</i> -----	145
Human Face Processing. -----	145
Toddlerhood as A Crucial Period for Face Processing Development. -----	149
Human Body Processing. -----	151
Toddlerhood as A Crucial Period for Body Processing Development. -----	154
<i>Individual Differences that May Shape the Environmental Input</i> -----	155
<i>Aims and Hypotheses</i> -----	164
METHODS-----	166
<i>Data Analysis Strategy</i> -----	166
RESULTS-----	167
<i>Presence of Siblings</i> -----	167
<i>Longitudinal Relations Between Input and Temperament</i> -----	170
<i>Concurrent Relations Between Input and Temperament</i> -----	173
DISCUSSION -----	178
<i>Sibling Status and the Naturalistic Input of Faces and Bodies</i> -----	180
<i>Temperament and the Naturalistic Input of Faces and Bodies</i> -----	183
<i>Overall Conclusion</i> -----	186
ADULT PERCEPTION OF NATURALISTIC, DYNAMIC EMOTIONAL EXPRESSIONS: CONSENSUS AND	
DISAGREEMENT -----	187

ADULT PERCEPTION OF NATURALISTIC, DYNAMIC EMOTIONAL EXPRESSIONS: CONSENSUS AND DISAGREEMENT -----	189
<i>Traditional Views on Facial Emotional Expressions</i> -----	190
<i>Contemporary Challenges to Traditional Views on Facial Emotional Expressions</i> -----	192
<i>The Importance of Naturalistic Research and Stimuli</i> -----	195
<i>The Importance of Understanding the Perception of Facial Emotional Expressions</i> -----	196
<i>Prior Research on Naturalistic Facial Emotional Expression Perception and Outstanding Questions</i>	
-----	198
<i>Agreement, Intensity and Confidence</i> -----	202
<i>The Influence of Emotion Category and Stimulus Model Age Group</i> -----	204
<i>Summary of Outstanding Questions</i> -----	206
<i>Aims and Hypotheses</i> -----	207
METHODS -----	209
<i>Participants</i> -----	209
<i>Stimuli</i> -----	210
<i>Apparatus</i> -----	212
<i>Procedure</i> -----	213
Set-Up, Instructions and Demographics. -----	213
Trial Structure. -----	214
<i>Data Analysis Strategy</i> -----	218
RESULTS -----	220
<i>Presence of Emotional Expressions in the Dataset</i> -----	220
<i>Consensus with Final Label Across Clips by Video Subject Age Group and Facial Emotional Expression</i>	
-----	222
<i>Average Intensity Rating by Video Subject Age Group and Facial Emotional Expression</i> -----	227
<i>Average Confidence Rating by Video Subject Age Group and Facial Emotional Expression</i> -----	232
<i>The Relationship Between Average Intensity, Average Confidence and Percentage Agreement</i> ----	238
<i>Clip Length as a Potential Confound</i> -----	241

DISCUSSION -----	243
<i>Limitations and Future Directions</i> -----	250
<i>Conclusions</i> -----	251
QUANTIFYING THE FREQUENCY OF EMOTIONAL EXPRESSIONS WITHIN THE TODDLER EGOCENTRIC VIEW	
-----	252
QUANTIFYING THE FREQUENCY OF EMOTIONAL EXPRESSIONS WITHIN THE TODDLER EGOCENTRIC VIEW-----	254
<i>The Importance of Emotional Expression Processing</i> -----	255
<i>The Importance of Infancy and Early Childhood for Emotional Expression Processing</i> -----	257
<i>The Hypothesised Role of the Environment in Infancy and Early Childhood</i> -----	261
<i>The Importance of Positive and Negative Emotional Expressions in the Second Year of Life</i> -----	265
<i>Prior Research on the Nature of the Environment</i> -----	267
<i>Outstanding Questions and Hypotheses</i> -----	268
METHODS-----	270
<i>Data Analysis Strategy</i> -----	271
RESULTS-----	272
<i>Presence of Emotional Expressions in the Toddler View</i> -----	272
<i>Presence of Emotional Expressions in the Caregiver View</i> -----	274
<i>Comparing the Toddler and Caregiver View in the Sample Subset with Both Perspectives</i> -----	276
DISCUSSION -----	282
<i>Findings and Analysis</i> -----	283
<i>The Infrequency of Negative FEE</i> -----	284
<i>Individual Variation and Future Research</i> -----	285
<i>Methodology and Limitations</i> -----	287
<i>Conclusions</i> -----	288
FACTORS AFFECTING THE FREQUENCY OF EMOTIONAL EXPRESSIONS WITHIN THE TODDLER	
EGOCENTRIC VIEW -----	290

FACTORS AFFECTING THE FREQUENCY OF EMOTIONAL EXPRESSIONS WITHIN THE TODDLER EGOCENTRIC VIEW -----	292
<i>Diversity in Facial Emotional Expression Processing Ability and its Implications</i> -----	293
<i>Prior Research on the Importance of Variation in the FEE Input</i> -----	298
<i>The Importance of Variation in Visual Input in Toddlerhood</i> -----	301
<i>Prior Research Characterising the Visual Input of Facial Emotional Expressions</i> -----	303
<i>The Role of Siblings</i> -----	304
<i>The Role of Temperament</i> -----	306
<i>Outstanding Questions and Hypotheses</i> -----	310
METHODS-----	311
<i>Data Analysis Strategy</i> -----	312
RESULTS-----	312
<i>Overall FEE Frequency and Variability</i> -----	312
<i>Presence of Siblings and FEE Frequency</i> -----	314
<i>Gender and FEE Frequency</i> -----	316
<i>Longitudinal IBQR Temperament and FEE Frequency</i> -----	318
<i>Concurrent ECBQ Temperament and FEE Frequency</i> -----	325
DISCUSSION -----	331
<i>The Implications for Differences in Socioemotional Processing Development</i> -----	337
<i>Overall Conclusions and Analysis</i> -----	339
 CHARACTERISING THE NATURALISTIC TODDLER VISUAL INPUT OF SOCIAL STIMULI: GENERAL	
DISCUSSION-----	341
 SUMMARY OF FINDINGS-----	 342
<i>Chapter 1</i> -----	343
<i>Chapter 2</i> -----	345
<i>Chapter 3</i> -----	347
<i>Chapter 4</i> -----	349

<i>Chapter 5</i>	350
<i>Summary</i>	352
IMPLICATIONS OF THESE FINDINGS	353
<i>Theoretical Implications</i>	353
<i>Methodological and Practical Implications</i>	360
LIMITATIONS AND FUTURE DIRECTIONS FOR RESEARCH	361
CONCLUSIONS	363
REFERENCES	366

Acknowledgements

There are a number of people whose insight, advice, care and support were essential in writing this document and making it through these past years in (more or less) one piece. First and foremost, Dr Elena Geangu, whose quick thinking, expertise, eye for detail and resilience was invaluable throughout. Many members of our lab and collaborators from outside it provided insight and plain hard work that made the impossible work of managing this video data somehow happen. There are too many individuals to name here, from individual undergraduate lab interns through to members of the computer sciences lab that we collaborated with, but all were essential.

Then, I need to thank some members of my family in particular for their kind words, patience and wisdom— LA, LJ, MA, AA and IJ. Each of you contributed support in your own way. A special thanks goes to The Exiles, who gave me a focus and community that kept me on an even keel. And last, but not least, my ever-supportive partner CM. Your inexhaustible patience, endless kindness and steady nerve have kept me grounded throughout these years.

Literature Review

Understanding what specifically toddlers could potentially extract from their visual input and how this may relate to factors driving development is essential. This is especially true in the domain of social processing; here, potential developmental trajectories are diverse, and this knowledge may be key in understanding the aetiologies of typical and atypical development. The environmental visual input could play a fundamental role. This is the information in the view of individuals as they navigate their environments. While the focus here is exclusively on the visual input, to the exclusion of vocal-auditory cues, it is important to be aware that these cues, such as vocal tone, exclamations and even semantic verbal content, are essential for typical socioemotional functioning (Morningstar et al., 2018; Scherer et al., 2003). However, these factors are beyond the scope of this present work. This literature review will focus on the emergence of processing skills relating to faces, facial emotional expressions (FEE) and bodies in the first years of life, as well as characteristics of mature adult processing. We will then discuss current research on the environmental visual input of toddlers with respect to faces, bodies and FEE, as well as factors that shape individual differences in this input, before outlining the purpose and structure of the remaining thesis chapters.

Known Characteristics of Face, Facial Emotional Expression and Body Processing

In order to assess the degree to which the environment may inform development of face, FEE and body processing, it is first necessary to outline key features of socioemotional stimulus processing in adults and toddlers.

Face Processing

Human faces are a rich source of information, giving perceivers cues relevant to many dimensions, from health (Henderson et al., 2016) through to internal affective and cognitive states (Ekman and Oster, 1979) through to identity (Russell et al., 2016; Shapiro and Penrod, 1986). First, we will discuss all aspects of face processing outside of FEE– which we will discuss in detail separately later– focusing particularly on identity processing from faces and factors that distinguish and make unique typical adult human face processing.

Typical Adult Face Processing.

Faces are (arguably) common within the environments of adults– we will discuss this in more detail later– and adults see a wide range of faces, with the average Western adult being familiar with around 5000 faces (Jenkins, Dowsett and Burton, 2018). Typically developed adults show a high degree of proficiency in processing faces for this wide range of informational cues-- indeed, it has been proposed that they form a special class of stimuli for adult humans (McKone and Robbins, 2011). For example, only humans and chimpanzees (the closest living primate relatives of humans) have been demonstrated to have cognitive systems specialised enough to holistic processing of upright facial configurations to demonstrate the ‘Thatcher effect’ (Weldon et al., 2013). Further, the existence of prosopagnosia, which is a selective deficit in face processing (usually restricted to identity recognition) in humans can be taken as evidence for a distinct and specialised modular “face processing” unit specialised for faces (Riddoch et al., 2008).

There are caveats on this face expertise: the expertise appears to be specific to certain categories of face (for example, familiar faces) (Burton, 2013), and adults can have difficulty parsing within-individual variability, especially for unfamiliar faces (Jenkins and Burton, 2011). Further, most evidence for this high degree of proficiency with certain categories of faces is taken from research focusing almost exclusively on identity processing, which neglects to acknowledge the wide variety of information that can be extracted from faces.

In addition to proficiency with faces being greatest for individually familiar faces, adult face processing is perceptually narrowed by category, meaning proficiency for categories of faces to which an individual is not typically exposed is lower. In the context of faces, these stimuli are individuals from cultural and racial out-groups. For example, adults show poorer recognition of faces belonging to people of races with whom they have less contact (the other-race effect, or ORE) (Chiroro and Valentine, 1995). Adults remember and recognise own-race faces with more ease than other-race faces (Meissner and Brigham, 2001). This extends to other metrics of face processing— e.g., facial emotional expressions, with poorer recognition of emotions in those from other cultures (Elfenbein and Ambady, 2002). These in-group superiority effects are all experience-dependent, being attenuated for racial minority groups, who typically have extensive perceptual experience with the majority racial outgroup; the ORE is also attenuated for racial majority group members who have more perceptual experience with the racial minority group tested (Walker and Hewstone, 2006). The ORE effect is a result of varying

perceptual expertise and experience with the categories of face tested, with early (pre-adolescent) experience seemingly being most formative (Zhou et al., 2019; Zhou, Elshiekh, and Moulson, 2019). So, the prior environmental input is highly predictive of adult face processing proficiency, with direct, causal relationships between aspects of this input and face processing.

One proposed characteristic feature of the adult cognitive system for face identity encoding, processing and recognition is the use of face “prototypes” or averages. This mechanism encodes faces in terms of their deviations from a cognitive face prototype, which is developed based on experience with faces; the deviation is mapped in a cognitive “face space” with faces able to deviate from the prototypical or average face along many different dimensions (Jozwik et al., 2022). Evidence for this can be seen in enhanced recognition for caricatures, the features of which deviate to an exaggerated degree from features from the prototypical face (Mauro and Kubovy, 1992); further, “distinctive” (i.e., non-prototypical) faces are more easily recognisable (Valentine and Bruce, 1986). This prototypical face structure could be learned from the environment, which may explain some of the findings in the domain of differing experience-related face processing skills levels dependent of the level of environmental exposure to different categories of face: separate prototypes exist for different categories of face, and these prototypes are less well-established for categories of face to which an individual has not been frequently exposed, especially over the course of development (Short et al., 2011). So, in this way, adult face processing is achieved through a process of establishing cognitive representations of normative faces and comparing encountered faces to these.

A final, notable characteristic of adult face processing is the degree of variation in skill level present within the general population. This variation can be within the normal range, unusual, or overtly a component of some pathology. Firstly, there exist prosopagnosics, individuals with very poor or non-existent face identity recognition, either as a result of brain injury or innate neurodevelopmental differences (de Gelder and Rouw, 2000). However, at the other end of the spectrum exist “super-recognisers”, people for whom face recognition and discrimination is markedly easier (Russell et al., 2009). In-between, there are various intermediate skill levels—for example, Autistic (Sasson, 2006) and Schizophrenic (Marwick and Hall, 2008) adults typically show a milder deficit in face processing.

The Importance of Understanding Face Processing Development in Infancy and Toddlerhood.

Face processing is an essential ability that infants must acquire, and this acquisition is likely to be greatly dependent on experience. Infant processing of faces is not merely a less practised version of adult face processing but involves differences in both the cognitive mechanisms and neural networks called upon—for example, in early infancy face processing recruits subcortical areas, with a shift to more cortically-based mature processing taking place (Johnson et al., 2015). Not only that, but the cortical networks of infants must undergo many years of maturation until they are functionally equivalent to those of adults (Uhlhaas et al., 2010). One way in which this can be seen is in the development of face-sensitive ERP components: while for

adults this is the N170 component, which reflects activity in the arguably face-selective fusiform area (Gao et al., 2019; Kanwisher and Yovel, 2006), in infants the equivalent components are N290 and P400, which do not show the same level of specificity for human faces early in young infants (3-months-old) as they do in older infants (12-months-old) (Halit et al., 2003). So, at a neural level, infant processing of faces is substantially different to adult processing and must undergo a major developmental change before it is functionally mature.

At a behavioural level, infant acquisition of face-related processing skills can be characterised by its speed and dependence on the environmental input. While we will examine the role of the environment in greater detail in a later section, the general principle of rapid and environmentally dependent development can be seen most clearly in infants with congenital cataracts. Maurer et al. (1999) followed the visual acuity of infants treated for cataracts at ages ranging from 1 week to 9-months. While immediately after treatment visual acuity was comparable to that of neonates, the rate of improvement with visual input was rapid, and within as little as one hour of visual input there was improvement. This demonstrates the enhanced plasticity of skills and the enhanced speed with which they can be acquired in infancy and early childhood, with early life constituting a sensitive period wherein the environmental input can result in great changes in proficiency. This unique pace of learning means that the only way in which we can ascertain how and in what manner these rapid changes unfold and what specific role the environment may play is through direct examination of infancy.

There is further evidence from infants with cataracts that reinforces this idea that disruption of the environmental input is crucial, especially for processing of faces and socioemotional input. Even relatively brief visual deprivation in infancy results in notable later differences in socio-emotional processing: infants born with bilateral congenital cataracts, which precludes visual exposure to faces in the early weeks of infancy, go on to show impairments in face identity recognition when aspects such as view orientation or facial expression vary (Geldart et al., 2002). While this is subtle, it is important, and we can see the behavioural differences reflected in imaging data, with those who were born with congenital cataracts showing reduced activation in components of the face network, and these neural differences correlating with their face discrimination ability (Grady et al., 2014). This demonstrates both the significant role of environmental input and the importance of infancy as a critical period for socioemotional development.

We can further understand the importance of the visual face input itself in early life through examining the effect of selective deprivation or variation in visual exposure on the qualitative nature of face processing. As an extreme example, infant monkeys deprived of exposure to faces during this apparent sensitive period and then exposed to either human or monkey faces only, show notable impairments in processing the category of faces to which they were not initially exposed (Sugita, 2008). This is analogous to less severe variations in visual exposure to different categories of faces present naturally for human infants. While typically, infants are exposed only to human faces and lose the ability to discriminate between individual primate faces as effectively as human faces by 9-months, exposure to primate faces in human infants

between 6-months and 9-months-old facilitates maintenance of the ability to discriminate this class of faces (Pascalis et al., 2005). This evidence suggests the existence of a sensitive period for learning about faces during which changes in input have a great impact on later proficiency.

The speed of development in infancy, as well as the formative nature of early input, highlights the importance of understanding this period and the environment. Given that even limited or brief deprivation or changes in input result in marked later function differences, small early differences in learning and input must have a disproportionate knock-on effect on later development, indicating a cascade-like process of face processing development. That is to say, changes in input earlier on result in a domino effect, with later stages heavily dependent on those that came earlier, and later input insufficient to fully compensate for early poverty of the environment. This idea of development following a cascade-like process has been explored and called upon in many domains of infant and toddler development previously, including socioemotional development (Tamis-LeMonda and Lockman, 2023). The importance and qualitatively unique nature of this period for all development, but particularly face development contingent on the environment, would therefore warrant direct investigation.

Characterising The Emergence of Face Processing in Infancy and Toddlerhood.

To understand what role the environment may play in development, we must understand the nature of this development. Extensive research has examined specifically how infants and young children respond to, understand and process faces. While an exhaustive review of this literature

would be beyond the scope of the thesis, here we will outline the major features of face processing development in infancy and childhood.

Face Processing at Birth.

A fundamental set of findings in face processing development is that from birth, infants show a looking preference for face-like stimuli, can discriminate between primate and human faces, and exhibit a looking preference for their mothers' faces (Mondloch et al., 1999; Bushneil et al., 1989). These seemingly simple abilities are highly specific, and there are many caveats to them.

Firstly, the apparent innate face-looking bias may be better characterised as a simpler phenomenon driven by perceptual properties, with infants simply showing a preference for "top-heavy" arrays (those with more elements in the upper half than the lower). This top-heavy bias is potentially driven by visual field sensitivity (Simion et al., 1998). Additionally, the ability of neonates to recognise their mothers may not reflect complex face-processing abilities. When external cues from outside of the face itself are removed (e.g., when the mother wears a headscarf to hide the hairline and head), neonates are unable to distinguish their mother from a stranger (Pascalis et al., 1995). This suggests that early "identity recognition" may in fact be driven by non-face elements. Face processing is also not species specific. While neonates can discriminate between human and monkey faces, not only do they not show spontaneous preference for either category, but they also show a preference for upright faces for both species (Di Giorgio et al., 2012).

A spontaneous preference for human adult faces over monkey adult faces emerges in infants aged 1-month and older, and this preference is only observed for infant human and monkey faces by 3-months-old (Sanefuji et al., 2014). This indicates that the proficiency and specificity of processing for and to human faces, a marker of mature and perceptually narrowed face processing, is not present from birth (Pascalis et al., 2002; Scott and Fava, 2013). Given the dependency of this perceptual narrowing on visual input, this aspect of neonate face processing highlights the significant learning from the environment yet to take place. Early face processing is thus not only less proficient than mature face processing but is less specialised for frequently seen categories of faces.

Face Processing Development Between 3- and 6-months.

In terms of specific developmental milestones, by age 3-months, many developments in face processing have taken place, extending beyond the improvements in visual acuity that are taking place (Cavallini et al., 2002). For example, at this age, infants show a preference for own-race faces, though not a processing advantage for them (Kelly et al., 2005). Additionally, a general preference for own-species faces emerges regardless of face age group-- prior to this point, infants only prefer own-species adult faces (Sanefuji et al., 2014).

Further, the genders represented in the visual input influence face processing at this age. For instance, 3-month-olds with female primary caregivers prefer female faces and can distinguish

face identities of female, but not male, faces (Quinn et al., 2002). From at least 3-months, if not earlier, infants show this preference and advantage for processing of female faces, which is seemingly underpinned by more extensive environmental experience with them (Ramsey et al., 2005; Sugden and Moulson, 2019). This suggests that by 3-months, processing of faces of different apparent genders has also been heavily shaped by the environmental visual input. It also suggests that the ability to perform equivalent complex processes for male faces is a marker of more mature processing, as infants initially demonstrate specific face processing skills only for more familiar categories of faces.

A milestone at 3-months is the emergence of holistic processing of faces, as opposed to featural processing. For example, Turati et al. (2010) demonstrated the composite effect for 3-month-olds and adults, but not in newborns. The composite effect occurs when the recognition of a partial face is impaired by alignment with the second half of another face, as this alignment induces holistic processing in those who have developed this feature of face processing (Rossion, 2013). Furthermore, there is evidence for this holistic processing of faces at 3-months from morphing research, with infants being more sensitive to configural changes than changes to individual features, especially when these configural changes occur in the upper half of the face (Quinn and Tanaka, 2009).

By 3-months, infants can recognise prototypical representations of groups of faces (de Haan et al., 2001). The ability to encode and recognize prototypes indicates that infants' responses to

faces are already being shaped by the types of faces they are exposed to. This ability may be essential for further milestones in infant development, allowing for recognition of categories of face.

Face Processing Development Between 6- and 9-months.

Further similarities to adult face processing emerge as the first year of life progresses. By 6-months, differences in visual scanning patterns for own- and other-race faces become apparent (Xiao et al., 2013), reflecting changes in behavioural engagement with familiar and unfamiliar categories of faces. This is reflected in the onset of perceptual narrowing in face identity processing to the ethnicities to which infants are predominantly exposed (Kelly, Quinn and Pascalis, 2007). While there is evidence for this other-race effect (ORE) from 6-months, it is not reliably evidenced before 9-months (Anzures et al., 2013). Thus, the ORE in identity processing, a characteristic of adult face processing, is emergent but not entirely established.

Between 6- and 9-months, similar perceptual narrowing for species is ongoing but plastic. Typically, the ability to discriminate between individual identities of primate faces as proficiently as between different identities of human faces is lost around 9-months. However, training with primate faces between 6- and 9-months for only a few minutes a day for three months is sufficient to facilitate retention of this skill (Pascalis et al., 2005). Therefore, at 6-months, the other-species effect is also emergent but not established. This indicates that the face prototype that infants develop is undergoing continual refinement over this period, and that this is directly

contingent on the environmental input. Broadly, this reflects the beginnings of the highly specialised face processing seen in adults.

Face Processing Development Between 9- and 12-months.

By 9-months, the ORE for face identity processing is well-established, with proficient recognition of faces restricted to own-race faces (Anzures et al., 2010; Kelly et al., 2007). This period also marks the establishment of species-specific face processing, with the ability to adeptly process primate face identity diminished (Pascalis, De Haan and Nelson, 2002). These developments highlight the substantial influence of the environment on face processing specialisation, even in only the first year of life.

At this timepoint, infants are also able to reliably distinguish between male and female faces, although this can still be confounded by differences in clothing and hair (Leinbach and Fagot, 1993). Additionally, the advantage and preference for female faces has attenuated by 9-months (Liu et al., 2015), indicating that infants' face processing skills are beginning to generalise beyond the typically familiar category of female faces to male faces.

By 10-months, more robustly holistic processing of faces is established. For example, 10-month-olds, but not 6-month-olds, process both the eyes and mouth of the face holistically (Schwarzer

et al., 2007). This demonstrates that features of mature face processing somewhat evident earlier are becoming more firmly established.

Face Processing Development Between 12- and 24-months.

By 12-months, infants' neural responses to faces become more qualitatively comparable to those of adults. For example, the infant face-sensitive ERP components (N290 and P400) show the same level of specificity to human faces as the adult face-sensitive ERP component (N170) (Halit, de Haan and Johnson, 2003). While a face-sensitive N170 does not emerge until much later in development (around 4-years) (Kuefner et al., 2010), this finding indicates that neural specialisation is already emergent by 12-months. Moreover, the speed and efficiency characteristic of adult face processing is developing; 12-months-olds can process faces 60% faster than 7-month-olds, using more adult-like patterns of attention and in a manner less dependent on face viewing angle (Rose et al., 2002).

From here, face processing ability continues to rapidly improve, with face processing possible even when visual stimuli are challenging. For example, 18-month-olds, but not 12-month-olds, show an upright face preference for visually degraded "Mooney faces" (Doi et al., 2009). This demonstrates that by 18-months, faces can be recognised as such even with incomplete visual information at 18-months, indicating rapid developmental change continuing into the second year of life.

An important milestone in face processing between 12- and 24-months (typically around 18-months) is explicit mirror self-recognition and visual preference for the self (Nielsen et al., 2003). Even 18-month-olds who have not quite reached this behavioural milestone exhibit stronger face-selective responses for own-face presentations (Stapel et al., 2017). This demonstrates two key elements of face processing development: first, recognition of the unique class of stimulus that is the self through face; and second, the emergence of skills at the level of covert processing before overt behavioural differences.

Face Processing Development in Toddlers.

Some plasticity in the face-processing system is maintained through toddlerhood. For example, while 3-year-olds without younger siblings are generally perceptually tuned to adult human faces, introducing experience with an infant face can significantly improve performance for this category of faces; this improvement is not seen in adults exposed to the same stimuli (Cassia et al., 2009). This demonstrates that some plasticity in the face processing system persists from infancy through to toddlerhood.

The early sensitivity and tuning of the face system for familiar faces, particularly the mother's face, declines in toddlerhood. Prior to 24-months, the greatest ERP Nc and P400 face responses are observed in response to the mother's face. However, between 24- and 45-months, toddlers show equal ERP responses to mother and stranger faces. Beyond this period, up to at least 54-months, the greatest ERP face response is in response to stranger faces (Carver et al., 2003).

This may reflect increased saliency of unfamiliar faces at a time when most toddlers are forming social connections outside of the household, seeing and interacting with many unfamiliar faces. Whether this change in neural response precedes or is a consequence of these environmental changes is unclear. However, it is evident that the face processing tasks of the typical toddler are likely to be very different from those of the typical newborn, with stranger faces being far more common—most of the faces young infants see are adult female faces, usually within the family home (Sugden et al., 2014). We will discuss the environmental input and how it may vary by age in more detail in later sections.

Face Processing Development in Childhood and Beyond.

While many of the major milestones in face processing development are reached in infancy and toddlerhood, the development of these abilities and related skills continues well into childhood and even adolescence (Chung and Thomson, 1995). For example, marked improvements in face identity recognition accuracy are still taking place up to age 10-years (Diamond and Carey, 1977). Further, the neural face-selective response in the N170 and P1 components matures and evolves into adolescence, with latencies in these components from face onset reducing over developmental time (Taylor, Batty and Itier, 2004). This demonstrates the protracted nature of acquiring mature face processing skills for the typical individual.

Theories of Face Processing Development.

The developmental trajectories of specific face processing skills have informed various theoretical models of development. One widely examined theory of face processing development is the CONSPEC/CONLERN model. Based on the vast differences between simple newborn face preference and the specialised face processing seen in adults, as well as the prolonged emergence of face processing skills, Johnson et al. (2005, 2015) formulated this dual-pathway model for face processing that is supported by much evidence from early development.

Initially conceptualised by Morton and Johnson (1991), the CONSPEC/CONLERN model proposes two face-processing mechanisms. CONSPEC guides simple face preference, providing a substrate for the development of CONLERN. The simple face orientation preference of CONSPEC results in infants seeing more instances of faces and for longer periods, allowing sufficient environmental input for CONLERN to develop. CONLERN is a set of specialised cortical networks and skills beyond face preference which develop because of this exposure to faces. This model suggests that while CONSPEC guides the input from birth, CONLERN represents the mature outcome of this process.

There is indeed evidence for a pathway for rapid, low spatial frequency subcortical processing of faces even in adults, though the face-specialised neural networks in adults are mostly cortical (Johnson, 2005). Further, as mentioned, neonates do show a visual preference for face-like

stimuli (Mondloch et al., 1999). Given this together with the protracted face processing already discussed, it is therefore plausible that face processing could be conceptualised as simple face preference at birth, while later infancy and toddlerhood are characterised by the emergence of processes specialised for face processing. This model posits that some information about the structure of face-like stimuli is present from birth, while other face-processing skills require environmental input to develop. This would mean that immature face processing is qualitatively different from adult face processing and highly contingent on environmental input. We can examine evidence speaking for and against this model and its claims.

First, the two groups of cognitive mechanisms are proposed to be supported by different neural structures. The subcortical face processing pathway (CONSPEK), which is fast and attuned to low spatial frequency, is supported by regions such as the amygdala and underpins the innate infant bias towards faces. In contrast, the slower pathway (CONLERN) is supported by cortical structures, tuned to mid and high spatial frequency, and is not fully developed until after infancy. There is indeed evidence for this. Cortical regions, particularly the fusiform face area (FFA) and superior temporal sulcus (STS), are proposed to support, and perhaps specialise in, face processing for adults (Kanwisher and Yovel, 2006; Nguyen and Cunningham, 2014). Cortical activity in response to faces within the cortical face network has been observed in infants as young as two months old (Kosakowski et al., 2024). However, in infants, this network is not as face-selective as in adulthood (Deen et al., 2017; Kadosh et al., 2007). Furthermore, some elements of this network, particularly in the temporal cortex, continue to mature in terms of face selectivity well beyond childhood and into adulthood (Golarai et al., 2015).

There is also evidence for subcortical face processing being key to face processing in infancy. Infants orient to face-like patterns only when they are presented in the temporal hemifields, which are thought to feed subcortically (Framorando et al., 2017). This indicates a subcortically driven initial face preference (Simion et al., 1998). Together, this evidence supports the idea of a qualitatively different and less specialised face processing system in infants compared to the mature system in adults, as well as the separate subcortical and cortical face-processing pathways of CONSPEC and CONLERN.

Understanding and evaluating theories of face processing development, such as this, requires comprehensive quantification of the environmental face input at different stages of development. This could reveal whether environmental input guided by CONSPEC alone is sufficient to develop the skills proposed to characterise CONLERN.

Facial Emotional Expression Processing

Humans can be in some way distinguished by their social cognitions (Saxe, 2006). Social cognition is cognitively complex: animals that engage in more sophisticated social processing tend to have larger neocortical ratios (Dunbar and Shultz, 2007). Indeed, humans are distinguished from other animals through the number and kind of social bonds that they form, with one mechanism facilitating this potentially being the broader range of communicative

signals employed (Dunbar, 2012). These signals can present in many forms, from words through to facial emotional expressions (FEE) (Buck et al., 1972). Here, we will focus on FEE processing.

FEE processing, which is at least partially distinct from identity recognition (Duchaine et al., 2003), is a fundamental skill for understanding others' internal mental states. In typically developing adults, FEE are processed rapidly and serve as a rich avenue for communication, even when only processed implicitly (Frith, 2009). Humans and other great apes are distinctive in the frequency with which FEE are used, with human faces particularly adapted for conveying a wide range of messages (Kret et al., 2020). Processing of facial expressions is significant in many contexts. For example, children's behavioural responses to FEE are significantly associated with theory of mind skills (Poulin-Dubois and Hastings et al., 2018). Deficits in these skills are central in disorders from Autism Spectrum Disorder (ASD) (Lombardo et al., 2007) through to antisocial personality disorder (ASPD) (Marsden et al., 2019). So, deficits in FEE processing are associated with wider deficits in social functioning. While in some senses it is difficult to consider FEE processing in isolation—real-world FEE processing is inherently multimodal, with auditory cues such as vocal expressions shaping perception and interpretation (Grossman, 2010)—this thesis will focus on information present only in the visual environment, to the exclusion of auditory cues and information.

FEE are (controversially) proposed to give information on the internal states of conspecifics, including their emotional states (Barrett et al., 2019). Given their influence on behaviour,

emotional states are likely adaptive, motivating individuals to engage in survival-enhancing behaviours (e.g., bonding with family members) and to avoid behaviours that risk survival (e.g., approaching a snake) (Izard, 2007). However, significant disagreement exists on the precise definition of emotion and FEE. Typically, emotions are defined as responses to stimuli, which can be external or internal. Emotional states are typically transient, lasting less than 60 minutes; they consist of multiple components, ranging from the subjective feeling to action tendencies, physiological responses, appraisal, and motor activity (Scherer, 1984). Emotions can be considered at varying levels of granularity, from broad categories to specific emotional states (Shuman and Scherer, 2014). Typically, skill in FEE processing is defined as accuracy in interpreting emotional states from expressions.

To fully understand other forms of social processing, it is essential to study the emergence of FEE processing skills and factors driving this. Necessarily, perceiving and understanding the expressions and communications of others is the first stage of responding to and decoding any social situation, and therefore the stage we might expect to emerge before others. Simple visual decoding of FEE seems to be associated with higher-order social cognition. For example, individuals with ASD, who have deficits in complex social processing like theory of mind, also show impaired FEE recognition (Lozier et al., 2014; Golan et al., 2007). Understanding how FEE recognition typically develops and what drives this may provide insights into both the emergence of FEE processing deficits and higher-order social processing deficits.

Typical Adult Facial Expression Processing.

First, we must understand the developmental endpoint of FEE processing. Typically developing adult humans are highly proficient in processing the emotional valences of faces. Presented with only the eyes of a face, typically developing adults can identify the emotions expressed by the model (Vogindroukas et al., 2014). Typically developing adults even show a remarkable degree of accuracy in identifying macaque emotional expressions (Linnankoski et al., 1994). This ability to extract emotional states with limited or unfamiliar information highlights the sophisticated nature of human FEE processing.

While there are some cross-cultural differences, many emotional expressions, like the smile conveying happiness, are claimed to be culturally universal (Ekman and Keltner, 1997). Crucially, no culture lacks emotional expressions, emphasising their essential communicative function for humans, so at least this element is universal.

There are some elements of FEE processing that are culturally mediated. Cross-cultural differences in emotional expressions presentation and processing, while subtle, are important and likely the result of environmental differences. When Eastern and Western cultures are contrasted, there are substantial differences in ratings of even the six “universal” emotions (Matsumoto, 1992). Correspondingly, there are subtle differences in the ways in which these expressions are presented by those from different cultures (Marsh, Efenbein and Ambady,

2003). This is a key piece of evidence, suggesting that emotional expression related skills are somewhat contingent on the environment and experience.

Within cultures, adults do show some diversity in FEE proficiency. These differences can be typical, pathological, or largely typical but modified by a separate pathology. Perhaps the most salient examples of pathological differences in emotional expression processing can be seen in Autism Spectrum Disorder (ASD), where even adults without comorbid intellectual disability demonstrate impairments in accurately ascribing emotions to expressions (Eack, Mazefsky and Minshew, 2015). This is directly related to the deficits exhibited by this group, as socio-emotional processing impairments are central to ASD (Travis and Sigman, 1998). Sources of this diverse proficiency are unclear.

Adults also show categorical perception of facial emotional expressions, perceiving and expressing several recognizable emotions. This is facilitated by conceptual knowledge in the form of verbal labels rather than structural information on expressions, demonstrating the interplay between cognitive concepts of emotion and FEE interpretation (Fugate, Gouzoules and Barrett, 2010). Categorical perception is characterised by discrete category boundaries in continua between emotional expressions and between-category discrimination advantage. Certain categories of FEE are more salient for adults than others— for example, there is a processing advantage for and an increased sensitivity to angry faces, which are threat-related (Öhman et al., 2001). This could reflect dissociable processing for different FEE categories.

For typical adults, interpreting and understanding FEE is essential for functional cognitive and emotional empathy. FEE can be used to make inferences about internal state, and therefore the ability to decode them is associated with cognitive empathy (Baron-Cohen, 2001). Not only this, but observing FEE in others can facilitate emotional resonance, a process that underlies emotional empathy (Decety and Meyer, 2008). In adults, this can be reflected in subtle, spontaneous mimicry of others' FEE, which is proposed to be linked to the mirror neuron system (Dimberg, 1982; Sato and Yoshikawa, 2007; Baird et al., 2011) and associated with empathy, especially emotional empathy (Drimalla et al., 2019). So, typical adults use FEE to facilitate a range of other cognitive and emotional functions, including those related to empathy.

Issues and Controversies in Defining and Conceptualising Facial Emotional Expressions.

Recently, there has been some debate among Psychologists about how to conceptualise FEE. Some adhere to the “classic” or “basic emotion” approach, while others advocate for new, constructivist models. These models differ in how they believe different combinations of facial action units (AUs) are produced and perceived by humans.

Basic Emotion Theories.

Traditional models of emotional expressions, often termed “basic emotion theories”, suggest that emotional expressions are universal, prototypical, and can be almost exclusively defined with reference to internal emotional states (Ekman, 1999). Basic emotion theories follow the

lead of Darwin in conceptualising FEE as discrete categories of emotion, with features serving evolutionarily adaptive functions (e.g., wide eyes for warning others of threats and for better threat surveillance) (Darwin, 1872; Lee and Anderson, 2017). Facial movements are thought to directly represent internal states, with emotions "leaking" onto the face even when undesired (Ekman and Friesen, 1969). It is in this context that we supposedly may even see micro-expressions, which are rapid and hard-to-detect full-face emotional displays which are proposed to be of great use in decoding internal emotional states (Ekman, 2009). So, basic emotions theorists consider the relationship between internal state and facial movements to be key.

This approach assumes the existence of a few "basic" emotions with culturally universal prototypes. To illustrate this universality, Ekman, Sorenson and Friesen (1969) selected 30 cropped static photographs of individuals believed to display a single emotion based on a scoring system. These expressions were also believed to be free of the influence of display rules (social conventions around how emotions should be exhibited). Participants from five cultures labelled these images, showing high cross-cultural agreement, which the authors proposed as evidence of universal and unambiguous FEE production and perception. Despite controversy, many concepts and methods from Ekman and colleagues are still used in emotion research. For example, the idea of action units and the Facial Action Coding System (FACS) remain standard methods for recording and (more controversially) categorising facial AUs. FACS has extensive applications in facial emotion research with clear training and guidelines, with a variation produced for infants (Prince et al., 2015).

This basic emotions approach has underpinned much of the past century or more of research on FEE. Many studies have focused on posed, artificial facial postures judged against a "correct" answer based on scoring systems for static images (Russell, 2015). These experimental images are assumed to represent prototypes of facial emotional expressions appropriate for real-world FEE. This is a rather top-down theory, and in many ways its assumptions have been difficult to test because of practical and technical constraints. Up until recently, which elements of the environment we can measure and how much rich, naturalistic data it is possible to analyse have been greatly limited. However, advances in recording systems, data storage, coding, and analysis now make it more feasible to directly test some of the assumptions of the basic emotions approach.

Modern Challenges to the Idea of Basic Emotions.

Alongside advances in technology, theoretical challenges to basic emotion theories have emerged. Contemporary models shift away from the idea of FEE as univariate signals of internal state that are universal and clearly identifiable. Many contemporary models frame FEE as a context-dependent form of nonverbal communication, which exists both to signal internal states and to convey messages to others (Barrett et al., 2019). These models often hold that many aspects of FEE production and processing are learned and culture-dependent, rather than innate biological signals. Specific criticisms levelled at traditional approaches by these models include the importance of understanding intent or context in shaping judgments of FEE, which

function as social tools (Fridlung and Russell, 2006; Aviezer et al., 2017); the use of highly artificial stimuli in research; and cultural differences in production, use and understanding of FEE (Porter and Samovar, 1996; Malsert et al., 2020).

Not only this, but some theorists entirely reject the notion that FEE are mostly driven by internal affective states. For example, one meta-analysis of 55 research papers found a low co-occurrence between emotion and whole face expressions for the basic emotions (.13) and partial expressions (.23) (Durán and Fernández-Dols, 2021). However, the available research on this topic is limited, with significant variability in operationalisation of dependent variables between different studies. Not only this, but it is difficult to accurately infer or directly measure internal affective state. While directly and fairly testing the affect-expression link is difficult, it is clear that real-world adult FEE are likely to be shaped in part or whole by factors other than emotional state. For instance, display rules learned over the course of childhood influence expressive behaviour (Malatesta and Haviland, 1982). Therefore, a one-to-one link between internal state and FEE seems unlikely.

So, we can reasonably claim that FEE must be considered at a minimum of two levels: the affective (how the facial movements relate to internal emotional state) and the socio-communicative (how the facial movement is used as a commonly understood “language”) (Shariff and Tracy, 2011). For example, the action units of a smile may indicate a positive

internal state or serve as a social signal for approach, or both. This is somewhat beyond what basic emotion theories would have suggested.

The Specific Role of Context in Judging Facial Expressions.

We rarely encounter the highly controlled, context-free exemplars of FEE used in much basic emotion research. Instead, we see faces and facial gestures alongside broader contextual information, including situational context, bodily expressions and other multi-sensory cues. Some have suggested that context is essential for judging FEE accurately (Barrett et al., 2011).

Bodily posture can be an essential source of contextual information for interpreting FEE (Wenzler et al., 2016; Huxter et al., 2023). They can even be more informative than FEE cues (Aviezer et al., 2012). So, judging facial movements in isolation without information from other modalities of expression is not representative of typical conditions for judging the internal affective states of others.

Another form of contextual information is prior knowledge on the states of an FEE model (Palumbo and Jellema, 2013). This demonstrates the importance of top-down information and cognitions in informing real judgements of real-world FEE (Otten et al., 2017). Social cognition, including social cognition around FEE, is at many levels bidirectional (Hadders-Algra, 2022), and utilises prior knowledge.

This is not to say that isolated FEE judgement never occurs, nor that categorisation of facial movements alone is a function without utility. It is a word of caution: context-free categorisation of facial movements is unlikely to be diagnostic of internal affective state. This role for context, whether prior knowledge or accompanying bodily gestures, is not accounted for in traditional “basic emotion” theories. So, examining FEE processing in isolation is an endeavour that must be approached critically.

Characterising The Emergence of Facial Emotional Expression Processing.

Infants and toddlers must learn how to distinguish different FEE across a wide variety of faces and extract socially meaningful information from them. Theoretically, FEE processing would emerge later than face identity processing. According to the dual-pathway CONSPEC/CONLERN model, CONSPEC is an orienting network triggered by face-like configurations, while CONLERN processes more complex cues, influencing behaviour only after 2-months of age (Slater, 2002). So, we would expect FEE processing to depend on CONLERN and emerge more gradually.

Facial Expression Processing in Neonates.

At birth, evidence for neonates' ability to distinguish facial emotional expressions is fairly limited. Early research suggested that neonates could mimic primary emotional expressions like happiness, sadness, and surprise (Field et al., 1983; Meltzoff and Moore, 1977). However, more rigorous recent replications indicate these findings may have been due to methodological errors

(Oostenbroek et al., 2016). Studies using motor responses, such as mimicry, introduce confounds related to motor ability, so modern research focuses on data from looking behaviour to minimise these confounds.

Neonates can distinguish happy from fearful expressions, selectively attending to happy faces though they cannot discriminate between fearful and neutral faces (Farroni et al., 2007). One proposed explanation for this is that neonates primarily see happy faces in the first few days of life, although the actual frequency of different FEE in the neonate environment has not been quantified. This would indicate that even very young infants can discriminate some expressions, even if this is limited to happiness. This could be related to very young infants likely primarily seeing positive FEE.

However, this preference for happy faces is limited only to happy faces with direct gaze (Rigato et al., 2011). Again, this may indicate an important role for very early experience, as potentially there is greater exposure to both direct gaze and smiling faces in the first hours of life— though this remains to be directly tested. The happy FEE preference is highly task-dependent, also not being spontaneously present when neonates are presented with dynamic happiness and dynamic disgust (Addabbo et al., 2018).

So, neonates demonstrate a spontaneous but fragile preference for happy FEE, with limited other FEE processing abilities, as we would expect from the CONSPEC/CONLERN model.

Research on neonates as outlined in this section can be seen in Table 1.

Table 1*Research on Facial Emotion Processing Pertaining Infants Birth to 3-months-old*

Study	Methodology	Age(s) Ps	Emotions Examined	Significant Result(s)	Null Result(s)
Oostenbroek et al. (2016)	Observed mimicry of adult models	1-, 3-, 6- and 9-weeks	Happy, sad	-	No imitation of expressions
Farroni et al. (2007)	Looking preference / habituation	Neonates	Happy, fearful, neutral	Preference happy	No fearful/neutral discrimination
Rigato et al. (2011)	Looking preference faces + gaze	Neonates	Happy, fearful, neutral (direct/averted)	Preference happy + direct over neutral	No preference happy + averted over fearful
Addabbo et al. (2018)	Looking preference / habituation (Perceptually controlled)	2-days-old	Happy, disgust	Discrimination (after habituation)	No discrimination (looking preference)
Nelson and Horowitz (1983)	Looking Preference (holographic)	2-months-old	Happy/neutral	Discrimination	-
Kuchuk et al. (1986)	Looking preference	3-months-old	Different intensities smiling	Preference more intense (sensitivity dependent on home environment)	-
Brenna et al. (2012)	Habituation (identity)	3-months	Happiness, fear, anger	Happiness facilitating identity recognition	No effect of fear, anger on identity recognition
Turati et al. (2011)	Habituation (identity)	3-months	Happiness, neutral	Happiness facilitating identity recognition	-
Hoehl et al. (2008)	EEG/ERP	3-months-old	Fearful, neutral (gaze directed at / away from object)	Nc amplitude enhanced for objects paired with gaze of fearful faces	No Nc amplitude change for objects not paired with gaze of fearful faces
Hoehl and Striano (2010)	EEG/ERP	3/6/9-months-old	Fearful, neutral (gaze directed at / away from object)	6m enhanced Nc fearful faces looking at objects (and for objects); 9m enhanced Nc fearful faces and objects	No difference 3m ERP fearful/neutral (any gaze); 6m ERP fearful faces looking away from objects

Note: This table provides a summary of the findings of current research on emotional expression processing birth to 3-months; provided are the citations, methods, age groups and findings.

Facial Expression Processing at 1- to 3-Months.

There is further evidence for infant abilities to process happy FEE specifically through the first 3-months— for example, Nelson and Horowitz (1983) find that 2-month-olds can discriminate between holographic happy and neutral FEE. However, towards 3-months, infants begin to demonstrate more sophisticated and varied FEE processing abilities. There appears to be some within-category FEE discrimination at 3-months. At this time point, infants have been shown to not only exhibit preference for smiling facial expressions, but for more intense smiling expressions when compared with less intense expressions, with sensitivity to smiling dependent on maternal behaviour (Kuchuk et al., 1986). This would imply that within-category facial expression discrimination becomes possible, but its emergence may be impacted by the environmental input. It also further demonstrates the earlier emergence of processing skills related to happy and smiling FEE.

This can be further seen in the interaction between FEE and identity processing at 3-months. Infants show enhanced identity recognition of faces exhibiting happy, smiling faces, seemingly driven by salient elements of these expressions (Brenna et al., 2012; Turati et al., 2011). This

may also demonstrate that some inherent elements of happy FEE drive the processing advantage and more rapid development of processing for these FEE.

There is also evidence for more complex processing and processing beyond happy, smiling FEE by 3-months– there is evidence that some meaning may be extracted from fearful expressions by these infants. Hoehl et al. (2008) found that Nc amplitude (an indicator of infant attention) was enhanced for objects that had been associated with an adult face that was both fearful (as compared to neutral) and had its gaze directed at the object. This is significant in indicating an ability to distinguish fearful faces from neutral at 3-months-old and interpret these as social cues about the environment, but limited comparisons mean we cannot be sure if there is understanding of the meaning of a fearful expression and the implications of it (which is to say, that the object is dangerous or a threat); further, it could be that fearful faces distracted from the objects, leading to a novelty response.

Interestingly, 3-month-olds do not show differential ERP responses to fearful and neutral faces alone (Hoehl and Striano, 2010). This neural differentiation between the objects associated with facial expressions, but not the expressions themselves, could be explained in several ways, but one possibility is that 3-month-olds do distinguish them, but not in a way recorded by the EEG. For example, differentiation at this age could be more heavily dependent on subcortical structures, which would not register activity as strongly at the scalp. This would be consistent

with the concept of more subcortically based face processing early in development (CONSPEC/CONLERN) but has not been directly tested.

In sum, by 3-months, infants continue to show more sophisticated processing of happy FEE, as demonstrated by both within-category processing and facilitation of identity processing for happy faces. There is also some evidence for emergent (and possibly subcortical) extraction of meaning from fearful FEE, as 3-month-olds show enhanced EEG Nc response to objects paired with fearful faces but not the fearful faces themselves. See table 1 for a summary of research up to this point. See table 1.

Facial Expression Processing at 4- to 6-Months.

From 4-months onwards, there is evidence for more sophisticated processing of a wider range of FEE, including negative FEE. This processing is often evident in restricted forms and contexts before reaching the sophisticated general FEE processing of adults.

There is evidence that by 4-months-old, infants can discriminate fear, anger and surprise (Serrano et al., 1992) and show increased skin conductance (a measure of arousal) to angry faces as compared to happy faces (Nava et al., 2016). Further, at 4-months, as at earlier time points, there is evidence that there is some meaningful processing of negative FEE based on gaze direction. Striano et al. (2006) compared 4-month-old infants ERP responses to faces that

differed in emotion (neutral, happy or angry) and in gaze direction (direct or averted). Nc, a metric of attention, was enhanced for angry faces with direct versus averted gaze, which would be a more meaningful threat signal.

As infants get older, their ability to distinguish different facial emotional expressions appears to become more sophisticated, with spontaneous preference for negative FEE becoming apparent. In contrast to 3-month-olds, 6-month-olds show differential ERP responses to fearful and neutral faces themselves, with an enhanced ERP response to fearful faces, though only when these faces are looking towards objects as opposed to empty space (Hoehl and Striano, 2010). These developments are somewhat experience-dependent. For example, infants at 6-months-old show spontaneous preference for fearful faces only in own-race faces (which they are exposed to often), but not in other-race faces (which they are exposed to less often), implying that experience with faces shapes emotional expression processing in the same way that it influences face identity processing (Safar et al., 2017). This is the beginning of an FEE ORE as seen in adults.

Infants at this age have also been shown to discriminate between neutral expressions and different intensities of happiness and sadness, from subtle to pronounced (Striano et al., 2002). This within-category discrimination for emotions beyond happiness may indicate development towards more mature emotional expression processing, especially for negative emotional

expressions. Again, this shows how skills that have already emerged for happiness emerge for other FEE, particularly negative FEE, later in development.

In sum, between 4- and 6-months, there are many FEE processing developments especially for negative FEE; further, perceptual narrowing is taking place for ethnicity, with an FEE ORE beginning to emerge. In general, the influence of elements of the environment on FEE processing development is becoming clear. See Table 2 for a summary of research between 4- and 6-months.

Table 2

Research on Facial Emotion Processing Pertaining Infants 4- to 6-months-old

Study	Methodology	Age(s) Ps	Emotions Examined	Significant Result(s)	Null Result(s)
Nava et al. (2016)	Skin conductance	3- to 4-months-old	Happy, angry (subliminal, supraliminal)	Skin conductance higher for angry	No effect subliminal vs supraliminal
Serrano et al. (1992)	Looking preference / habituation	4-months-old	Fear, anger, surprise	Preference anger and surprise; discrimination all (habituation)	No discrimination of fear after habituation surprise
Striano et al. (2006)	EEG/ERP	4-months-old	Neutral, happy, angry (direct/averted)	Nc enhanced for angry faces direct (over averted)	No difference in direct/averted Nc for neutral/happy
Safar et al. (2017)	Looking preference	6/9-months-old	Happy, fear (White / East Asian)	Preference fear 6m (own race); preference fear 9m	No preference for fear 6m (other-race)
Hoehl and Striano (2010)	EEG/ERP	3/6/9-months-old	Fearful, neutral (gaze directed at / away from object)	6m enhanced Nc fearful faces looking at objects (and objects); 9m enhanced Nc	3m no difference in ERP response for fearful/neutral (any gaze); 6m no enhanced Nc

Study	Methodology	Age(s) Ps	Emotions Examined	Significant Result(s)	Null Result(s)
				fearful faces and objects	for fearful faces looking away from objects
Striano et al. (2002)	Looking Preference	6-months-old	Neutral, smiling, frowning (varying intensities)	Discrimination varied intensities from neutral (preference ~ maternal depression)	-

Note: This table provides a summary of the findings of current research on emotional expression processing 4- to 6-months; provided are the citations, methods, age groups and findings.

Facial Expression Processing at 7- to 9-Months.

New skills continue to emerge as existing skills generalise to new categories of FEE. As at 6-months, there is extension of within-category FEE discrimination at 7-months. Beyond happiness and sadness, there is evidence that 7-month-old infants can discriminate between mild and extreme fear expressions (Ludemann and Nelson, 1988). This would demonstrate that finer-grain elements of negative FEE are becoming easier to recognise and discriminate from 7-months.

In this paper, Ludemann and Nelson (1988) also found some evidence for categorical perception of happiness and surprise. This categorical perception is a step towards adult-like categorical

processing. In examining infant responses to happy and fearful expressions, Nelson et al. (1979) demonstrated that, with sufficient exposure to expressions across different models, infants can recognise a new model posing the same facial expression as familiar. However, the order in which stimuli were presented affected this ability. This would suggest that an understanding of FEE as categories of pose that can be recognized regardless of the specific model is emergent, if not yet established.

In addition to generalisation of expressions to new models, adult-like categorical processing would entail clear category boundaries along expression continua with cross-category processing advantage. Indeed, further research using a continuum of faces from happy to fearful in this age group has demonstrated categorical discrimination between facial expressions, with infants showing reliable category boundaries between the two expressions and discrimination between expressions from either side of them— but, this effect was only observed when infants were habituated to happy faces, not fearful faces (Kotsoni et al., 2001). One explanation proposed by the authors for this effect not being demonstrated after habituation to fearful faces draws upon the idea of happy expressions being more familiar to the infants and the associated processing advantage for these faces. Overall, this is further evidence for an emerging but not established categorical perception of FEE.

The attentional bias for fearful faces becomes less context-specific and more established. 7-month-olds exhibit an enhanced Nc for angry faces with direct gaze, with gaze direction not

mediating the Nc response to fearful faces (Hoehl and Striano, 2008). There are similar findings in Hoehl and Striano (2010), with 9-month-olds (but not 6-month-olds) demonstrating enhanced Nc for fearful faces regardless of gaze direction. This could suggest a shift to cortical processing of emotional expressions, as well as more mature processing of negative emotional expressions. This could emerge as a result of more exposure to them at this time point, as infants begin to navigate the world independently.

Infants at this time exhibit behavioural threat biases that may reflect innate processes, with 7-month-olds demonstrating a preference for looking at fearful faces but not happy or novel expressions (Peltola et al., 2008). While previous authors have focused on the role of experience with faces, proposing that this response to fear is a novelty preference stemming from lack of experience with fearful faces, Peltola et al. (2008) finds evidence that the threat bias may be experience independent. In comparing responses to fearful faces to happiness and artificial novel expressions, each paired with a control, it was found that infants showed a looking preference and reduced ability to disengage attention only for fearful faces. Had this threat bias been driven by the novelty of fearful expressions, the same effect should have been observed for the novel facial expressions. So, the increased saliency of fearful faces at 7-months at least is driven by factors other than novelty.

In sum, at this age, more sophisticated within-category processing abilities begin to emerge as clear evidence of adult-like categorical processing of FEE becomes apparent. A behavioural and

physiological bias for processing threat-related FEE like fear and anger becomes more established and less context-dependent. There is some evidence to suggest that this bias cannot be wholly explained by novelty preference. See Table 3 for a summary of research between 7- and 9-months.

Table 3*Research on Facial Emotion Processing Pertaining Infants 7- to 9-months-old*

Study	Methodology	Age(s) Ps	Emotions Examined	Significant Result(s)	Null Result(s)
Ludemann and Nelson (1988)	Habituation	7-months-old	Happy, fear, surprise (mild/extreme)	Discrimination mild/extreme happy and fear; categorical perception happy and surprise	No categorical perception fear, or discrimination from happy/surprise; no dishabituation surprise/happy
Peltola et al. (2008)	Looking preference	7-months-old	Fear, happy, novel, control (scrambled)	Preference fear over control; disengaged less from fear	No preference happy/novel over control; no novelty bias
Hoehl and Striano (2008)	EEG/ERP	7-months-old	Fear, angry (direct/averted)	N290, P400, Nc sensitive to gaze and emotion	-
Nelson et al. (1979)	Habituation	7-months-old	Happy, fear	Discrimination of happy/fear across different models	-
Kotsoni et al. (2001)	Looking preference / habituation	7-months-old	Fear, happiness (continua)	Categorical perception; preference fear	-
Hoehl and Striano (2010)	EEG/ERP	3/6/9-months-old	Fearful, neutral (gaze directed at / away from object)	6m enhanced Nc fearful faces looking at objects (and objects); 9m enhanced Nc fearful faces and objects	No 3m ERP fearful/neutral discrimination (any gaze); 6m no ERP for fearful faces looking away from objects

Note: This table provides a summary of the findings of current research on emotional expression processing 7- to 9-months; provided are the citations, methods, age groups and findings.

Facial Expression Processing at 10- to 14-Months.

Many developmental changes take place between 10-months and 14-months, but the pace of change is especially rapid up to toddlerhood. To establish typical adult-like FEE processing, infants and toddlers must do more than perceptually discriminate between a handful of emotion categories. They must disentangle subtle and small differences between facial poses and movements. They must also actually extract information from the facial poses and movements that they see and use this information to enhance their navigation of the environment, including the social elements of the environment.

By 12-month-olds, there is a more adult-like enhanced sensitivity to angry faces with occipital activation in response to them. Grossman et al. (2007) compared to responses of 7- and 12-month-old infants to happy and angry expressions. While in 7-month-olds there is a larger ERP response to happy faces (at the frontal, central temporal and parietal sites), in 12-month-olds negativity was greater for angry faces. The response of 12-month-olds is more like that observed in adults, and the results suggest that development of this processing entails a shift in the ways in which the same FEE are processed over time to reflect adult patterns of processing.

A further important component of interpreting emotional expressions is resonating with them (Decety and Meyer, 2008). This can allow for a range of adaptive social functions, including emotional empathy. In adults, one element or sign of resonating with an emotion can be subtle mimicry; there is subtle and spontaneous mimicry of the emotional expressions of others in adults, potentially a mechanism related to the proposed mirror neuron system (Dimberg, 1982; Sato and Yoshikawa, 2007; Baird et al., 2011). Degree of mimicry is associated with both cognitive and emotional empathy (Drimalla et al., 2019). There is some evidence of this phenomenon in infants ages 9- to -10-months, with infants at this age demonstrating lip-corner raising in response to happy faces and brow-lowering in response to angry faces (Hashiya et al., 2019). This is indicative not only of some ability to differentiate these expressions but is also potentially the foundations for emotional resonance with others' expressions through extraction of social information from FEE.

The ability to extract information about the cognitive states of others from FEE is also emergent in the first year of life. There is evidence to demonstrate that 12-month-olds can connect information from gaze and emotional expressions to intentional action, and that this ability is continuing to develop at 14-months. Phillips, Wellman and Spelke (2002) found that 12-month-olds dishabituated to events in which gaze, expression and subsequent action were not consistent. The infants were habituated to models positively regarding one of two toys and then shown holding one. However, this was not an established ability. When habituated to only the positive regard and not the action, 12-month-olds no longer exhibited surprise to trials wherein the toy which was regarded positively and the toy held were not the same. By 14-months, there

was surprise towards these inconsistent trials regardless. This demonstrates an emergent understanding of the relationship between expression, gaze and action at 12-months that is further developed by 14-months. So, there is evidence for inferring mental states and intention in the first year of life, but this ability is still developing.

At 12-months, infants demonstrate an understanding of emotional expressions as a form of communication about the environment and modulate their behaviour accordingly. Sorce et al. (1985) demonstrated that 12-month-olds use the facial emotional expressions of their mother to determine the correct action in an ambiguous situation (the visual cliff), with expressions of joy or interest leading most infants to cross the “cliff” and fear or anger leading few to do so. This modulation of behaviour was selectively observed in ambiguous situations, demonstrating that 12-month-olds can recruit information from the emotional expressions of others and change their response accordingly only when this is necessary for navigating the environment. This use of social referencing from 12-months has been replicated for ambiguous objects (Kim and Kwak, 2011). Social referencing using FEE is adult-like in the sense that emotional expressions are utilised as an avenue of communication.

In sum, from 10- to 14-months, developmental milestones extend well beyond simply distinguishing FEE and through to combining information from FEE with other sources of information to formulate and act upon hypotheses about other social actors and the environment itself. As verbal skills emerge, implicit research using physiological measures and looking behaviour is no longer the only way to assess FEE understanding. Development of

socioemotional skills around FEE processing continues well into childhood and adolescence. See Table 4 for a summary of research presented here.

Table 4

Research on Facial Emotion Processing Pertaining to Infants from 10- to 24-months

Study	Methodology	Age(s) Ps	Emotions Examined	Findings	Null Result(s)
Hashiya et al., 2019	Mimicry	9- to 10-months	Happy, angry	Lip-corner raising when observing happy, brow-lowering when observing angry	-
Grossman et al. (2007)	EEG/ERP	7/12-months	Happy, angry	7m larger ERP response happy; 12m greater negativity angry	-
Sorce et al. (1985)	Visual cliff	12-months	Joy, interest, fear, Anger	Most crossed if mother posed joy/interest; most did not for anger/fear	No significant effect emotion when no depth to cliff
Kim and Kwak (2011)	Positive, negative and ambiguous toys, happy/fearful adult	12-months; 16-months	Happy, fear	More frequent and faster looking to adults for ambiguous toys; regulation based on FEE only for these	-
Phillips et al. (2012)	Habituation	12-months; 14-months	Positive	Dishabituate to trials wherein gaze, affect and action are not consistent after habituated to event	12-months do not dishabituate after habituated to gaze and affect alone
Chiarella and Poulin-Dubois (2013)	Behavioural response to justified and unjustified emotions	15-months; 18-months	Negative	18-month-olds respond differently to justified and unjustified	15-month-olds do not respond differently

Study	Methodology	Age(s) Ps	Emotions Examined	Findings	Null Result(s)
Scott (2017)	Behavioural response to actors' false beliefs being disproved	20-months	Surprise	Expected actor to express surprise	-

Note: This table provides a summary of the findings of current research on emotional expression processing from 10- to 24-months; provided are the citations, methods, age groups and findings.

Facial Expression Processing at 15- to 24-Months.

By 18-months, infants appear to use information from the environment to assess whether emotional expressions are justified. This necessitates relatively complex skills in terms of both cognitive and emotional empathy using information from FEE. Chiarella and Poulin-Dubois (2013) found 18-month-olds, but not 15-month-olds, reacted differently to negative emotional expressions following positive (unjustified) and negative (justified) events, with more concerned reactions in the justified condition and more hypothesis testing (e.g., looking between the object and the face) in the unjustified group, demonstrating that by 14-months a more sophisticated understanding of what internal experiences facial expressions represent may be present. This would require an ability to understand the links between expressions, feelings and events, as well as the formulation of hypotheses around which events should be associated with which emotions.

In keeping with the theme of older infants exhibiting understanding of the relationship between the world, actors' internal states and actors' FEE, 20-month-olds expect an actor interacting with objects to exhibit surprise when a false belief of the actor is disproven (Scott, 2017). This demonstrates an understanding that others can have false beliefs, as well as an understanding that internal factors like beliefs can elicit emotions. It demonstrates that information can be extracted from FEE and compared to internal hypotheses around how actors should respond to and assess events in the environment emotionally. See Table 4 for a summary of research presented here.

Facial Expression Processing in Toddlers and Children.

While at earlier time points implicit measures must be used to examine FEE processing development, as infants become toddlers and begin to communicate verbally, more explicit measures can be used. An adult-like categorical perception of emotional expressions can be explicitly established using verbal report in toddlers, having been observed for happy and sad faces in 3.5-year-olds (Cheal and Rutherford, 2011); however, even at this time point, responses of toddlers are somewhat heterogeneous when compared to the responses of adults, and responses tended to be less strongly categorical in the toddlers than in the adults. This would suggest that while many aspects of adult-like emotional expression processing are present by the time children start preschool, this processing is still subtly different from adult processing and there are still some developmental milestones to be reached.

Many changes in FEE processing continue to take place beyond toddlerhood into childhood and even adolescence (Batty and Taylor, 2006; Herba and Phillips, 2004). While it would be unnecessary to outline all these changes here, generally they pertain to fine-grain qualitative differences in FEE processing such as differences in the strength and timing of ERP components associated with processing different FEE. However, there are also general improvements in the speed and accuracy of processing that can be observed through to adulthood. See Table 5 for a summary of research presented here on toddlers and children.

Table 5*Research on Facial Emotion Processing Pertaining to Toddlers and Children*

Study	Methodology	Age(s) Ps	Emotions Examined	Findings	Null Result(s)
Cheal and Rutherford (2011)	Discrimination, labelling, morphed continua	3.5 years	Happy, sad	Category boundaries reflect adult category boundaries; enhanced discrimination between boundaries	-
Kammermeier et al. (2019)	Eye tracking, attachment measures	5-year-olds	Neutral, anger, fear, sad and happy	Attachment security predicts visual attention to FEE; securely attached children look longer at fear	-
Batty and Taylor (2006)	EEG, implicit processing	4- to 15-years	Six basic emotions + neutral	Age-related changes in several FEE-related ERPs	-
Herba and Phillips (2004)	Literature review	Childhood and adolescence	Six basic emotions + neutral	Continued behavioural and physiological changes into	-

Study	Methodology	Age(s) Ps	Emotions Examined	Findings	Null Result(s)
				adolescence, especially as related to accuracy and speed	

Note: This table provides a summary of the findings of current research on emotional expression processing in toddlers and children; provided are the citations, methods, age groups and findings.

Summary of Facial Expression Processing Development.

To summarise, infants develop from the neonatal ability to discriminate and preference for happy faces to an ability to discriminate and utilise information from a wide range of facial expressions that is like that of adults, but not fully adult, by preschool. The ability to discriminate and a preference for fearful faces begins to emerge around 6-months, and adult-like processing of anger is recorded at around 12-months, despite a preference for and enhanced attention to it emerging much earlier on (around 4-months). Categorical perception is observed to some degree from around 7-months, but still emergent into preschool. The ORE for FEE begins to emerge at around 6-months. Many authors posit a role for the infant visual environment in shaping these developmental changes, particularly those centred around negative emotional expressions. The suggestion is that younger infants are frequently exposed to positive FEE like happiness, but are not exposed to negative, threat-related FEE like fear and anger, which may make them novel. However, older, mobile infants may be exposed to expressions like fear and anger as they explore the environment and commit transgressions.

This could explain the earlier emergence of skills related to emotions like happiness and the later emergence of skills related to these negative emotional expressions. However, up to this point, the infant visual environment with respect to emotional expressions has not been quantified, so these ideas remain suggestions. That said, the ORE that we see develop for FEE would suggest some role for the environment in shaping FEE development.

Body Processing

We will briefly consider adult processing of bodies and how body-processing skills may emerge over infancy and toddlerhood. While not the primary focus of this thesis, and we will not extract information on what emotional cues are available for toddlers within body presentations, we will discuss the frequency of body presentations within the infant visual field in papers 1 and 2 for the purposes of frequency comparison. For this reason, it is necessary to briefly outline what purpose observing bodies may serve for human adults and infants, as well as characteristics of body processing.

Typical Adult Body Processing.

Human bodies are an important class of stimulus for adult humans: they are frequently present and a rich source of information. There are many similarities between the factors that make faces “special” and attributes of bodies— both are an important source of information on the self and others (Slaughter et al., 2004). This is reflected in the ability of adults to extract

information on emotional states from body movements, whether in the form of recordings or merely point-light displays (Dittrich et al., 1996).

Adult Recognition of Bodies.

As with faces, there is an apparently selective neural locus for non-motion, perceptual processing of bodies in the lateral occipitotemporal cortex, analogous to the FFA (Downing et al., 2001). Like faces, bodies are processed somewhat holistically by adults (Butti et al., 2002; Willems et al., 2014). This is reflected in a similar inversion effect in bodies as that observed in faces (Reed et al., 2003).

The existence of a specialised body recognition “module”, analogous to that seen in faces, is also evidenced by the existence of autotopagnosia, which is a disorder defined by a selective inability to recognise the individual body parts. This deficit has been associated with everything from temporal brain lesions to Alzheimer's (Ogden, 1985; Sirigu et al., 1991). While less commonly recorded, this seems to be a deficit similar to prosopagnosia and demonstrates the existence of a distinct and dissociable body processing skill.

Processing of Motion and Posture Information from Bodies.

Adults seem to utilise information from both faces and bodies to assess emotional expressions, with mismatch between the two confounding assessments of facial emotional expressions (Meeren et al., 2005), meaning that bodies are a valuable source of socioemotional information. Not only may bodies add to assessments of emotional state, they may in some instances be

decisive: for peak positive and negative emotional expressions, Aviezer et al. (2012) established more ambiguity for faces than bodies, and that decisions on the emotional context (positive or negative) for composite images of faces and bodies were primarily contingent on body posture; this demonstrates that holistic, real-world decisions on emotional state can, at least in some contexts, depend primarily on bodily posture.

Adults have neural and cognitive systems highly proficient in extracting information from bodily movements and postures. This is reflected in specialised visual mechanisms in the superior temporal sulcus for biological motion (that is, motion resembling that of a body) (Puce and Perrett, 2003).

Further evidence for the existence of contexts in which bodies may be a primary source of emotional state information over faces can be seen in Pollux et al. 2019. This paper demonstrates a tendency for adults to show a strong bias for looking towards the face for extracting emotional state when presented with dynamic displays, but a bias for upper body regions when the face is not visible or merely static, with differences in specific pattern by emotional expression (for example, for static displays, there was a preference for hands and arms for happy and fearful displays). This again demonstrates dependence on bodily cues for extracting emotional states in some contexts, here specifically when the image is static, or the face is occluded.

Characterising The Emergence of Body Processing in Infancy and Toddlerhood.

Bodies are also common stimuli in the environments of infants. However, there are many differences in the timeline for body processing development. Further, there is far less evidence for specific developmental milestones in body processing. Here, we will outline some major elements of body processing development.

Body Processing at Birth to 6-Months.

Even neonates and very young infants exhibit a visual preference for biological over nonbiological patterns of motion (Fox and McDaniel, 1982; Bidet-Ildei et al., 2014). Infants at 3-months-old can discriminate walking and running motions in point-light displays, but respond primarily to motions within a single limb, whereas by 5-months infants interpret movements on the basis of relationships between limbs (Booth et al., 2002); 5-month-olds also demonstrate some evidence of the holistic body processing of adulthood (Hock et al., 2015). Together, this demonstrates that bodily motion processing methods change and become somewhat more mature by 5-months, with the body processed as a whole and movements interpreted on the basis of this.

In terms of early visual specialisation for processing of bodies, this is less pronounced and emerges later for bodies than for faces. For example, at 3-months, only Nc is observed as a marker of specialised cortical processing of bodies, and this Nc is smaller for bodies than it is for faces (Gillmeister et al., 2019).

By 6-months, infants show understanding of the actions of adults being directed towards goals in the environment, demonstrating dishabituation to new targets but not to new directions of movement (Woodward, 1998), showing that an understanding of intentionally bodily action is emerging. Not only this, but intentional action is understood within the context of bodily biomechanics: 6-month-olds, but not 4-month-olds, can discriminate between hands grasp movements that are biomechanically possible and impossible, and an emergent tendency to look selectively towards the goal of a possible action (Geangu et al., 2015). This relatively sophisticated understanding of hand movements in the context of intentional action and what is physiologically possible may be specific to hands, as infants at 9-months are unable to discriminate between possible and impossible movements for a physiologically closely related body part, the elbow (Morita et al., 2012).

Body Processing at 7- to 12-Months.

At 7-months, infants demonstrate a preference for looking at the upper body as compared to the lower body and demonstrate differing fixation patterns depending on the emotional “expression” conveyed by body posture: for example, when presented with fear, infants showed an increased preference for the upper body (Geangu and Vuong, 2020). This demonstrates that 7-month-olds, like adults, demonstrate gaze behaviour for bodies contingent on emotional expression; this may suggest not only some discrimination but attention selectiveness for areas that are most information-rich, as the upper body contains more of the emotional cues.

At 8-months, as with adults, presentation of incongruent bodily emotional expressions can hinder infant identification of facial emotional expressions (Rajhans et al., 2016), suggesting both discrimination of bodily emotional expression and some knowledge of which facial emotional expressions “match” or convey the same state.

While an equivalent ability does not emerge for images of bodies until much later in developmental time, there is some evidence that by 9-months, infants are able to discriminate intact from scrambled bodies when presented as real human bodies (Heron and Slaughter, 2010). This demonstrates later emergence of these skills than the equivalent for faces, and demonstrates that for body processing, whether these skills can be elicited is highly contingent on the specific task elements.

By 12-months, like adults, infants can discriminate between possible and impossible arm movements, looking longer at the elbows when presented with physiologically impossible body movements (Morita et al., 2012); however, unlike adults, there was no difference in emotional response when 12-month-olds were presented with the possible and impossible body movements. So, while this demonstrates an emergent understanding of biomechanics that extends beyond the hand, this ability may not yet be fully mature by 12-months of age.

Body Processing at 13- to 24-Months.

While there are many early developments in understanding of bodily motion and posture, some elements of body processing emerge much later in development than their corresponding face processing abilities. For example, in contrast to early recognition of and preference for face-like stimuli as a class of stimulus distinct from others, infants do not discriminate between images of intact and scrambled bodies until between 15- and 18-months (Slaughter et al., 2004).

Further, they do not show any preference between intact and scrambled human bodies until 18-months (Slaughter et al., 2002). Not only this, while neonates show preference for intact over scrambled faces, when a looking preference for bodies emerges at this much later age, infants prefer scrambled bodies. Presumably, this reflects a novelty preference, as bodies should be common in the environments of most infants. However, it also reflects fundamental differences between the emergence of face and body processing, with equivalent perceptual skills emerging much later for bodies.

Further evidence for the comparatively slower emergence of body processing can be seen in the inversion effect. This effect for bodies does not emerge until around 14-months, when differential neural responses to upright and inverted bodies begin to present (Gillmeister et al., 2019). Again, this is much later than the equivalent effect for faces, which emerges by 4-months (Turati et al., 2004). This demonstrates that the developmental timeline for body visual processing development is very different to that of faces, and in many ways is much slower.

Temperamental Differences and Socioemotional Development

One important variable in shaping infant responses to socioemotional stimuli is their disposition or temperament (Martinos et al., 2012). Temperament can be separated into finer-grain factors. For example, there exist two very well-validated measures of temperament, the IBQR (for <18 months) and ECBQ (for 18-36-months) (Garstein and Rothbart, 2003; Putnam et al., 2006; Putnam et al., 2008). These tools split temperament in 14 and 18 dimensions respectively that each load on three factors (surgency, negativity and regulation). Two crucial factors in understanding early temperament are emotional reactivity and emotional regulation (Rothbart, 1986; Rothbart and Bates, 2007). Here, we will outline the significance of temperament in socioemotional development. Infant temperament has been associated with important aspects of later social behaviour, as well as concurrent socioemotional stimulus processing.

Associations Between Early Temperament and Childhood Social Development

There is evidence that even infant disposition may be related to later social behaviour in combination with other factors. For instance, infant boys who exhibit both right frontal EEG asymmetry and negative reactivity at 9-months-old have been demonstrated to show higher social wariness at 4-years-old (Henderson et al., 2001). This demonstrates that temperament can predict later social behaviour in conjunction with variables like physiological disposition, demonstrating that both that temperament is important for socioemotional development and that these relationships can be highly multivariate.

There are also direct longitudinal relationships between early temperament and social development. Infants who are both highly emotionally reactive and highly negatively reactive later demonstrate higher levels of social reticence at 3 to 4-years-old (Fox and Henderson, 1999), and less reactive 4-month-olds go on to show less empathic behaviour towards strangers at 2-years-old (Young, Fox and Zahn-Waxler, 1999). This suggests a direct role for temperament in later social behaviour.

In terms of socioemotional stimulus processing, in older children, emotional regulation has been associated with response to emotional expressions. Using EEG and children ages 5- to 9-years-old, Dennis et al. (2009) found that larger P1 and Nc amplitudes in response to fearful and sad faces were associated with higher emotional regulation. P1 amplitude is associated with processing of different emotional expressions (Batty and Taylor, 2003), while Nc is associated with attention (Richards, 2003). This would suggest that the qualitative nature of emotional expressions differs with disposition. Possibly, larger Nc responses reflect negative FEE being more salient for children scoring higher in emotional regulation.

Potential Explanations for the Relationship Between Temperament and Socioemotional Development

One potential mechanism for the above relationship between infant temperament and later social behaviour may be that infant temperament affects how emotional expressions are

processed early in life. Given the evidence that temperament is related to FEE processing in childhood, this would seem plausible.

There is indeed evidence that infant temperament is related to concurrent emotional expression processing. For example, using fNIRS, Ravicz et al. (2015) found preferential activation of the left prefrontal cortex in response to smiling facial expressions in 7-month-olds demonstrating low negative emotionality. Further, Martinos et al. (2012) found a slower peak latency of Nc for infants (ages 3- to 13-months) with poorer regulation in response to emotional expressions and enhanced Nc responses to fearful faces for those with better regulation. They also found enhanced Nc to happy faces in those with higher negativity. This suggests a relationship between infant temperament and neural response to FEE. According to the Nc responses elicited in these studies, negative FEE such as fear could be more salient for infants with higher regulation, and happy FEE could be more salient for those higher in negativity. These differences could shape input and therefore development.

Beyond FEE, infant temperament is associated with concurrent general face processing, which could also shape long-term relationships between temperament and socioemotional development. In infants 6- to 10-months, surgency and orienting interact with characteristics of a face to predict how they scan it (Rennels et al., 2020). Not only does this demonstrate that temperament is directly associated with face processing, but this moderation of face scanning

behaviour also provides a potential mechanism for a causal relationship between temperament and FEE processing.

The relationship between aspects of temperament and socioemotional stimulus processing extends to bodies. For example, infant fearfulness at 8-months is negatively associated with neutral fearful body expressions, with this relationship moderated by maternal sensitivity (Rajhans et al., 2015). While there is less research on the relationship between temperament and body processing, this provides some evidence for an equivalent role for temperament in this domain.

Differences in concurrent attending to and scanning of faces, bodies and FEE highlight one way in which infant temperament could shape development— through shaping the way that infants attend to and engage with stimuli at earlier time points. For example, it could be that in response to a fearful face, infants would experience negative arousal. Then, in infants with higher reactivity and poorer self-regulation, this arousal could be both higher initially and more likely to remain high. This could then affect looking behaviour and the way in which infants interact with the environment, thereby shaping the environmental input. Assuming that environmental input is formative of socioemotional processing development, this would then mean that temperament shapes social development through mediators including visual input.

There are ways in which temperament could shape socioemotional visual input beyond looking behaviour. Particularly, toddlers, being more advanced in their motor and verbal development, have at their disposal a wider range of potential ways to interact with their environments and others in it. This provides more avenues for internal factors such as personality to shape the input, and by extension toddler development. As outlined above, a widely used measure of personality is the ECBQ (early childhood is the Early Childhood Behaviour Questionnaire) (Putnam et al., 2006). Scores on the ECBQ have been associated with risk of developing socially based disorders such as ASD (Pijl et al., 2019). The ECBQ returns scores on a variety of dimensions, some of which are more directly related to executive functioning (for example, inhibitory control, or other dimensions that load on regulation) and others which are more immediately related to social behaviour (for example, cuddliness). These dimensions could be related to and shape socioemotional development in many ways.

Firstly, it could be that aspects of personality that are more closely related to executive functioning shape the environmental input through shaping the responses of individuals around the child as the child commits transgressions. Low levels of inhibitory control, which is closely related to both executive function and self-regulation, are associated with higher levels of aggressive behaviour (Suurland et al., 2015). This is behaviour that typically transgresses social norms and parent instruction and so could change both the degree to which children are exposed to faces and bodies and the degree to which they are exposed to different emotional expressions. It could be that children with low levels of inhibitory control are exposed to faces and bodies more frequently, as adults often correct their transgressions, or less frequently, as

adults prefer to spend less time interacting with them. It could also be that these children are exposed to more negative emotional expressions as these transgressions elicit negative social feedback. These executive control mechanisms may also impact on the ways in which toddlers attend to the faces and bodies of themselves or others, as depending on developmental level, they may be more or less able to selectively attend to pleasant or useful socioemotional stimuli.

There are other dimensions of temperament that may impact on the attitudes and behaviours of adults around the child, and therefore the socioemotional input of the child in ways similar to regulation. For example, elements of negativity may influence the responses of those around the child. Toddlers high in negativity may be more challenging for parents, as they express negative emotions more frequently and for longer durations, being harder to soothe. Maternal stress is positively associated with toddler negativity scores (Oddi et al., 2013). Adults experiencing negative emotions around and possibly directing these emotions towards a child may then accordingly modulate their social behaviour and time spent around the child.

More directly socially related aspects of early personality, which often load on surgency or regulation, such as cuddliness and approach, could directly impact on the frequency with which children are exposed to faces, bodies and facial emotional expressions through moderating their tendency to approach and engage with other individuals in the environment. Indeed, lower cuddliness in toddlerhood in high-risk children has been identified as a component of a personality profile predictive of later ASD diagnosis at 3-years (Clifford et al., 2012). While it

could be that the reduced scores on cuddliness are a marker of an ASD profile already emergent, it could also be that this relationship has a causal component, with children less likely to engage with other individuals in the environment receiving less socio-emotional input, culminating in these later social deficits.

While we have discussed these elements of temperament shaping the environment (and therefore development) as being concurrent, there are many ways in which similar processes could take place longitudinally. There is a significant degree of stability in temperament from infancy to toddlerhood (Carranza Carnicero et al., 2000), so toddlers with “challenging” temperaments are more likely to have been infants with “challenging” temperaments. Long-term, temperament could shape relationships with and behaviours of others in the environments of toddlers and children. So, the impact of temperament on very early formation of social relationships may then shape elements of the socioemotional input long-term.

While there are many aspects of infant and toddler temperament that have been associated with concurrent and later social development and behaviour, any proposed causal mechanism through shaping environmental input remains hypothetical, as there is currently no measure of how naturalistic socioemotional input is related to temperament. To fully assess this hypothesis, it will be essential to examine relationships between the environmental input and both concurrent and longitudinal temperament.

Processes of Development: Maturation Versus Environmental Input

Given that in this thesis we will be examining the nature of the environmental input, a matter of importance is how the environment may shape processing of emotional stimuli. From a maturational standpoint, the factor of key importance would be an automatic, gradual “activation” of skills and neural regions over the course of development, producing a standard developmental timeline regardless of input. Alternatively, some put a greater emphasis on the role of the environment. We will evaluate evidence for these approaches in the context of face processing here, as there is more evidence in this domain.

We can find evidence in favour of maturation from preterm infants. There is rapid neural development in utero, thus preterm and full-term infants are first exposed to social stimuli such as faces at different stages of neural development. Preterm infants are more likely to develop socio-emotional deficits (Spittle et al., 2009), and in comparing EEG responses to emotional faces in preterm and full-term infants at 4- and 8-months-old (age corrected for premature infants) Carbajal-Valenzuela et al. (2017) found that the responses differed between the groups. Namely, at 8-months-old, full-term infants exhibited higher occipital power while processing positive faces as compared to neutral and non-faces, while premature infants showed no difference, suggesting impaired processing of emotional stimuli. Therefore, the level of cortical maturation at which infants first are exposed to the social world impacts the way in which social stimuli are processed; however, there are many additional physical challenges that preterm infants may face beyond differences in cortical maturation at birth, resulting in confounds like

more time spent removed from the mother in intensive care units that could explain these differences (Markestad et al., 2005).

Another way that we can establish support for a maturational view is through examination of early brain lesions. For example, Farah et al. (2010) report the case of an individual who acquired bilateral occipital and occipito-temporal lesions at 1 day of age and subsequently never developed the ability to recognise faces despite typical exposure. This failure of neural plasticity to compensate supports a maturational view of specifically face processing, as it suggests that there are certain neural substrates for face processing, and damage to them can preclude this ability despite ample experience. However, it is important to note that facial emotion recognition was not tested in this individual, so we cannot say whether deficits extended to this domain; further, it may be that while some specific neural substrates are necessary for aspects of socioemotional stimulus processing, in typical individuals the environment is still highly relevant and formative.

There is evidence against a maturational view of face processing in favour of an interactive specialisation view, which accounts for some aspect of environmental influence: Joseph et al. (2011) used fMRI to examine responses to faces in children aged 5 to 12-years-old, as well as adults. While increased specialisation in some brain regions was observed with developmental time, consistent with the maturational account, a network of regions was recruited for this processing, and there were changes in the mapping of functions as well as regressive changes in

networks activated. This supports an interactive specialisation account of face processing, rather than a maturational account, which is not compatible with diverse networks being recruited for the processing of faces or regressive changes.

There are many ways in which we can directly see how the environment is directly related to the development of socioemotional stimulus processing. One approach to this focuses on atypical socioemotional input. For example, infants of mothers exhibiting depressive symptoms during pregnancy and the early postpartum period show differences in emotional expression processing, with a bias in attention towards fearful faces (Kataja et al., 2018). Since mothers with depressive symptoms will likely provide an atypical early environment, this could suggest that processing is shaped by the environmental input. However, we do not know exactly what differences would be critical, if this were the case. It could be a matter of simple differences in the frequency of exposure to emotional expressions, or it could be related to style of interaction. Further, these outcomes could have been shaped in alternative ways: for example, shared genetic risk factors for depressive symptoms may affect socioemotional processing (Shih et al., 2009). It could also be that depression is shaping the uterine environment during gestation through changes in anything from hormones to nutrition.

Institutionalised children can also provide a rich avenue for observing the impact of highly atypical socioemotional input. Children exposed to institutionalised care are raised in group homes which are typically characterised by large group sizes, business-like caregiver behaviour,

and a lack of a nurturing and stimulating environment (Van Ijzendoorn et al., 2011). Differences in development have been observed in these children—Sloutsky (1997) observed, in comparing children in orphanages to those in kindergartens, that children from institutions correctly identified emotions in photos and stories significantly less frequently. Further, comparing the children who had been institutionalised for less than 2 years to those who had been institutionalised for 4-6 years, the former showed significantly less impairment, providing evidence for the environment being key. From this, we may conclude great environmental influence on facial emotional expression processing. However, performance on the Weschsler IQ test correlated with emotion identification and differed between the orphanage and kindergarten groups; further, neglect in these contexts extends beyond a lack of sensitive social interaction. This may implicate general developmental delays as a confound.

There is evidence that in more “typical” environments, with fewer confounds related to nutrition or developmental delays, selective deprivation of visual input may result in socioemotional stimulus processing deficits. For example, infants born with bilateral congenital cataracts, which preclude visual exposure to faces in the first weeks of life, later show impairments in face identity recognition when aspects such as view orientation or facial expression vary (Geldart et al., 2002). This demonstrates a subtle but marked deficit in processing for infants deprived of visual input, even for a short period and in otherwise presumably typical environments.

In sum, there does appear to be some significance for the environment in driving the development of face and facial emotional expression development. At the least, there is an association between the input and whether these skills develop appropriately, with extremely impoverished environments associated with markedly impaired later socioemotional stimulus processing. However, there is much that we do not know in terms of what characterises the typical infant visual environment.

Known Characteristics of the Visual Socioemotional Environment

The visual socioemotional environment here is taken to mean the faces, facial emotional expressions and bodies within the view of an individual. While “seeing” a socioemotional stimulus does not necessarily imply attending to it or extracting information from it, the first necessary stage of any social processing or learning about a visual stimulus is seeing it.

Known Characteristics of the Adult Visual Environment

To fully understand the significance of the toddler visual environment, it is necessary first to understand the developmental endpoint: the adult visual environment. How often do adults see faces, bodies and facial emotional expressions? While this may seem like a simple, even basic question, there is surprisingly little direct evidence that can be called upon to answer it.

There has been some attempt to directly measure the degree to which faces are in the view of adults. Oruc et al. (2019) used egocentric, first-person cameras embedded in eyewear to

measure the degree to which faces were in the view of 30 adults undergoing a diverse range of everyday activities. On average across participants, 22.5% (SD = 12.5%) of frames contained faces, with most of those faces being familiar faces rather than those of strangers. Male and female faces were roughly equally represented, and about 82% of faces were own-race. Faces were most commonly viewed at a $\frac{3}{4}$ angle (i.e., turned away slightly), followed by a frontal angle. The median size for familiar faces was 7.1° , which translates to a median viewing distance of 109 cm. This is consistent with intentionally face viewing in the context of social interaction being a driver for face presence in the adult view. This research appears to be the only attempt thus far at directly recording the visual input of adults with respect to faces.

However, there has been no direct assessment of the adult visual input of bodies and FEEs.

While these are assumed to be common components of the adult visual environment, this has not been directly tested.

Known Characteristic of the Infant and Toddler Visual Environments

Bearing the above in mind, we must outline what is known about the infant and toddler visual input. At three-months-old, first-person exposure to faces makes up a large proportion (25%) of the visual environmental input (Sugden et al., 2013).

While we might intuitively expect infants to be exposed to stimuli like faces often throughout the developmental timeline, there is evidence that the frequency of faces declines over the first 18 months of life, and that this decline is strictly for faces (Jayaraman et al., 2017). Fausey, Jayaraman and Smith (2016) examined the frequency of faces and hands in the infant view using naturalistic egocentric recordings taken from 34 infants ranging in age from 1-month to 2-years, finding a decline in the presentations of faces (24% or the data for the youngest infants, 7.5% for the oldest) and an increase in presentations of hands over developmental time. Jayaraman, Fausey and Smith (2017) examined egocentric head camera data for 36 infants, including one to 24-months-old, with an average usable recording time of 4.2 hours per infant and sampling at 1/5 Hz. Overall, 39.6% of frames contained individuals other than the infant; the overall percentage of frames with faces was 23.2% for the youngest infants and 8.83% for the eight oldest infants (18-months and older). This would seem to suggest that by toddlerhood, faces are relatively infrequent, and that this decline in frequency is for faces alone.

Interestingly, prior lab-based research utilising more structured play interactions often yields the opposite age-related pattern for face frequency, with the youngest infants seeing faces the least frequently, possibly because they are typically held against the body of the caregiver during these play interactions (Frank et al., 2013). This could indicate that the kinds of interactions caregivers (and later, their toddlers) choose to engage in over the course of everyday life is a driving force in shaping these age-related changes in the visual input.

One theoretical proposal is that the visual input changes over time through the changing actions of the infant or toddler, creating a de-facto “curriculum” for learning about the world, including faces (Fausey, Jayaraman and Smith, 2016; Jayaraman, Fausey and Smith, 2017; Smith et al., 2018). There is evidence that infants play a role in shaping their own perceptual input for other skills: for example, in object learning, infants shape the input by manipulating the objects that they are viewing in a way that optimises object learning as compared to adult views (Bambach et al., 2016). While younger infants have limited scope in terms of controlling their own view of faces, they have some direct control (in the form of eye/head movements) and some indirect control (in their influence over the behaviour of their parents). However, their ability to independently locomote improves drastically over the first months and years of life, with toddlers able to navigate the environment and seek out input independently.

The motor activity of infants seems to constrain and facilitate learning from the visual environment. For example, the manual object exploration skills of 7- to 8-month-olds are associated with visual prediction proficiency for explored objects (Kubicek, Jovanovic and Schwarzer, 2017). Typically, infants transition from an inability to locomote (resulting in much time spent prone) to sitting and crawling to walking (Franchak, 2019). With the transition from crawling to walking, visual input from caregivers and other individuals is more readily available, particularly when it comes to faces (Kretch, Franchak and Adolph, 2013). This illustrates the importance of considering separately the visual input for children at these three stages of physical ability. Particularly, independent shaping of the input both directly and indirectly may be most highly relevant for walking toddlers.

Considering the aforementioned research on the input of faces and bodies alongside the idea of distinct visual input for crawling and walking children, there are limitations that we must consider with regards to our understanding of toddlers. Toddlers, whose motor skills are typically well-developed, particularly for those approaching the 24-month mark, are underrepresented in previous research samples. Only thirteen of the thirty-six participants in the headcamera study of Jayaraman et al. (2017) were 12-months or older, with only seven older than 16-months. Given that 12-months is average onset of walking, that infants may not walk consistently from this point, and that we would expect the visual input of walking infants to differ from that of crawling infants, this under-representation of older infants is significant and could be partially driving findings on decreased incidence of faces in the visual environments of older infants.

Further, prior research has typically only considered real-life presentations of the faces and bodies of other individuals. Given that infants are often exposed to the bodies of themselves and others through independent play, media and mirrors, that extracting information about the structure of faces and bodies is equally plausible this way, and that this exposure especially may be self-directed, this is an important avenue of input for visual learning that while qualitatively different, should not be discounted. In addition, because of the very rich nature of this data, a variety of approaches have been utilised for coding. Namely, Jayaraman et al. (2016, 2017) sample images once every five seconds. This coarse sampling may not fully represent the visual

input. So, face presentations may be underrepresented in these previous datasets as a result of both the face definition used and the method of quantifying face frequency in the toddler view.

Crucially, previous research on the toddler environmental input is also limited in that FEE are not quantified. Much of the research on emotional expression development referenced in this review finds that the ability to recognise different expressions emerges at different points, with explanations often depending on assumptions about which expressions infants would be exposed to most and least often. A direct measure would provide evidence for or against these assumptions, but currently no characterisation of the infant visual input regarding facial emotional expressions exists.

Overall Research Aims

So far, we have discussed what is known and unknown about how infants develop their processing of faces, bodies and the ability to distinguish between and process different emotional expressions. This project intends to investigate some as-yet unexamined questions, broadly speaking on the visual input of toddlers in terms of FEE, as well as to re-visit and explore with a wider sample and more detailed coding the extent to which they see faces and bodies. Not only this, but we also intend to examine individual differences that may be associated with variation in this input, as well as corresponding caregiver visual input within the same environment.

The first aim of this thesis is to address the question of whether previous findings on the rates at which toddlers are exposed to faces and bodies are replicated with a bigger sample, a wider definition of “face” and frame-by-frame sampling. Secondly, we wish to quantify the input of FEE for these toddlers. There remains no quantification of this so far. Further, we aim to establish what factors may shape the nature of this input, both with respect to the frequency of faces, bodies and FEEs. We will examine the individual differences by toddler temperament (both concurrent and longitudinal) and whether they have child siblings. Finally, we will assess corresponding caregiver face, body and FEE input within the same environment. We will directly assess this in walking toddlers in the second year of life ($M = 29.37$ months, $SD = 0.50$).

The Projects at Hand

The proposed aim of the five studies presented in this thesis is to advance understanding of the naturalistic socioemotional visual input of toddlers. This will be done through analysis of egocentric view head camera data recorded from toddlers in the second year of life and their caregivers while they navigate their home environments in a naturalistic and unconstrained way. This data will be processed, coded and analysed for five studies. For full methodological details, see the general methods section.

Study 1.

Study 1 aims to quantify what percentage of the egocentric view of infants in the second year of life contains faces and bodies. The purpose of this study is to re-test conclusions from previous

research that toddlers at this age are exposed to faces infrequently and examine the corresponding visual input of caregivers within the same environment.

Previous research has suggested a decline in faces (but not bodies) in the visual field from birth to toddlerhood. However, the research that exists has its limitations. For example, sample sizes of toddlers old enough to be confidently walking are limited. Further, there is no clear evidence as to whether any decline in face frequency is a temporary, unique feature of toddlerhood or whether this is reflected in the caregiver view within the same environment, which would be informative in understanding the decline. We hypothesise that the frequency of faces in the toddler view will be lower than the frequency of bodies in the toddler view, but the percentage of frames containing faces in the toddler view will be higher than that specified for toddlers in the second year of life in prior research. We also hypothesise the frequency in the caregiver view will be lower than body frequency in the caregiver view, but face frequency in the caregiver view will be comparable to that in the toddler view.

Study 2.

Study 2 will assess what factors influence the prevalence of faces and bodies in the egocentric toddler view using the same sample. Currently, there is no research that directly addresses whether and why face and body frequency may vary. Variation could come about because of external or internal individual differences. We will examine the presence of child siblings for the former and temperament for the latter. We hypothesise that presence of child siblings,

temperament concurrent to recording and temperament measured longitudinally ahead of recordings will be related to individual differences in the degree to which toddlers see faces and bodies.

Study 3.

Study 3 builds on validation data used for FEE detection models used in studies 4 and 5. To validate a computer vision model capable of FEE detection with naturalistic footage, we asked typical adults to assign categorical labels to 311 dynamic, short instances of facial movements extracted from our video corpus. We also asked them to give corresponding intensity and confidence ratings. This kind of data can give good insight into typical adult perception of the kinds of real-world FEE that are lacking in traditional perceptual studies. Prior research has focused almost exclusively on posed, static and artificial emotional expressions, which are highly unrepresentative of the kinds of judgements adults make in their everyday lives. We hypothesise that adult ratings of real, dynamic FEE would be far more varied and ambiguous than research using posed FEE would suggest.

Study 4.

Study 4 will aim to assess what facial emotional expressions toddlers actually see in their naturalistic home environments and analyse corresponding caregiver visual input. We hypothesise that while both positive and negative FEE will be present in the view of the toddlers, positive and neutral FEE will be more frequent, and the least common FEE in the

toddler view will be anger, fear and disgust (which would be more frequent in the caregiver view).

Study 5.

Study 5 will utilise the same data as study 4 to establish to which factors may shape variation in the degree to which toddlers see FEE. There is diversity in mature ability to extract FEE from faces, as well as diversity in frequency of FEE in the toddler view. However, no attempts have been made to characterise the degree to which visual FEE input may vary between individuals, including toddlers, or which factors may shape this variation. This may be especially notable in toddlerhood, at which point individuals are more able to shape their visual input through their newfound motor independence. We hypothesise that the frequency of FEE will be related to both external factors (the presence of siblings) and internal factors (toddler temperament).

General Methods

Participants

Parents of 41 toddlers (22 female, 19 male) were recruited from in and around the York area (UK) and invited to take part. Of these, 26 (16 female) recorded usable data from the child perspective and were included in the final analysis. The average age of participants was 29.26 months ($SD = 0.56$ months); for female participants, this was 29.37 months ($SD = 0.50$), and for male participants this was 29.08 months ($SD = 0.62$). Of the $N = 26$ families that recorded usable data from the child perspective, $N = 24$ also recorded usable data from at least one caregiver perspective (inclusion criteria for footage are specified after the procedure). The data for $N = 26$ participants recorded footage from the toddler perspective, while a subset ($N = 24$) recorded footage from the caregiver perspective also. For the analysis, for papers 1 and 4, we decided to first consider the full $N = 26$ for the presence of socioemotional stimuli in the toddler view, and then separately, for the comparison of the toddler and caregiver view, to only analyse the $N = 24$ participants with corresponding caregiver footage.

All 26 participants provided sibling data, while $N = 23$ also completed the ECBQ, and $N = 22$ completed the IBQ. So, for papers 2 and 5, we consider 26 participants for sibling status and 23 and 22 participants respectively for temperament analysis.

Apparatus

Families were sent a pack that included a head-mounted camera to record the egocentric view of the child, a head-mounted camera to record the egocentric view of the parent or caregiver, and one room camera that gave a view of the entire room as context for the

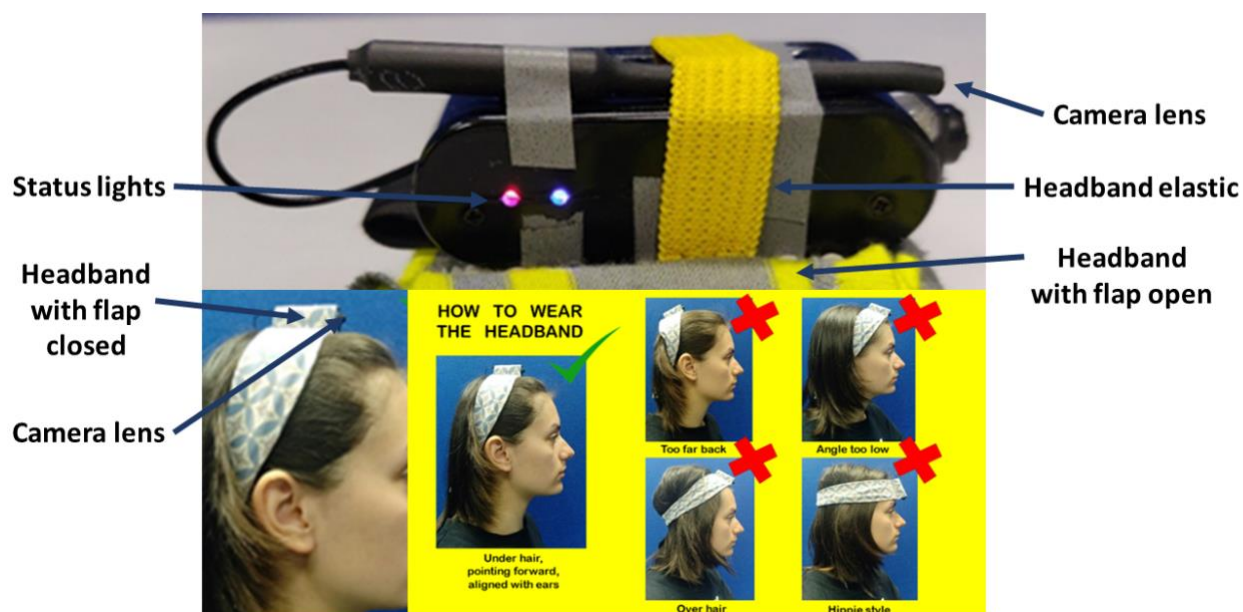
egocentric perspectives. Packs also included instructions for how to operate the devices, SD cards and all other relevant materials.

The two head-mounted cameras were small, light and portable SANPROV 1080P cameras, selected for their ability to capture data unobtrusively and safely. They recorded audiovisual data using a 90-degree wide-angle lens at HD 30 fps 1080P quality to an SD card.

Figure 1 demonstrates a head-mounted camera inside a headband strap.

Figure 1

An example of a head-mounted camera placed in a headband and instructional image given to families.



Note. The top image demonstrates the cameras use to record toddler and caregiver perspectives placed correctly on one of the headbands, and the bottom left image

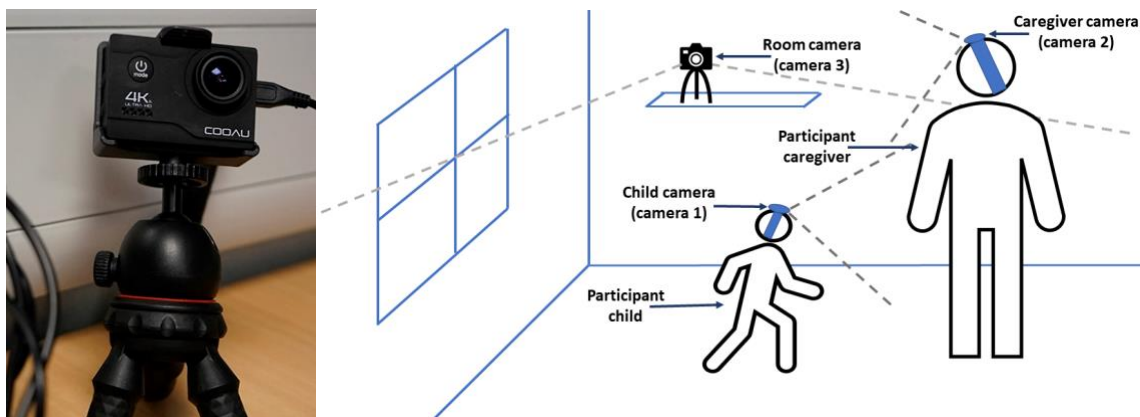
demonstrates correct placement of a headband on a head. The bottom right image demonstrates correct and incorrect placements for the head-mounted camera; this image was given to families ahead of testing.

The cameras were mounted on fabric headbands in a breathable material made to fit the measurements of the individual participants. In the instructional materials participants received ahead of participation, they were given guidance on headband placement—the example of ideal placement that they were given is shown in figure 1.

Finally, the Cooau brand room camera was an action camera bigger than the head-mounted camera but still compact, and recorded 4k 30fps video digitally to an SD card; families were given tripods to allow them to place it optimally for capturing the entire room (Figure 2). The three camera perspectives allowed for more information to be utilised when coding (e.g., the caregiver and room camera perspectives made it clearer when a child head camera had been removed or placed incorrectly).

Figure 2

The room camera in one possible placement on the supplied tripod and a schematic depiction of a recording session



Note. The left image shows the room camera placed in the supplied tripod, which was flexible and allowed for versatile placement; the right image is a schematic depiction of a recording session, demonstrating how the room camera would be placed to record a wider perspective on the interaction.

Procedure

The families in this dataset were part of a longitudinal data collection and had already provided IBQ data at an earlier age, when the participant children were 13-months of age (September 2019 – August 2020). Video data collection took place from January to December 2021, over the course of the COVID-19 pandemic, necessitating minimal contact with participants. For a week prior to data recording, parents of participants were asked to fill in a diary of what their infant was doing and when over the seven days. On this basis, days and times for recording were suggested. Experimenters met with participants virtually to discuss

instructions. Parents received packs, and participants were instructed to record for one hour per day for seven days, before the apparatus was retrieved by a member of the lab. The ECBQ was completed online by parents at the end of the recording week.

Data Exclusion

Of the 32 families that recorded data, 26 recorded analysable data from the child perspective; this was classed as footage without persistent image quality impairment that was recorded from an egocentric camera positioned correctly (forward facing, at the front of the head) on the participant child. Any footage that was blurry, not recorded from the perspective of the correct individual, or recorded from an incorrectly positioned camera was excluded. Of these 26 families, 24 recorded analysable data from the caregiver perspective: this was classed as footage from the caregiver-perspective that was free from blur or other image impairment and recorded from the correct egocentric camera position (top of the head, forward-facing) from the caregiver perspective.

This usable footage totalled 140 separate recordings from the child perspective, with an average of 278,683.69 frames (approximately 5.20 hours) of raw footage recorded per child. Each separate recording was one segment of continuous recording that lasted longer than 15 minutes. In total, this child-perspective data contained 7,245,776 frames of raw footage (135.26 hours). Within this, frames wherein the head camera was not being worn were excluded; 624,384 frames in total (11.66 hours) were excluded on this basis— an average of 0.45 hours, or 27 minutes— leaving 123.61 usable, codable hours of footage from the child perspective— an

average of 4.75 per participant. For the caregiver perspective, there were 6,896,475 analysable frames overall (approximately 128.74 hours, or an average of 5.36 hours per participant). A summary of recordings generated can be found in table 1.

Table 1

Video data inclusion metrics summarised

Metric	Toddler Perspective	Caregiver perspective
Participants recording successfully	26	24
Number of recordings	140	141
Recordings from each perspective	Toddler = 140	Mother = 123 Father = 16 Grandmother = 16
Total frames	7,245,776	7,259,896
Total hours	135.26	135.53
Excluded hours	11.66	6.78
Included hours	123.61	128.74

Note: This table provides a summary of the recordings collected and analysed from each perspective.

Algorithmic Coding

With such a rich dataset, computer-driven, automatic detection of relevant visual features within the video footage was determined to be the feasible course of action. However, the accuracy of such methods was uncertain for the videos recorded by cameras worn on the children and adult heads, requiring that selected algorithms were tested against manual coding for accuracy.

For body detection we chose to use the pre-trained YOLOv6-S deep neural network model by Meituan Inc. (Li et al., 2022) due to its balance in efficiency and accuracy in general object detection. Their approach is an evolution of single-stage neural network architectures where an image is passed through sequential convolutional layers for extracting features and subsequently decoded for regressing object bounding boxes and classes simultaneously. Specifically, they refine the backbone, neck and head architecture originally formulated in YOLOv4 (Bochkovskiy et al., 2020). A RepBlock (Ding et al., 2021) based backbone is used to extract rich features from the input image while remaining computationally inexpensive. The neck subsequently uses a path aggregation neural architecture (Liu et al., 2018) to aggregate features from different levels of abstraction, also enhanced with RepBlock layers. Finally, the head leverages individual convolutional branches to decode these aggregate features into object class and bounding boxes. Additionally, Li et al. boost performance of YOLOv6-S by utilising advanced training techniques (e.g. self-distillation and data augmentation) and by using efficient loss functions for training the model.

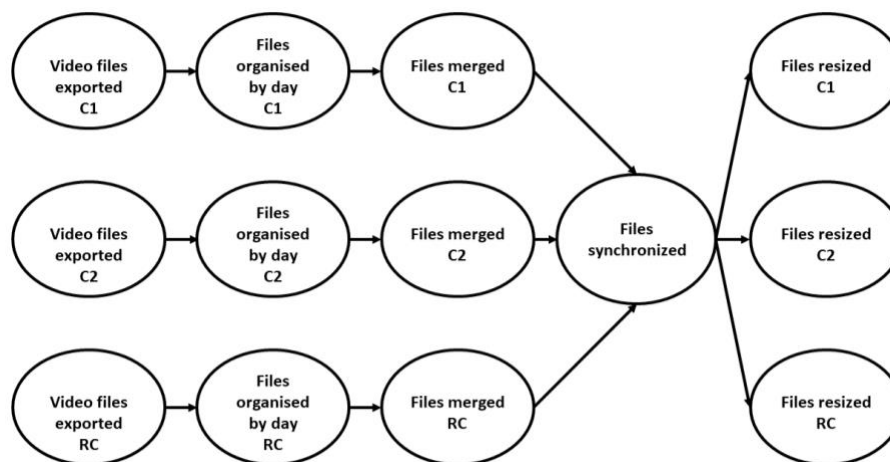
Face detection was largely conducted using Retinaface and Arcface (Deng et al., 2019, 2020) with 1-dimensional convolutional neural network applied to per-frame maximum detection confidence values to output per-frame probability that a frame contains a face; further identity information was extracted to allow for identification of other variables in a frame, such as the number of faces (Kalogeiton and Zisserman, 2020).

Manual Coding for Face and Body Algorithm Validation

First, all video files were pre-processed (see figure 4). Then, 14.3% of child-perspective videos were coded for the presence of faces and bodies using a Matlab program allowing for manual labelling of each frame. These videos overall contained 990,027 frames (18.47 hours) of child-perspective data taken from six families and were manually coded frame-by-frame. Several researchers contributed to frame coding. To ensure accuracy, each frame was viewed a second time by another researcher who checked. This manual coding was used as the basis for validating the face and body computer vision models.

Figure 4

The data processing pipeline prior to coding



Note: This chart illustrates the stages of video processing prior to coding. Videos were first exported from SD cards to the lab drive and sorted by day for each family. Video files for each day were merged, before videos from the three perspectives were synchronised temporally. Finally, files were resized for compatibility with the processing pipeline.

Frames were labelled “face” if they contained any real or photorealistic internal feature of the face (eyes, nose, mouth), regardless of source (self, other, media); frames were labelled “body” if they contained any identifiable real or photorealistic body part regardless of source (self, other, media). Frames for all videos were also coded for whether the head camera had been removed. Beyond the validation sample of videos, the ‘camera removed’ frames were manually identified for every video.

Validity of the Computer Vision Models for Face and Body Detection

To test the classification accuracy of the face and body detection models, their results were analysed in terms of agreement with the human coding. Commonly used measures of model accuracy are recall, precision and F1. Recall is the proportion of actual positives that are detected; precision is the proportion of reported positives that are true positives; and F1 is a metric based on precision and recall (Sundaram and Mani, 2016). The face detection model was successfully trained on naturalistic data and validated with a recall, precision and F1 of 0.72, 0.76 and 0.74 respectively on this validation subset of videos specifically. The body detection model also performed well. Specifically, on the subset of validation videos that were manually coded, the precision across videos was 0.89, the recall was 0.75 and the F1 was 0.81. See table 2 for a summary of these statistics.

In this context, the ground truth was manual coding for face and body. Therefore, lack of correspondence between algorithmic and manual coding does not always imply an error on the

part of the model, given the errors and value judgements inherent in manual coding of such a large subset of data. However, the close correspondence between human and model coding here does indicate that the model replicates careful human coding well.

Table 2

Face and Body Detection Models Summarised

Metric	Face model	Body model
F1	0.74	0.81
Recall	0.72	0.75
Precision	0.76	0.89

Note: This table provides a summary of the accuracy statistics for the face and body detection models.

Validity of the Computer Vision Models for Facial Emotional Expression Detection &

Categorisation

We used the confidence-weighted “winning” labels extracted from the human behavioural data collected for chapter 3 to validate our FEE categorisation model. Human annotations were collected for these 311 validation clips, and then the winning and runner-up confidence-weighted human labels were compared to the corresponding algorithm labels. The model identified face frames and outputted a likelihood for each FEE label for each frame.

Frames above 70% confidence for face detection were analysed for emotional expression content. An average likelihood for each FEE label was generated for each validation clip, with the “winning” model FEE label being the one with the highest average likelihood, and

the runner-up the second-highest likelihood. For 62.06% of the clips, the winning human label was selected as the winning model label. See table 3 for a breakdown of how human ratings corresponded with model assignments. In sum, the algorithm selected the winning human label as its winning or runner-up answer for 81.03% of clips and selected the runner-up human label as its winning or runner-up label 40.52% of the time. Only 5.47% (17) of the clips did not meet any of these criteria. On this basis, particularly considering the high rate of disagreement amongst human raters of these clips, we decided that this model was appropriate.

Table 3

Facial Emotional Expression Model Agreement with Human Coding Summarised

<i>Percentage of validation stimuli in each category (%)</i>	Model winning label	Model runner-up label
Human winning label	62.06	18.97
Human runner-up label	16.08	24.44

Note: This table provides the percentage of validation stimuli wherein the model and human participant confidence-weighted winner and runner-up labels agreed.

Participant Variables

Demographics and information on the presence of child siblings in the home was verbally provided by parents and recorded by experimenters; the accuracy of this information was double-checked through viewing of video recordings. This information was present for all participants included in the final sample based on valid video recordings.

Temperament Assessments

Child temperament was assessed via parental report, using the ECBQ and the IBQ. IBQ data was obtained for 23 of the participants who provided analysable video recordings; parents had provided this information when the participant children were 13-months of age, over a year before the video recording stage of the experiment. Out of 26 participants who provided analysable video recordings, 24 also completed the ECBQ. The ECBQ was filled in by parents online, at the end of the recording week. The link to the online questionnaire was only sent to parents at the end of the week, to not influence their spontaneous behaviour. Complete ECBQ data was available for 24 participants.

Quantifying the Frequency of Faces and Bodies in the Egocentric View of Toddlers

Emma J Jackson, Bruce Muller, William Smith and Elena Geangu

This paper has not yet been submitted for publication

Abstract

Faces and bodies are an important feature of the toddler environment, and the ability to recognise and extract information from them is fundamental. While seemingly simple, these skills have protracted developmental courses. Information from the environment may be crucial in development. Prior research has suggested a decline in faces (but not bodies) in the egocentric view from birth to toddlerhood. However, walking toddlers are underrepresented in the extant data; further, there is no clear evidence as to whether any decline in face frequency is a temporary, unique feature of toddlerhood or whether this is reflected in the adult view, which would be informative. To address these gaps in understanding, we set out to quantify and compare the visual input of faces and bodies in the toddler and caregiver egocentric views. We instructed 26 families in and around York (UK) to record at home using egocentric head-mounted cameras, for one hour a day for one week. We hypothesised that: 1) The frequency of faces in the infant view would be lower than the frequency of bodies in the infant view; 2) The percentage of frames containing faces in the infant view would be higher than that specified for infants in the second year of life in prior research; 3) Face frequency in the toddler view would be lower than that in the caregiver view; 4) Face frequency in the caregiver view might be lower than that previously reported for adults; and 5) Body frequency would be comparable between toddler and caregiver view. Results suggest that the relative frequency of faces present in the view of toddlers is low in comparison to the frequent presentations of bodies. Further, faces and bodies are both more frequent in the caregiver view than the toddler view. This would suggest that infrequent faces are a unique feature of the toddler visual input.

Keywords: naturalistic; face; body; infant; head-mounted camera; egocentric view

Quantifying the Frequency of Faces and Bodies in the Egocentric View of Toddlers

The skills that infants must acquire over the first years of life in order to decode faces and bodies are complex and crucial for typical functioning. Faces and bodies contain a variety of information for infants, children and adults: this information can range from the identity of the observed individual through to internal states of the individual; faces are an especially rich source of social information for humans (Kret et al., 2020). Learning how to decode these rich sources of information is essential for typical social functioning. Examples of deficits in these skill sets and how these deficits can have a profound negative impact on social functioning can be seen in disorders from prosopagnosia (Damasio, 1985) through to Autism Spectrum Disorder (ASD) (Castelli, 2005). Understanding how these skills emerge over the course of typical development, both in terms of timeline and causal mechanisms, can help us to understand both typical and atypical functioning.

Learning to extract information from faces and bodies is essential for infants, and the development of the skillset is protracted, with changes taking place rapidly within the first months of life. From birth, infants show a looking preference for face-like stimuli (Mondloch et al., 1999; Simion and Giorgio, 2015). However, they continue to acquire new skills related to faces throughout infancy and childhood (Phillips et al., 2012; Hoel and Striano, 2010; Booth et al., 2002). Bodies are also an important stimulus, being common in the environments of infants, children and adults and providing a wealth of information about the identity and internal states of other individuals. At birth, neonates exhibit a visual preference for biological bodily motion

(Fox and McDaniel, 1982; Bidet-Ildi et al., 2014), but development of the ability to extract socially relevant information from bodies continues through infancy (Booth et al., 2002).

These changes could be driven in whole or part by the environment. Deprivation of visual input certainly impairs face processing development (Geldart et al., 2002). Prior research examining audiovisual recordings from the toddler perspective appears to suggest a shift from frequent presence of faces in the visual field in the first months of life (Sugden et al., 2014, 2019) to infrequent face presentations infants move towards 24-months (Fausey, Jayaraman and Smith, 2016), with this decline strictly being for faces and not bodies (Jayaraman, Fausey and Smith, 2017). This could be related to changes in locomotor ability and modality of locomotion (Kretch et al., 2014). It is also unclear whether infrequent faces are a unique feature of the toddler perspective, or universal to walking humans in the home environment—there are currently no characterizations of face or body frequency in the corresponding caregiver view, though prior research would suggest that adults see faces at around the same rate as newborns (Oruc et al., 2019). The scope and number of prior research papers have been limited by the labour-intensive manual coding often required for this rich video data. Particularly, mobile and walking infants are underrepresented.

Here, we present data building on this prior research and seeking to replicate its findings in a sample of 26 confidently walking toddlers. Understanding the input toddlers see is important for informing our understanding of the mechanisms of development, and the research on visual input in terms of faces and bodies in this age group is limited. We seek to

establish whether the infrequent faces relative to bodies in the toddler view is robust, and whether this is unique to toddlers or extends to caregivers in the same environment.

Face and Body Processing Development

While the developmental timelines for face and body processing development are protracted, most people are proficient in processing faces and bodies (Diamond and Carey, 1977; Dittrich et al., 1996; Slaughter et al., 2004; Young and Burton, 2018) as holistic elements of the environment (Butti et al., 2022; Weldon et al., 2013; Willems et al., 2014) by adulthood. This means that the features of the face are bound together and processed with reference to encoded face prototypes (Cabeza et al., 1999). By adulthood, processing of faces is perceptually narrowed for the categories of face to which an individual has been most often exposed in phenomena such as the Other-Race Effect (ORE), wherein proficiency is lost for unfamiliar races and ethnicities (Chiroro and Valentine, 1995; Meissner and Brigham, 2001). The vast majority of these elements are not present at birth.

At birth, processing of faces is limited to a preference for faces (or, at least, top-heavy visual arrays) (Mondloch et al., 1999; Simion et al., 1998; Simion and Giorgio, 2015) and a preference for happy faces (Farroni et al., 2007; Rigato et al., 2011) as well as an ability to discriminate between human and monkey faces (Di Giorgio et al., 2012). While neonates may appear to recognise their mothers' faces, this is in fact driven by external features such as the hairline (Pascalis et al., 1995). Skills emerge very rapidly over the first years of life. By 3-months,

infants prefer own-race and own-species faces (a precursor to adult perceptual narrowing) (Kelly et al., 2005; Sanefuji et al., 2014), and demonstrate an emerging ability to process faces holistically and extract cognitive face prototypes (de Haan et al., 2001; Quinn and Tanaka, 2009; Turati et al., 2010). They also prefer individuals of the same gender as their primary caregiver (Quinn et al., 2002).

A true ORE and other-species effect is evident by the second half of the first year (Anzures et al., 2010; Anzures et al., 2013; Pascalis et al., 2005; Pascalis, De Haan, and Nelson, 2002; Xiao et al., 2013) and a wider variety of facial emotional expressions, including threat-related expressions, are discriminated at this point (Grossman et al., 2007; Hoehl and Striano, 2010). Neural responses to faces more clearly resemble those of adults by 12-months (Halit et al., 2003; Rose et al., 2002). Face processing developments continue into the second year of life, with self-recognition and the ability to extract information from more challenging, degraded face images emerging (Doi et al., 2009; Nielsen et al., 2003). In fact, face processing development continues into childhood and adolescence (Chung and Thomson, 1995).

In some ways, body processing development is similar: behavioural and physiological milestones such as specialised and holistic processing must emerge. As well, body processing begins with simple visual preferences at birth, such as preferences for body-like movements presented as point-light displays (Fox and McDaniel, 1982; Bidet-Ildei et al., 2014). However, equivalent body processing milestones take much longer to emerge. It takes until 5-months for infants to demonstrate evidence of holistic body processing and interpretation of motion based

on relationships between limbs. (Booth et al., 2002; Hock et al., 2016). This shift indicates that by 5-months, infants are starting to process the body as a holistic whole. Similarly, physiological milestones are more protracted in their emergence. Nc, an EEG marker of attention, is evident in response to bodies by 3-months but is markedly smaller than that for faces (Gillmeister et al., 2019). This could suggest that bodies are less salient for infants than faces.

In the second half of the first year of life, infant processing of bodies extends to an understanding of goal-directed actions and bodily postures as conveying internal states (Geangu and Vuong, 2020; Woodward, 1998; Zieber et al., 2013). Infants at 8-months can discriminate between congruent and incongruent bodily and facial expressions (Rajhans et al., 2016). It is only at 9-months that some evidence of holistic processing of bodies is evident (Easterbrook et al., 1999; Heron and Slaughter, 2010) and at 12-months that consistent discrimination of possible and impossible bodily motions can be observed (Morita et al., 2012). Holistic body processing, as evidenced by the inversion effect, is not observed until 14-months (Gillmeister et al., 2019; Turati et al., 2004) and intact are not consistently preferred over and discriminate from scrambled bodies until 18-months (Slaughter et al., 2002; 2004). So, many elements of body processing emerge much later than their face counterparts.

Overall, face and body processing both progress from very simple biases at birth through to complex, holistic adult processing of many cues over infancy and childhood. However, these developmental timelines differ somewhat. The environment may play a role in the emergence of both face and body processing abilities.

The Importance of Environmental Input

So, there are many different skills that infants and toddlers must acquire, and the environmental visual input could be a driver of this. Much of this input may come from other individuals; infants and toddlers are rarely alone and receive information from a variety of individuals across different social settings (Elicker et al., 2014). We will outline current evidence for the significance of this environmental socioemotional input for face and body development, both direct and indirect, here.

There is evidence from atypical development that very early visual input shapes socioemotional stimulus processing development. For example, infants born with congenital cataracts, which preclude visual input, demonstrate consistent but subtle deficits in face identity recognition well into adulthood when the recognition task is more challenging through viewing angle or expression changes (Geldart et al., 2002). This indicates a profound impact of early visual input, especially given the fact that cataracts are typically treated rapidly (Lambert, 2004).

In typical individuals, face perceptual narrowing also presents a clear example of variation in the visual input shaping proficiency with different categories of face. Adults are perceptually narrowed to own-race (Meissner and Brigham, 2001) and own-species (Scott and Fava, 2013) faces. The former phenomenon, the ORE, is moderated by the degree of experience with a particular racial outgroup (Hancock and Rhodes, 2008). Neither effect is present at birth, with elements slowly becoming apparent over the first 6-months (Di Giorgio et al., 2012;

Sanefuji et al., 2014) before more established processing advantages observed between 6- and 12-months (Anzures et al., 2010; Anzures et al., 2013; Pascalis, De Haan, and Nelson, 2002; Xiao et al., 2013). This perceptual narrowing is dependent on the visual input and can be modified by changing the statistics of the visual input with respect to species (Pascalis et al., 2005) and race (Heron-Delaney et al., 2011). Even after infancy, research on cross-race/ethnicity adoption would suggest some malleability is maintained (De Heering et al., 2010; Sangrioli et al., 2005). The experience-dependent nature of this perceptual narrowing makes clear how central the visual input is to face processing development.

There is less evidence for the specific impact of the environmental visual input on body processing. For example, the visual deprivation arising from congenital cataracts does not have such a notable impact on later processing of biological motion as it does for faces (Hadad, Maurer and Lewis, 2012). This may indicate that at least some elements of body processing are not as dependent on the environmental input, though the impact of visual deprivation on other aspects of bodily processing has not been examined. It could also be that the very early visual input that would be affected by congenital cataracts is not so essential— many key milestones in visual body processing do emerge later, and so if the environment does shape this processing development, it could be that the first weeks of life are not so critical. Indeed, prior research on the input would indicate that bodies are denser in the visual input of older infants and toddlers, so perhaps this is when the environment is most crucial for body processing development (Fausey et al., 2016; Jayaraman et al., 2015, 2017).

Likely changes in the environmental input of faces also coincide with important changes in face processing. For example, as infants become toddlers, they become more embedded in a variety of social networks outside of their immediate families and caregivers (Lewis and Takahashi, 2005). This could mean that they are exposed to a wider range of individual faces. While this social change takes place, toddlers demonstrate changes in their processing of the faces of different identities. Specifically, Nc response to faces (an ERP component indexing attention) is greatest in response to the mother's face prior to 24-months; from 45-months, the Nc response is greatest for stranger faces (Carver et al., 2003). Given that this coincides with a likely increase in stranger faces, this could either be a maturational adaptation, or perhaps a result of this wider experience.

The environment may vary across developmental time in several ways and for a variety of reasons. One reason could be that parents respond to and interact with infants and toddlers at different ages in different ways. Another reason could be that the pragmatic situations that infants and toddler find themselves in differ— for example, older toddlers are more likely to be at nursery or in school. Another factor could be that infants' and toddlers changing behaviours change the nature of the environmental input. Finally, and on a related note, motor ability and method of locomotion may be highly formative of the visual input (Kretch, Franchak and Adolph, 2014).

Age, Motor Ability, Motivation and the Environmental Input

Here, we will outline the likely relationship between motor ability developments with age and the changing environmental visual input. A key distinguishing factor of our sample is that it consists of confidently and independently walking toddlers, who are underrepresented in previous samples. Differences in motor ability are likely to be highly formative of the visual input (Kretch, Franchak and Adolph, 2014; Kubicek, Jovanovic and Schwarzer, 2017).

Neonates are capable of only very limited motor actions and must depend entirely on caregivers for traversing the environment and meeting their basic needs (Ainsworth, 1969). Neonates progress from an inability to independently change posture, to stable head control (Thelen and Spencer, 1998). They transition to being able to sit at around 6-months before being able to traverse the environment somewhat independently at 13-months (Adolph and Robinson, 2015). Since the earliest forms of independent locomotion, whether walking or shuffling, often require use of the hands and arms for balance and movements, the ability to manually explore the environment is still limited up to the point of confident walking. Each turning point in infants' development of motor skills can shape the visual input of faces and bodies in many ways: pragmatically (through changes in head position and orientation); through allowing independent exploration; and through allowing more indirect influence over the behaviours of others. The transition from being unable to locomote to being able to crawl (Robson, 1984) and then later to being able to walk (Karasik et al., 2011) shapes the environmental input in many ways. This ranges from changes in perspective to increased independence.

Firstly, walking changes aspects of the environment that are visually accessible to infants and toddlers. Given that pre-mobile infants spend their time almost exclusively prone or supine, entirely unable to support their own heads (Konishi et al., 1994), they typically only see elements of the environment above them. They require active input from their caregivers for everything from survival to learning from the environment (Trevathan and Rosenberg, 2016). So, interactions with young infants typically entail caregivers leaning over the infant (Davis and Thoman, 1988) or holding the infant close to the body (Blehar et al., 1977), which would make many aspects of the body less visually accessible.

When infants learn to crawl, and then to walk, which elements of the environment are visually accessible changes dramatically. Kretch et al. (2014) captured the visual input of crawlers and walkers within the lab, characterising distinct visual experiences. While both walking and crawling infants must look up to visually access the upper visual fields, where adult faces and bodies typically are, this process is a little easier for walking toddlers. However, height differences between toddlers and typical adults still mean that faces are much harder to access for toddlers than they would be for a much older child or adult (Kim et al., 2020). So, at birth, faces are very accessible for babies so long as adults choose to interact with them. Faces and some elements of bodies become less accessible for crawling toddlers, and then slowly become more accessible as toddlers learn to walk.

Beyond pragmatic constraints, motor skills are likely to change the visual input in another way: walking toddlers are much more independent and have more control over their

visual input. Theoretically, it has been proposed that the actions of infants and toddlers directly shape the input according to their motivation, creating a “curriculum” for learning about the world; this would include faces and bodies (Fausey et al., 2016; Jayaraman et al., 2017; Smith et al., 2018). Walking toddlers can move efficiently and independently with hands left free for exploration according to their interests and motivations. And indeed, there is evidence that the face preference (or preference for face-like stimuli) present at birth declines somewhat over developmental time (Johnson et al., 1991). So, this may be consistent with a face decline in the view of toddlers.

There is indeed evidence that infants directly shape their own perceptual input in line with these interests and motivations in the domain of object learning. Infants have been demonstrated to shape the visual input by manipulating the objects that they are viewing in a way that optimises object learning (Bambach et al., 2016). While independent control of the head and arms allows for some degree of this, the ability to independently traverse the environment with hands left free for manual exploration introduces a much greater degree of direct control over the input. So, when infants can walk, their ability to change the input directly means that this input will change.

Finally, the ability to locomote more independently may shape the visual input by allowing toddlers to change the behaviours of those around them. Premobile infants can only elicit responses from others in limited ways— for example, the infant cry is a form of communication that can elicit behaviour in adults around the infant (Murray, 1979), and there is

some research to suggest that there exist distinct forms of cry to elicit different responses from parents (Brennan and Kirkland, 1982). Even very subtle nonverbal cues from premobile infants can serve as a form of communication (Als, 1977; Messinger, 2002; Messinger and Fogel, 1998) to elicit a range of responses from these caregivers (Beebe et al., 2016). However, walking toddlers can directly engage in a much wider range of behaviours that can elicit responses from others. Toddlers can break a wider range of social rules, which shapes corresponding parental response (Dahl et al., 2014; Smetana, 1989). Toddler and parent behaviour is bidirectionally, causally related (Del Vecchio et al., 2010; Reid et al., 1994). Independently walking toddlers can therefore change their environments by changing the behaviours of those around them.

So, as infants become walking toddlers, the nature of the environmental visual input likely changes greatly. This is a result of pragmatic factors such as bodily posture alongside the more far-reaching ability of toddlers to shape their visual inputs both directly and indirectly.

Prior Quantification of the Input

There have been some previous attempts made to quantify the frequency with which infants and toddlers see faces and bodies using egocentric head-mounted cameras (Fausey et al., 2016; Jayaraman et al., 2015, 2017; Sugden et al., 2013). This egocentric perspective research allows for characterization of the elements of the environment that are within the view of an individual; simply knowing how many people are present from a third-person perspective would not allow for inferences to be made on the degree to which the infant or toddler actually sees faces or bodies. However, evidence of this kind is relatively scarce, possibly

because of the methodological difficulties that arise in collecting it, from recruiting a wide sample of compliant children through to coding very rich video data (Borjon et al., 2018; Long et al., 2020).

The prior research that does exist would suggest a decline in the frequency of faces specifically from birth to toddlerhood. Sugden et al., (2014) examined 14 Canadian 1- to 3-month-olds, finding that faces are frequent, being present in roughly a quarter (25% and 26% respectively) of infant-perspective frames. Jayaraman et al (2015) establish a decline in faces over developmental time from 1- to 11-months. Fausey et al. (2016) extended these findings to new age groups, examining changes in this input in 34 infants from the USA ranging in age from 1-month to 2-years, finding a decline in the presentations of faces and an increase in presentations of hands (both their own and other) with age. The proportion of the recorded child-perspective frames containing faces was 0.24 (24%) for the youngest infants and 0.075 (7.5%) for the oldest toddlers. So, the frequency of faces in the view of young infants is a consistent finding, with evidence for a decline in face frequency as infants become toddlers.

Extending this research on the change in visual statistics with age, Fausey, Jayaraman and Smith (2016) examined the frequency of faces and hands in views of 34 infants and toddlers ranging in age from 1-month to 2-years, finding that as infants become toddlers there is a decline in the presentations of faces (24% of the input for the youngest infants, 7.5% for the oldest) and an increase in presentations of hands (from less than 10% of the input to around 25% of the input). Jayaraman, Fausey and Smith (2017) examined egocentric perspective video

recorded from a similar sample of 36 infants, including one to 24-months-old, finding that frames with faces was 23.2% for the youngest infants and 8.83% for the eight oldest infants (18-months and older).

However, there are several limitations in using this data to draw definitive conclusions about the visual input of toddlers with respect to faces and bodies. Firstly, the prior research outline here includes only a limited number of individuals in each age group. Toddlers around the 24-month mark are especially underrepresented. Only 13 of the 36 participants in the research of Jayaraman et al. (2017) were 12-months or older, with 7 older than 16-months. Given that 13-months is the average onset of walking, that infants may not walk consistently from this point, and that we would expect the visual input of walking infants to differ from that of crawling infants, this under-representation of older infants who we would expect to traverse the environment upright is significant.

Secondly, prior research has typically only considered real-life presentations of the faces and bodies of other individuals. The previous attempts to characterise the infant environment (e.g., Jayaraman et al., 2017) do not consider presentations outside of these categories (e.g., presentation from the self or media) to be presentations of faces and bodies, while it may be that these actually provide an important source of information and learning. By 12-months, prior to explicit self-recognition (Stapel et al., 2017), infants have formed some representations of their own faces (Nitta and Hashiya, 2021). Further, infants use knowledge of their own bodies and others' reciprocally (Meltzoff and Marshall, 2020), so visual input from the bodies of others

is unlikely to shape development in isolation. Socioemotional environmental input from multimedia is also important– toddlers actively use this source of social data, imitating social models observed in these modalities (Meltzoff, 1988) and learning from media socio-emotional expressions (Mumme and Fernald, 2003). So, visual input from the media and from mirrors should not be excluded. Given that this exposure in particular may be self-directed, this is an important avenue of input for visual learning in mobile toddlers in particular.

Finally, because of the very rich nature of naturalistic head camera data, a variety of coarser coding methods based on sampling have been used. For example, Jayaraman et al. (2015, 2017) sample images once every five seconds. While in these papers analysis of this coding frequency would suggest that it is sufficient, it is possible that coarse coding of this nature would misrepresent the statistics of the visual input.

Understanding the Face Decline in the Context of Adult Visual Input

There is very little data on the visual input of adults with respect to faces and bodies. But understanding the statistics of the adult visual input could help us to understand and contextualise the possibly infrequent faces within the toddler view. Whether the decline in faces and increase in elements of bodies within the egocentric view is unique to toddlerhood or continues into adulthood would be essential in building models explaining this phenomenon.

There are many potential explanations for the decline in faces alone from infancy through to toddlerhood, but broadly they can be separated into two categories: pragmatic and

motivational. In terms of the former, it could be that toddlers are *less able* to access faces in their environments, as a result of a mixture of underdeveloped motor control and a plane of view not level with the faces most adults in the home (Kretch and Adolph, 2015); in terms of the latter, it could be that toddlers are *less interested* in accessing faces in their environments. Knowing how the toddler visual input compares with that of a caregiver within the same environment is essential here.

There is surprisingly sparse evidence on the day-to-day visual input of adults in terms of faces and bodies, though the term “face diet” for the former was coined over two decades ago (Rhodes et al., 2003). Recently, there has been some quantification of how frequently adults see faces, with Oruc et al. (2019) finding that in egocentric recordings taken from 30 adults undergoing daily activities, on average 22.5% (SD = 12.5%) of frames contained faces. This is on par with the face frequency found in prior research for the youngest infants, suggesting that perhaps the decline in faces into toddlerhood is temporary.

However, this data is not sufficient for our purposes. Firstly, without knowing the frequency with which adults see bodies, we cannot say whether adults see faces specifically more frequently than toddlers, or people as a whole. Secondly, adults in the presence of their dependent children may attend to— prior research has demonstrated a relative caregiver preference for attending to the infant face as compared to other elements of the environment (Franchak et al., 2018). This same paper also suggests that ease of motor access to faces is associated with looking behaviour towards them, and infant faces could be easier to access for

adult caregivers than vice versa. So, the looking behaviour and thus visual input of caregivers may be shaped by unique motivational and pragmatic factors when interacting with dependent toddlers. Finally, the prior adult data was recorded during the everyday lives of working adults, which may be an environment in which faces are more easily accessible because of the number of same age/height interaction partners.

Quantifying the frequency of faces and bodies in the view of adults and toddlers in the same environment would allow us to understand whether infrequent faces as compared to bodies within the visual input is unique to toddlers, or whether this is an effect that generalises to adults within the same environment. This could help us to assess the developmental significance of the face decline over infancy and toddlerhood, assuming that we replicate infrequent faces in the toddler view here.

Aims and Hypotheses

The aim of this study is to quantify the frequency with which faces and bodies are present in the egocentric view of toddlers and caregivers as they freely navigate the home environment. This will allow us to understand factors that may facilitate infant socioemotional development, and differs from previous research in four key ways: firstly, we will examine data from a larger sample of toddlers than has been previously studied, focusing on reliably quantifying the visual experiences at 2-years-olds; secondly, we seek to compare the frequency of faces and bodies in the visual field of both toddlers and their caregivers; thirdly, we seek to

evaluate methods for handling rich, real-world data through comparing the validity of frame-by-frame analysis of video data as compared to previously used five-second sampling and evaluating the viability of using an automated method of coding in order to label naturalistic videos frame-by-frame; finally, we define a “face presentation” as an instance of any real or photorealistic internal feature of the face (eyes, nose, mouth), regardless of source (self, other, media), and we define a “body presentation” as any instance of an identifiable real or photorealistic body part regardless of source (self, other, media).

We believe that a combination of frame-by-frame coding (allowed by our computer vision models) better representing the statistics of the visual input and the wider definition of face and body may lead to a modest increase in the percentage of coded toddler-view frames containing faces as compared to prior research. Specifically, we think that sampling frames may not well represent the very infrequent faces in the toddler view. However, this percentage is still very unlikely to eclipse the body percentage in the toddler view. Pragmatic elements of height differences do seem to limit access to faces for walking toddlers to at least some degree (Kretch and Adolph, 2015) and there is prior evidence to suggest that faces are more accessible for caregivers than toddlers in walking toddler - caregiver dyads engaged in lab-based free play (Franchak et al., 2018).

This paper is highly novel in examining comparable face and body frequencies in the caregiver view within the naturalistic home environment. In terms of the caregiver-view faces, we expect that the nature of the home environment for these families (fewer individuals, and

those individuals varying in height) may lead to a lower face frequency in the caregiver view than the adult face frequency established by Oruc et al. (2019), which established that 22.5% (SD = 12.5%) of frames recorded in a variety of environments contain faces. However, given that it would be easier for caregivers to access the faces of toddlers than vice versa (it is easier to look down from height), we still expect a higher face frequency in the caregiver view than in the toddler view. Given that there is no prior research on body frequency in the adult view, and bodies are more frequent in the view of upright toddlers than pre-mobile infants, we expect caregiver body frequency to be roughly equivalent to that in the toddler view.

So, with the above in mind, we hypothesise that: 1) The frequency of faces in the toddler view will be lower than the frequency of bodies in the toddler view; 2) The percentage of frames containing faces in the toddler view will be higher than that specified for infants in the second year of life in prior research, possibly as a result of frame-by-frame coding; 3) face frequency in the toddler view will be lower than face frequency in the caregiver view; 4) Though, the face frequency in the caregiver view may be lower than that specified for adults previously; 5) The body frequency in the caregiver view will be comparable to the body frequency in the toddler view.

Methods

See general methods section for full details.

Data Analysis Strategy

To maximise the sample size and statistical power here, we first consider the full $N = 26$ for the presence of faces and bodies in the toddler view, and then separately, for the comparison of the toddler and caregiver view, the $N = 24$ participants with corresponding caregiver footage. As a preliminary exploratory assessment of whether the 5-second sampling of prior research may have misrepresented the frequency of faces, we also compared the face frequency in the toddler perspective at different simulated sampling frequencies.

Results

Incidence of Faces and Bodies in the Toddler View

Overall, for the 26 families with data from the child perspective, there were over 6.6 million child-perspective frames (approximately 131 hours) coded for the presence of faces and bodies. This footage was recorded over the course of 145 separate recording sessions, with an average of 5.6 sessions per participant. On average across participants using frame-by-frame coding, 7.98% of frames ($SD = 4.75$) contained faces while 53.73% of frames ($SD = 6.70$) contained bodies. A Wilcoxon signed rank test showed that there was a significant difference ($Z = 4.46$, $p < 0.001$) between face and body frequency in the toddler view.

Testing the Adequacy of Coarser Sampling Rates for Face Frequency

The percentage of frames containing faces (taken over only included, camera-on frames) was compared across these five sampling frequencies. These sampling frequencies were generated, averaged and compared across all videos, rather than across participants. See Table 1 for a summary of face and body frequency by sampling rate.

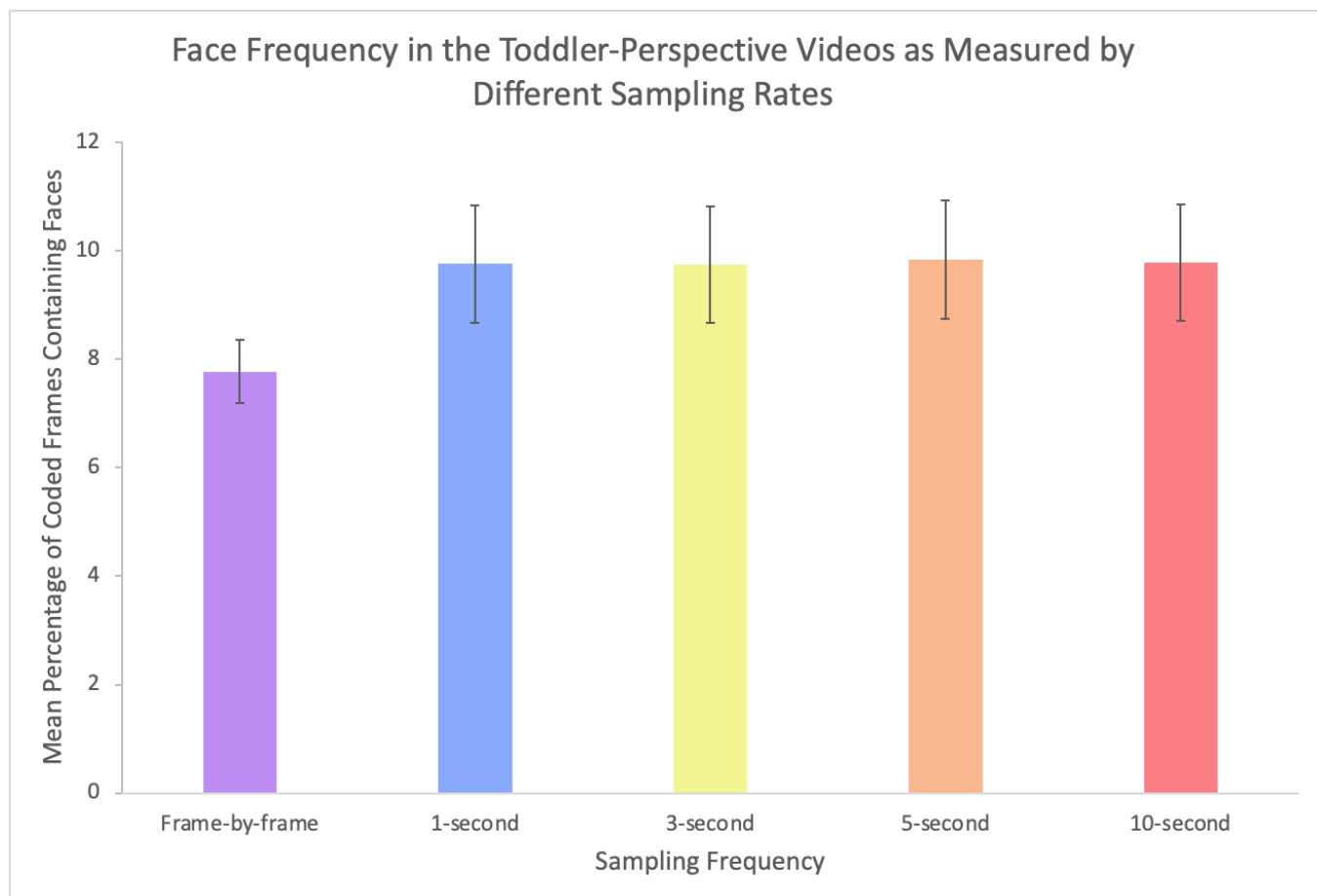
Table 1*Face and Body Frequency by Sampling Rate*

Sampling Rate	Face Frequency (Toddler View)	Face Frequency (Caregiver View)	Body Frequency (Toddler View)	Body Frequency (Caregiver View)
Frame-by-frame	7.77% (SD = 6.98)	14.53% (SD = 11.41)	53.67% (SD = 11.91)	72.30% (SD = 14.08)
1-second	9.75% (SD = 13.01)	15.40% (SD = 12.40)	53.69% (SD = 11.97)	72.30% (SD = 14.07)
3-second	9.73% (SD = 12.89)	15.40% (SD = 12.43)	53.70% (SD = 11.98)	72.31% (SD = 13.96)
5-second	9.83% (SD = 13.02)	15.34% (SD = 12.56)	53.61% (SD = 11.99)	72.33% (SD = 14.00)
10-second	9.78% (SD = 12.87)	15.16% (SD = 12.52)	53.63% (SD = 11.97)	72.21% (SD = 13.93)

Note: This table provides a summary of the face and body frequencies in the toddler and caregiver views at different sampling frequencies— 1 frame per second, 1 frame every 3-seconds, 1 frame every 5-seconds, and 1 frame every 10-seconds.

Figure 1

A bar chart to show the face frequencies from the toddler perspective at different sampling frequencies.



Note: This bar chart displays the mean percentage of frames containing faces from the toddler perspective across videos at different sampling frequencies, from frame-by-frame through to every ten seconds. The error bars represent standard errors across videos.

The face frequencies in the toddler perspective were then compared at different sampling rates (see figure 1). To test whether face frequency in the toddler perspective varied with sampling frequency, a repeated-measures ANOVA was used to test the effect of sampling

frequency on the percentage of faces identified in toddler-perspective footage. The Greenhouse-Geisser adjusted values were used. There was a statistically significant effect of sampling frequency on the percentage of faces identified in the toddler view ($F(1.108, 158.414) = 6.137, p = .012$). However, post hoc pairwise comparisons using the Bonferroni correction revealed no statistically significant differences between any individual levels of sampling frequency. This may indicate a spurious main effect.

Comparison with Caregiver-View Frequency

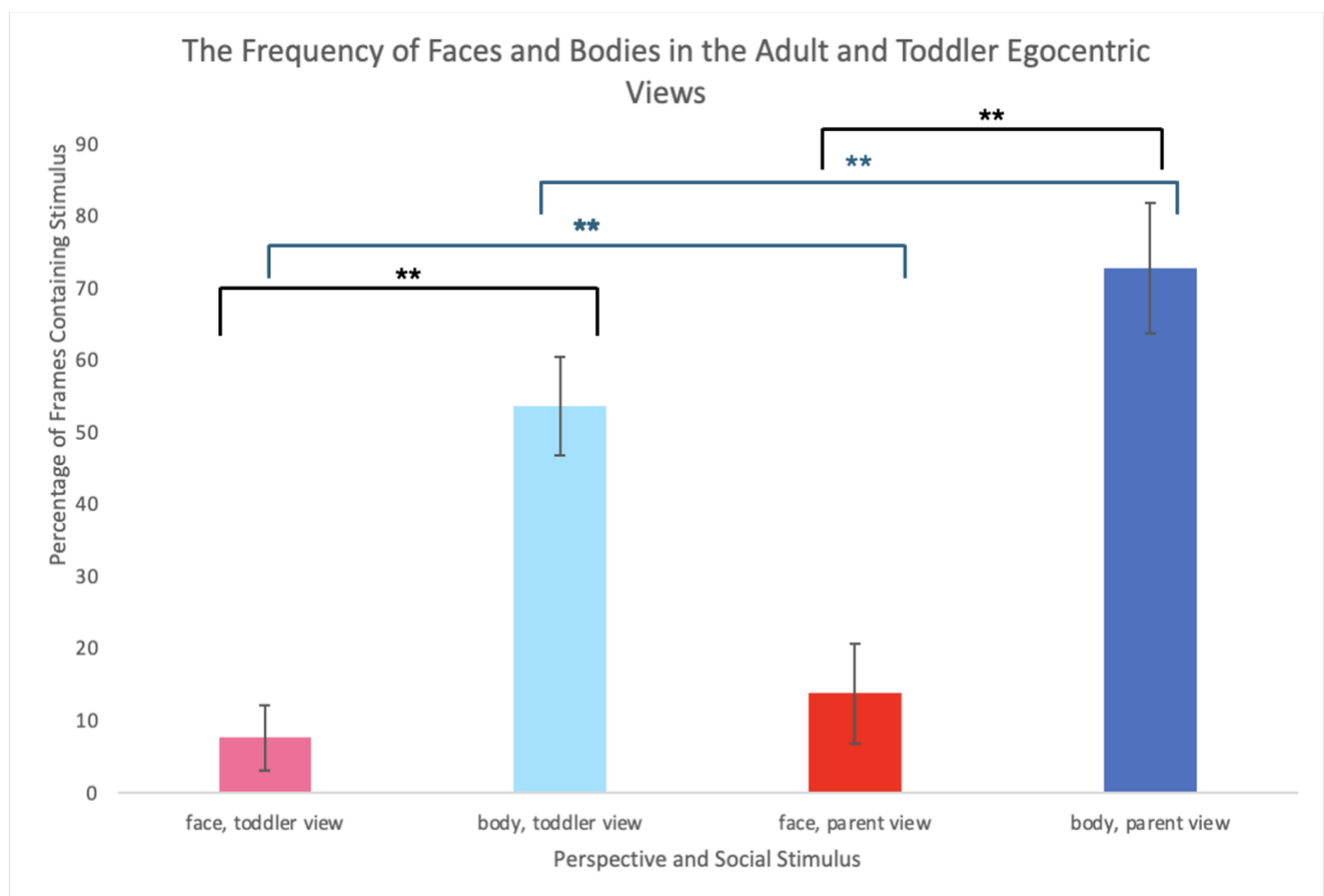
Of the 26 families that recorded from the child perspective, 24 recorded a total of 143 recording sessions from the caregiver perspective: this amounted to over 6.8 million frames codable from the caregiver perspective (approximately 136 hours). There were approximately 6.5 million frames (129 hours) of corresponding toddler-perspective footage. On average across participants who recorded both perspectives, using frame-by-frame coding, 13.81 % (SD = 6.92) of caregiver-perspective frames contained faces, and 72.88% (SD = 9.06) of frames contained bodies; and on average, 7.67% (SD = 4.49) of toddler-perspective frames contained faces, and 53.71% (SD = 6.87) of frames contained bodies (Figure 2).

A 2 (perspective: toddler, caregiver) x 2 (stimulus: face, body) repeated measures ANOVA revealed statistically significant main effects of perspective ($F(1, 23) = 123.842, p < .001$) and stimulus ($F(1, 23) = 939.954, p < .001$). These were further qualified by their interaction, ($F(1, 23) = 52.523, p < .001$). In order to disentangle this interaction, paired-samples t-tests were conducted to compare the effect of perspective (toddler, caregiver) and stimulus (face, body) on

frequency within the egocentric view. After applying a Bonferroni correction, the significance level was set at $p < 0.0125$. These post-hoc analyses showed that toddlers' egocentric views of their environment contain a significantly higher percentage of bodies ($M = 53.71\%$, $SD = 6.87$) than faces ($M = 7.67\%$, $SD = 4.492$), $t(23) = 28.07$, $p < .001$.

Figure 2

A bar chart to show the face and body frequencies from the toddler and caregiver perspective.



Note: This bar chart displays the mean and standard deviation percentage of frames containing faces and bodies from the child and caregiver perspectives. ** indicates significance at the $<.001$ level.

The same pattern of results was observed for the caregivers' egocentric views, with a higher percentage of frames containing bodies ($M = 72.88\%$, $SD = 9.06$) than faces ($M = 13.81\%$, $SD = 6.92$), $t(23) = 26.95$, $p < .001$. Further, the post hoc comparisons revealed that the egocentric views of toddlers contained lower percentages of faces ($M = 7.67\%$, $SD = 4.492$) than the caregivers' views ($M = 13.81\%$, $SD = 6.92$), $t(23) = 4.99$, $p < .001$. The toddler view also had lower percentages of bodies (53.71% , $SD = 6.87$) than the caregivers' views ($M = 72.88\%$, $SD = 9.06$), $t(23) = 11.68$, $p < .001$).

The Frequency of Faces in the Caregiver Perspective

Given that prior research on the infant and toddler visual input does not state standard deviations for each age group, direct comparisons were not possible. However, Oruc et al. (2019) do report these figures (face frequency $M = 22.5\%$, $SD = 12.5\%$) for the 30 adults that they recorded the visual input of. We compared this to the frequency of faces recorded in the view of the 24 caregivers here ($M = 13.81\%$, $SD = 6.92$) using a between-subjects t-test. The face frequency in the caregiver view that we report is significantly lower than the face frequency reported by Oruc et al. (2019), $t(52)=3.05$, $p = 0.004$. Given differences in research design, this comparison should be treated as exploratory in nature and interpreted with caution.

Discussion

We aimed to characterise the visual input of toddlers and their caregivers as they freely navigated the home environment by quantifying the frequency with which faces and bodies were present in their egocentric views. This data was coded using computer vision models to determine the frequency of faces and bodies in the toddler and caregiver perspectives.

We confirmed our first hypothesis: bodies were significantly more frequent than faces in the toddler's view. This is consistent with our expectation that while toddlers encounter relatively few faces, bodies are frequent features of the visual input. However, contrary to our second hypothesis, we did not find a higher frequency of faces in the toddler's view compared to previous research. While direct statistical comparison was not possible, our results closely match the range reported in earlier studies (7.5% to 8.83%). We found that in the toddler view, 7.98% of frames ($SD = 4.75$) contained faces while 53.73% of frames ($SD = 6.70$) contained bodies. With a wider sample of confidently walking toddlers in the second year of life, we have established more evidence for infrequent faces in the toddler view. We also find that infrequent faces are unique to the toddler perspective. While it is surprising that a wider definition of face and frame-by-frame coding did not result in a greater incidence of faces in the toddler view, to an extent this finding makes sense considering our analysis comparing frame-by-frame to 1-second and 5-second interval coding. We were able to code every frame using computer vision methods, and then we were able to code simulated more coarse coding frequencies. While face frequency did vary by sampling frequency, there were only very small absolute differences in

the face data here regardless of coding frequency used; these small differences are unlikely to markedly change face frequency.

Importantly, we have established the adult visual input within the same, potentially face-scarce environment of the family home. Our third hypothesis was supported, with face frequency being significantly higher in the caregiver's view compared to the toddler's. Not only this, but we found evidence for our fourth hypothesis: while comparisons should be interpreted with caution given the different methodologies used, face frequency in the caregiver's view was lower than that reported for adults in other settings (Oruc, 2019; $M = 22.5\%$, $SD = 12.5\%$). However, our fifth hypothesis was not confirmed; bodies were observed more frequently in the caregiver's view than in the toddler's. We have established that infrequent faces are somewhat unique to toddlers within this environment, with adult caregivers seeing faces significantly more frequently— however, somewhat less frequently than adults out in a variety of other environments such as the workplace. This suggests that the face decline from infancy to toddlerhood is temporary, with either unique pragmatic or motivational constraints on the toddler's socioemotional visual input with respect to faces. Further, we have established that toddlers see all human body parts, not just faces, less frequently than their caregiver counterparts.

One methodological reason for the low frequency of captured faces could be discrepancies between the visual input captured by the camera and the visual input to which infants are exposed. Necessarily, the camera captures a limited range of input that does not

correspond precisely to the visual field of the infant. This may mean that observation of faces, especially faces high above the toddler (e.g., from a standing adult) at proximity, when this discrepancy is the most impactful, is not always recorded by the head cameras. This is a problem with both this research and all prior attempts to quantify the nature of the visual input and is compounded by the lack of meaningful standardisation of head-mounted recording apparatus for infants. Further, head movements and eye movements may not always identically correspond-- though prior research suggests close alignment between head and eye gaze direction during visual exploration (Schmitow et al., 2013). There is also an error rate inherent to any form of video coding. We utilised a rapid, automated algorithmic coding method to code video footage for faces and bodies. While carefully validated using a subset of data, any errors will undoubtedly have had some impact on the results. However, this was a more time-efficient approach than manual coding, and may have reduced the overall error rate, given that manual coding can be a mentally fatiguing task and require many raters for good reliability.

Another methodological limitation of this research is also that the egocentric view of toddlers and caregivers was only recorded within the home environment. The home environment contains relatively few individuals, and for toddlers and caregivers, a large proportion of these individuals in the home environment are of markedly different heights. This could mean that faces in the home environment are less easily visually accessible, especially for toddlers, than they might be in other environments—for example, preschool. We can see in comparisons of the caregiver view to prior research on adult face input that the home

environment is associated with scarcer faces in the adult view; the same could be true for toddlers.

However, the home environment is likely one of if not the most important learning environment in infancy, toddlerhood and early childhood based on the proportion of time spent there alone. The quality and nature of this early home environment has been associated with everything from language and literacy outcomes (Pinto et al., 2013) to later general cognitive outcomes (Biedinger, 2011) to social outcomes (Foster et al., 2005). Time spent with parents makes up a great proportion of early childhood, especially time with mothers: on average, mothers in the UK take 9.3 months of parental leave (Kaufman and Grönlund, 2021); further, mothers in the UK spend on average around 150 minutes with their children per day (Dotti Sani and Treas, 2016), which (while there is no direct measure) is likely to be higher for younger children and toddlers, who are not yet in full-time education. For socioemotional development, the home environment can be a source of highly influential protective factors (Whittaker et al., 2011) or risk factors (Deater-Deckard et al., 2009), often outweighing the influence of educational environments (Bobbitt and Gershoff, 2016). Not only this, but by quantifying visual input within the home environment, we are able to make direct comparisons with prior developmental research characterising the toddler and infant visual inputs.

It may be that in reality, the views of mobile and walking toddlers contain a relatively low incidence of faces. This dataset was very rich, containing many scenes taken from a much wider sample of walking toddlers than previous data. With this sample, we replicate the

infrequent faces in the toddler view and find data suggestive of this being unique to toddlers as compared to caregivers in the same environment. Given that pragmatic constraint of number of people (and number of people of the same height) in the environment should apply equally to toddlers and their caregivers, this would suggest that there is something in particular about the way in which toddlers interact with their environments that produces this relative scarcity of faces and reduction in overall body frequency. This could be motivational— that is, toddlers are less interested in people and their faces and bodies than adult caregivers. However, it is also possible that pragmatic constraints presented by height differences more significantly affect toddlers than vice versa.

In terms of pragmatic constraints, there is evidence that faces are relatively difficult to access for walking infants as compared to bodies (Kretch and Adolph, 2015) and that faces are more accessible for caregivers as compared to toddlers within the same environment (Franchak et al., 2018). So, it could be that the smaller height and lower visual perspective of toddlers leads to infrequent faces relative to bodies, particularly within the toddler view.

Motivational factors may also play a role in infrequent faces within the toddler view. It could be that with a transition from learning about the structure of faces per se to extracting information from faces on topics like how to best explore the environment (Feinman, 1982; Klinnert et al., 1986) less frequent or prolonged looking at faces is necessary. This would be indirectly supported by research indicating a general increase in cognitive processing speed over developmental time (Hale, 1990), which may allow for more efficient extraction of necessary

information from faces. It is also supported by research suggesting the emergence of joint attention relying primarily on non-face cues between parents and infants in later infancy (Yu and Smith, 2013) -- in these interactions, faces are infrequently accessed to complement information on attended objects extracted from hands. This would explain why bodies are so frequent in the view of toddlers as compared to faces, and why prior research would suggest that this shift in the visual statistics represents a particular increase in the frequency of hands in the toddler view (Fausey et al., 2016). Our research would be consistent with toddlers relying on bodies, hands and other non-face cues for engaging socially with other actors in the same environment.

Since toddlers are still learning about faces (Chung and Thomson, 1995), it must be that either this learning is driven by very infrequent presentations of faces, that toddlers have already received most of the necessary substrate for development, or that innately driven maturational processes play a role in driving face processing development, as proposed by authors including Carey (1980). To fully assess these proposals, it would be necessary to understand just how toddlers attend to the infrequent faces that they do see, as well as the context in which they see these faces. However, a role for maturational processes in development would be consistent with the very infrequent faces that toddlers see.

Variability and Future Research Directions

It was clear that even within toddlers, the statistics of the visual input were highly variable. This could be driven by many factors, from differences between home environments to

motivational differences between individual toddlers. Future research should explore the sources and significance of this variation in the visual input. This could be key in understanding how the visual input of infants and toddlers shapes their development. Further research on the adult visual input is also required, including research exploring sources of variation in the adult visual input ranging from individual to situational.

As well as this, further research is required to establish what input is available in faces and bodies within the environment and what information toddlers actually extract from these presentations. This may help to explain the changes in face and body density in the egocentric view that take place with developmental time. For example, it could be that toddlers extract specific information on the environment from the bodies (and especially hands) that they frequently observe (Fausey et al., 2016). Older toddlers are more able to freely explore objects manually (Karasik et al., 2011), which may mean that they look to the bodies of others in the environment as a source of learning and information to inform this exploration. Knowing how and when toddlers look to faces will be an essential focus for future research,

Beyond the toddler perspective, there is a need for more comprehensive data on adult visual input. While it seems from our research that adults view faces much more frequently than toddlers, the specific factors that drive and shape the frequency of faces and bodies in the adult view remain unclear. Adults are able to process faces far more efficiently than children and toddlers (De Sonneville et al., 2002), so intuitively we would expect adults to need to view faces for less time overall than toddlers. So, understanding the task-dependent motivations and

needs of the adult observer may also be essential in building models of how, why and when visual data from faces is intentionally accessed and used. Whether the increased frequency of faces in the caregiver view as compared to the toddler view is driven by pragmatic factors, motivational factors or both will be an essential question for future research.

Conclusions

In conclusion, our study replicates previous findings on the low frequency of faces compared to bodies in toddlers' views, even when analysing data taken from using a wider sample of toddlers, frame-by-frame coding, and a broader definition of faces. This study is novel in comparing this to the visual input of caregivers within the same environment, finding that caregivers see both faces and bodies more frequently than toddlers. However, we do find that adult caregivers in the home environment see faces less frequently than adults navigating a variety of environments do according to previous research. It is unclear precisely why faces are so comparatively infrequent in the views of toddlers, though both pragmatic and motivational factors may play a role. Future research should examine what information may be available to and extracted by toddlers from the faces and bodies within their view.

Factors Affecting the Frequency of Faces and Bodies in the Egocentric View of Toddlers

Emma J Jackson, Bruce Muller, William Smith and Elena Geangu

This paper has not yet been submitted for publication

Abstract

Visual processing of faces and bodies is fundamental, but there is diversity in proficiency between individuals. This could be due to differences in early visual input, which may be shaped by many internal and external factors. Faces and bodies are an important component of the toddler visual environment, but this input could vary. Variation could come about because of many factors, but two likely candidates are aspects of the environment external to the toddler relating to the presence of siblings and internal individual differences in temperament. We sought to establish factors associated with the frequency of faces and bodies in the egocentric view of toddlers, specifically examining the presence of siblings and toddler temperament. We analysed data from 26 families in and around York (UK) who recorded the toddler visual perspective at home using egocentric cameras for an average of 5.20 hours each. While there is no prior research in this domain, temperament and number of siblings have previously been associated with social functioning, and so possibly are related to individual differences in the visual input. We hypothesised that toddlers with siblings would see faces and bodies more frequently than only children. We also hypothesised that both concurrent and longitudinal temperament would correlate with the frequency of faces and bodies in the toddler view. Our findings revealed no significant effects of sibling presence on the frequency of faces or bodies in the toddlers' view. Additionally, no robust relationships were found between temperament and the frequency of either faces or bodies. Future research should explore factors shaping the visual input further with larger samples and consider finer-grain aspects of the visual input, but the statistics of face and body input appear to be unrelated to these individual differences.

Keywords: naturalistic, face, body, toddler, visual

Factors Affecting the Frequency of Faces and Bodies in the Egocentric View of Toddlers

Processing of faces and bodies is a foundational set of skills that infants must acquire, and diversity in these skills is present (Yovel, Wilmer and Duchaine, 2014). These individual differences are theoretically and clinically significant, with variation at the extreme associated with a range of clinical disorders (van Boxten et al., 2017; Sasson, 2006). One potential explanation for this variation in later ability to extract information from faces and bodies lies in variation in the frequency with which individuals are exposed to faces and bodies early in life. While this variation is likely to be shaped by many factors, at least one factor, age, is known to shape the frequency of face and body presence in the egocentric view (Jayaraman et al., 2015; Jayaraman et al., 2017). Knowledge of which other factors shape the socioemotional input would be essential in building a model of how antecedent variables such as temperament and family environment are associated with later socioemotional communication.

Here, we will focus on two factors in particular: presence of siblings and temperament.

Temperament has been associated with differences in social competence (Calkins and Mackler, 2011) and differences in face attending behaviour (Martinis et al., 2012). Similarly, sibling status has been associated with differences in social competence (Downey et al., 2015), and there is evidence that presence of a sibling shapes face processing proficiency (Cassia et al., 2013). So, these factors are likely candidates for variables related to and possibly shaping the visual input of toddlers with respect to faces and bodies.

Toddlerhood represents a vital period for socioemotional development. It is also a period of increased independence. By this age, most toddlers are proficient walkers, which fundamentally alters their interaction with their environment. Unlike infants, who are often presented with faces while prone, walking toddlers must actively engage in visual exploration, including looking up to see adult faces in an effortful action (Karasik et al., 2011; Kretch et al., 2014). This shift in perspective likely contributes to the relative scarcity of faces in the toddler's visual field (Fausey et al., 2016; Jayaraman et al., 2015, 2017; Sugden et al., 2014). In chapter 1, we established both further evidence for this relatively infrequency of faces in the toddler view as compared to bodies and we also noted great individual variation in the frequency with which toddlers see faces and bodies. Critically, as toddlers gain mobility, they may exercise greater autonomy over what they see, increasing the potential for individual differences in visual experiences. These differences may stem from intrinsic motivations or external environmental factors that either facilitate or hinder access to faces and bodies (Smith et al., 2018). These differences in input then in turn may play a significant role in shaping unfolding socioemotional development (Oakes, 2023).

This paper will examine factors that may correlate with individual variation in the toddler visual input of faces and bodies. Here, we will focus on two factors in particular: the presence of siblings (i.e., whether a toddler has siblings or is an only child) and temperament, which refers to individual differences in tendency to react and tendency to regulate these reactions (Garstein and Rothbart, 2003; Goldsmith, 1982; Rothbart and Bates, 2007). Temperament has been

associated with differences in social competence (Calkins and Mackler, 2011) and differences in face attending behaviour (Martinis et al., 2012). Similarly, sibling status has been associated with differences in social competence (Downey et al., 2015), and there is evidence that presence of a sibling shapes face processing proficiency (Cassia et al., 2013). So, these factors are likely candidates for variables related to and possibly shaping the visual input of toddlers with respect to faces and bodies. We will outline what is currently known about the visual input of faces and bodies as well as developmental trajectories for face and body processing that may be influenced by this visual input.

Prior Research on the Socioemotional Visual Input

First, we will outline current research on the frequency with which faces and bodies appear in the toddler view. There does exist some quantification of this frequency (Fausey et al., 2016; Jayaraman et al., 2015, 2017). This research typically quantifies and characterises the faces and bodies in the views of infants and toddlers using egocentric head-mounted cameras that capture scenes in front of the infant or toddler face. Though this research does not focus on or even explore individual differences in the input, there is clear evidence for wide individual differences in the visual input that can be seen in this existing research and in our data from chapter 1. Here, we will outline what each piece of extant data can tell us about potential variation in the environmental input.

Sugden et al. (2014), using this head-mounted camera methodology, established that young infants (30 infants 1- to 3-months) frequently see faces, with about 25-26% of the input containing faces. While this paper does not report individual variation, they do present the range of face identities present in each video— and there is wide variation in this. While on average there were 2.19 face identities in the infant view per video, this ranges from 1 to 19 face identities even when excluding very high face density videos (e.g., videos taken as the infant is moved through a crowd). There was also a wide range in the duration of face presentations— while the average face presentation lasted 4.12 seconds, there was a wide range from 1.89 to 8.90 seconds. This variation in face input extended to specific aspects of face identity. For example, 70% of time spent viewing faces was spent viewing female faces, but this varied widely, from 7% to 100%. So, though the focus here is on very young infants, and there is no direct quantification of the variability in time spent viewing faces overall, this paper would suggest that there is a great degree of variability in specific aspects of the visual face input of young infants.

There is further research including additional information, both in the form of characteristics of the visual input of bodies and providing data on a wider age range of participants. Research by Fausey et al. (2016) and Jayaraman et al., (2015, 2017) provides evidence for decline in the presence of faces in the egocentric view from early infancy through to toddlerhood. Jayaraman et al. (2015) provided evidence for a decline in face frequency in the infant view from approximately 25% of the input containing faces at 1-month to 8.3% at 11-months. Fausey et al. (2016) used similar methodologies but a sample extending into the

second year of life, finding that from 1 month to 2 years, the proportion of faces seen by infants decreased from 24% to 7.5%, while the presence of hands increased. Jayaraman et al. (2017) also found evidence for decline in face visibility, from 23.2% at 1-month to 8.83% for infants 18-months and older. However, they found no age-related decline in the overall presence of other individuals in the infant view. This indicates that while faces become less frequent in the visual field as young infants become toddlers, this effect is specific to faces and does not reflect a general decline in the presence of others.

While the authors of these papers do not report directly on individual variation in the frequency of faces and bodies present, there does appear to be significant variation around each age-point. For example, the presented data in Jayaraman et al. (2017) suggest that for the five toddlers around 24-months, the face frequency ranges from approximately 4% of the input to 20% of the visual input, while the overall person frequency ranges from approximately 32% to 43% at this age. So, in toddlerhood, there is notable variation in face and body presentations, but especially in the frequency of the scarce face presentations.

This is in line with the previous analyses carried out on our dataset, suggesting that faces are infrequent in the views of toddlers and that the toddler visual input is highly variable (see chapter 1). In examining face and body frequency in the views of those 26 toddlers whose data we continue to analyse here, we found that across participants, $M = 7.98\%$ of frames ($SD = 4.75$) contained faces while $M = 53.73\%$ of frames ($SD = 6.70$) contained bodies. This suggests a great degree of variation, particularly proportionally for the frequency of faces within the toddler

view. Explaining the degree of variation that we observed could be crucial in understanding the developmental significance of the face and body visual input, especially when it comes to shaping individual differences in socioemotional developmental trajectories.

However, there is currently no research exploring what factors, if any, may be related to this variation in the visual input. This is a necessary first step in building models of how variation in the visual input may come about and what its significance may be. The wide range in the frequency of faces and bodies in the toddler views could be developmentally significant in shaping later variation in developmental trajectories. Further, the variations in visual input could be related to pre-existing individual differences in developmental.

This is especially true for toddlers. Studying individual differences in the visual input of toddlers is crucial because, as infants' and toddlers' motor abilities develop, they are increasingly able to shape their own visual experiences. There has been prior suggestion that infants create an individual "curriculum" for their visual learning, including their learning about faces and bodies (Fausey et al., 2016; Jayaraman et al., 2017; Smith et al., 2018). There is evidence that infants actively optimise their visual input in domains like object learning, adjusting their interactions to facilitate visual learning (Bambach et al., 2016). As toddlers gain greater motor independence, their ability to seek out and control their visual input increases significantly, which changes the nature of the visual input (Kubicek, Jovanovic, and Schwarzer, 2017; Franchak, 2019; Kretch, Franchak, and Adolph, 2014). Given that toddlers are more able to directly and indirectly shape

their own input, individual differences in internal and motivational factors may be especially formative of visual input as infants become toddlers.

Face and Body Processing Development and Variation in Ability

There are many elements of face and body processing that develop from birth and continue to emerge into toddlerhood over varying timescales. To fully understand what the significance of individual differences in the environmental input of faces and bodies may be, it is first necessary to have some understanding of the skills that infants and toddlers must acquire over the first months and years of life. While development of the ability to recognise, understand and extract information from faces and bodies shows many commonalities in individuals who are typically developing (Bjornsdottir et al., 2017; Furl et al., 2011), there are differences in processing ability and qualitative nature even outside of clinical populations— and, of course, there are many disordered developmental trajectories that include some form of face or body processing deficit or difference.

Human Face Processing.

Faces are a particularly rich source of information on other people for humans. Humans (and chimpanzees, to some extent) are especially sensitive to subtle differences in facial feature spacing, allowing for discrimination between individuals based on facial features with greater ease (Parr, 2011). Beyond this, the movements or poses of these facial features can reflect internal states, motivations and intentions. While there is evidence for the use of facial movements to reflect emotional or social states across mammalian species, from rodents to

primates (Zych and Gogolla, 2021), humans have faces especially adapted for communication (Kret et al., 2020). Over the course of infancy, toddlerhood and early childhood, humans must develop the ability to extract and understand this very rich information from the face.

The face processing abilities emerge rapidly over developmental time. From birth, infants show a preference for face-like stimuli, possibly driven by basic perceptual features (Mondloch et al., 1999; Simion et al., 1998; Simion and Giorgio, 2015), but abilities beyond this are limited.

Although neonates can distinguish human from monkey faces, they must rely on external cues to recognize familiar faces like their mothers (Di Giorgio et al., 2012; Pascalis et al., 1995).

Morton and Johnson's (1991) two-process theory would suggest that infants initially rely on a preference for face-like configurations (CONSPEX) before beginning to extract more complex visual features from around 2 months as a result of environmental input (CONLERN). By 3-months, infants show nascent differences in face processing dependent on environmental input (Bar-Haim et al., 2006; Hayden et al., 2007; Kelly et al., 2005; Quinn et al., 2002) which become more established as infants approach 12-months (Pascalis et al., 2005; Grossman et al., 2007; Xiao et al., 2013; Anzures et al., 2013). Categorical perception of faces is also evident around this point (Ramsey et al., 2005). The specialisation for frequently observed categories of faces and the role for caregiver input would suggest that the socioemotional input may be key for early face processing development.

While on average, adults are familiar with approximately 5,000 faces, the degree of expertise in processing these faces differs widely (Jenkins, Dowsett, and Burton, 2018). Typical

adults process faces holistically, evident in phenomena such as the 'Thatcher effect,' (Weldon et al., 2013), but there does exist variability in face processing ability (Lander, Bruce and Bindemann, 2018; Tardif et al., 2019; Wilmer et al., 2010; Palermo et al., 2013).

The variation in adult face processing ranges from those with prosopagnosia to "super-recognizers," who have exceptional or highly exceptional face recognition abilities (Ramon et al., 2019; Russell et al., 2009). Super-recognisers markedly outperform the average person in facial identity recognition, including for other-race faces (Robertson et al., 2016, 2020; Russell et al., 2009; Noyes et al., 2017). There is limited information on the developmental timeline over which these abilities emerge, but they appear to be established by adolescence (Bennetts et al., 2017). While there is not an equivalently widely researched category of people demonstrating superior performance for facial emotional expression recognition, face identity recognition and facial emotional expression recognition ability are associated, even at extreme ends of the ability spectrum (Ali and Davis, 2020).

Prosopagnosia is a condition wherein individuals struggle specifically with face recognition due to either brain injury or developmental differences, (Damasio, 1985; Riddoch et al., 2008). This modular deficit provides evidence for a specific and dissociable set of face processing cognitive processes. Developmental prosopagnosia (DP) affects about 2% of the population (Susilo and Duchaine, 2013). While case studies with details of how DP presents in infants and young children are limited, the information on this would suggest clear evidence of divergence from the typical developmental pathways, with infants that go on to be diagnosed

showing early signs of an inability to discriminate familiar from unfamiliar individuals (Dalrymple et al., 2012). While not universal, facial emotional expression processing impairment is not uncommon in those with this diagnosis (Stephan et al., 2006; Dobel et al., 2007; Biotti and Cook, 2016).

Between these extremes, typical adults also show varying levels of skill. Some individuals experience mild to moderate deficits associated with other disorders. Deficits in face processing have been implicated in a range clinical disorders such as Autism Spectrum Disorder (ASD) (Sasson, 2006), Conduct Disorder (CD) (Fairchild et al., 2014), Schizophrenia (Marwick and Hall, 2008) and Antisocial Personality Disorder (ASPD) (Dawel et al., 2012). Understanding variation in the visual input of faces and bodies could shed light on the aetiology of this broad range of face processing abilities.

Critically, face processing proficiency is greatly influenced by experience and familiarity with different categories of faces. Adults are generally better at recognizing familiar categories of faces, such as those from their own race or cultural group (the other-race effect, ORE). This effect is experience-dependent and can be attenuated with increased exposure to other-race faces, especially when that experience occurs in early life (Chiroro and Valentine, 1995; Walker and Hewstone, 2006). This would suggest that specific aspects of variation in face experience is critical.

The influence of this experience is further evident in the research on face “prototypes”. Evidence for these prototypes can be seen in the more efficient recognition of distinctive faces, which deviate from these prototypes (Mauro and Kubovy, 1992). Further, these prototypes are more well-developed for familiar categories of faces (Short et al., 2011). These prototypes indicate that face processing skills are highly dependent on the individual's environmental input and prior exposure to faces.

So, adult face processing is typically proficient and experience dependent. However, there is great variation in mature face processing ability, and typical adults can be selectively impaired in processing categories of face with which they are unfamiliar. Direct data on factors shaping the environmental input may clarify the aetiologies of diverse trajectories in face processing development.

Toddlerhood as A Crucial Period for Face Processing Development.

As infants become toddlers, there is evidence that they become capable of more complicated inferences about faces and more adult-like processing. One key event in toddlerhood is the emergence of more sophisticated language abilities, which mean that testing of explicit understanding is possible (Horvath et al., 2022). This allows toddlers to engage with more varied experimental paradigms beyond simple looking behaviour.

By 24 months, toddlers can integrate gaze cues with body language to learn from the environment (Moore and Povinelli, 2007; Paulus and Fikkert, 2014). By 3.5 years, perception of emotional expressions is more mature if not fully adult-like (Cheal and Rutherford, 2011). While these new skills develop, the face-processing system retains its plasticity, allowing new environmental experiences like exposure to infant faces to moderate face processing development (Cassia et al., 2009). Further, sensitivity to the mother's face diminishes between 24 and 45 months, with toddlers showing stronger ERP responses to stranger faces (Carver et al., 2003). This could indicate a growing role for new and unfamiliar categories of face in the toddler visual input.

Significant individual differences in face processing emerge during toddlerhood. Typically developing toddlers exhibit an attentional bias for faces, but at 32-months a notable attenuation of this bias is noticeable in Autistic toddlers (Katarzyna et al., 2010). Variation in scores on theory of mind (ToM) tasks also predict variations in how toddlers visually scan facial expressions (Poulin-Dubois et al., 2018) -- this is significant given the relationship between ToM and Autism. This demonstrates how diverging developmental paths in face processing become especially apparent in toddlerhood.

Face and facial emotional expression processing development does not end with toddlerhood. Development of the ability to process faces and facial emotional expressions continues into childhood and adolescence (Batty and Taylor, 2006; Chung and Thomson, 1995; Thomas et al., 2007). For instance, face identity recognition accuracy and speed continue to

improve with age up to at least age 10 (Diamond and Carey, 1977), and specialisation of ERP components continues into adolescence (Taylor et al., 2004). However, many of the changes that take place in childhood and adolescence involve increasing skill in existing skills or changes in the qualitative nature of existing processes (for example, changes in cortical activity and functional connectivity in response to facial emotional expressions; Wu et al., 2016).

Overall, it is clear that face processing is still emergent in toddlerhood, with this being a critical period for face learning just as infancy is. This period also reveals the emergence of significant individual differences in face processing. Which factors are associated with and may drive the emergence of these individual differences remains unclear.

Human Body Processing.

While there is less research on the visual processing of bodies than faces, humans are also capable of extracting a wide range of information from bodies, ranging from identity to emotional state (Chouchourelou et al., 2006; Johansson, 1973; Rice et al., 2013). In the same way that faces are processed holistically, bodies are processed as more than the sum of their parts, with whole-body information contributing to analyses of bodily posture and bodily “expressions” (Harris et al., 2016; Willems et al., 2014). Not only this, but as with faces, bodies are processed contextually and alongside other multisensory cues to infer a range of information about others within the environment (Jessen and Kotz, 2011).

Simple elements of body processing emerge very early in development, though many skills develop later. This to some extent reflects the face processing timeline. However, often, comparable body processing abilities are evident later than their face-processing counterparts. From birth, infants show a behavioural preference for attending to biological motion (Fox and McDaniel, 1982; Bidel-Ildei et al., 2014). This attentional bias may then guide the visual input that provides the substrate for more complex development.

In early infancy, some holistic processing of bodies emerges. By 3-months, infants can distinguish between different types of bodily motion when presented as point-light displays; while both 3-month-olds and 5-month-olds exhibit this ability, only 5-month-olds are more sensitive to between-limb than within-limb relationships., which is a marker of holistic processing (Booth et al., 2002). At 5-months, infants distinguish whole-body postures from scrambled body postures, which again provides evidence for emerging holistic processing of bodies (Hock et al., 2016). They also begin to exhibit signs of recognising when vocal and bodily emotional expressions match and demonstrate processing of bodily actions as goal directed (Heck et al., 2018; Woodward, 1998), which indicates that intentional and state-based information is being extracted from these bodies.

While for faces the inversion effect, a key indicator of holistic bodies, emerges early in development, this does not emerge for bodies until around 14-months (Gillmeister et al., 2019; Turati et al., 2004). The ability to discriminate between intact and scrambled bodies is somewhat robust by 18-months, and at the same point, infants demonstrate learning from

adult body movements about tool use (Slaughter et al., 2002; Rat-Fischer et al., 2012). This indicates that infants at this point can now use information from bodies to learn about their environments. So, infants progress from simple attentional biases for bodies and biological motion to a complex ability to extract a variety of information about bodies.

While there is not as much research on the variability in the developmental trajectories of body processing development, there is some research to suggest that there are significant individual differences in body processing skills. There is no equivalent concept of “super recognizer” for bodies, so much of the evidence for individual differences comes in the form of evidence for deficits in body processing in certain groups. However, these are often associated with more domain-general deficits than the face-specific deficits reported above.

In DP, there has been some research to indicate a co-occurring deficit in body identity recognition, potentially as a result of general deficits in holistic stimulus processing (Biotti et al., 2017); however, these effects are subtle, with less sensitive tasks indicating simple differences in response time rather than accuracy (Rivolta et al., 2017), which would be consistent with the fact that individual patients do not report particular problems with body perception, and in fact often use bodily features to offset their face processing deficits (McConachie, 1976).

A similar, subtle deficit in body identity processing has been identified in those with schizophrenia, with slower processing of bodies on matching tasks and a slightly reduced inversion effect as compared to controls (Bauser et al., 2012); however, given that this

generalises to object classes, this likely reflects a domain general deficit in configural processing rather than a specific deficit in identity processing for bodies.

In terms of bodily emotional expressions, there are noted deficits in extraction of this information from bodies in socially related disorders where body processing is not necessarily central. Individual differences in body processing have been implicated in ASD (van Boxten et al., 2017), with deficits in extracting emotional information from body posture (Philip et al., 2010) and with the absence of the body inversion effect for posture observed in typical controls (Reed et al., 2007).

Overall, there is some variability in the ability to recognize and extract information (for example, affective information) from bodies, but the evidence here is mostly limited to the deficits present in disorders. It is unclear whether an equivalent concept to the face “super recognizer” exists for bodies, or even whether the ability to process bodies varies as widely as the ability to process faces.

Toddlerhood as A Crucial Period for Body Processing Development.

In toddlerhood this development continues, though again research on body processing development is more limited than equivalent research for face processing development. By 20-months, toddlers exhibit mimicry of bodily action not possible for younger infants (Jones, 2007), though it is very possible that this is in part driven by improving motor ability. There are many developments not specific to body processing that may indirectly allow for the extraction of

more sophisticated body processing— for example, more established awareness of the self-other distinction (Lewis and Brooks-Gunn, 1979).

While research on processing bodily emotional expressions or interpretation of bodily postures in infants and toddlers is also less extensive than on facial expressions, the existing research suggests that these skills mature relatively early with limited age-related variation (Vuong and Geangu, 2023). This contrasts with the more gradual development of facial emotional expression processing.

Overall, body processing is also very limited at birth and skills in this domain emerge over infancy and toddlerhood, from simply perceiving bodies holistically to interpreting the intentionality of actions. However, perhaps aside from bodily emotion processing, often equivalent elements of body processing (for example, holistic processing) develop later than their counterparts for face processing. This could indicate that, if the environmental input were an important driver of body processing, maybe input later in developmental time would be more key for these developments.

Individual Differences that May Shape the Environmental Input

So, there is variation in socioemotional stimulus processing development, and there is some evidence that variation in early visual input shapes later processing ability, but we do not know what factors may shape the frequency of face and body exposure for toddlers. There is

limited direct evidence for which variables may be important in shaping the visual input.

However, there is indirect evidence for which internal and external variables may be significant.

It is important theoretically to understand factors shaping the social environment that are external to the child. Even identical children placed in different countries, cultures and homes may be exposed to socioemotional stimuli at vastly different rates because of differences in those environments. Even a factor as simple as the number of individuals around a child could shape their visual input in terms of faces and bodies. For example, there is a relationship between early childcare attendance and social competence (Volling and Feagans, 1995) -- this could be due to these toddlers being exposed to a greater number and variety of peer faces early in life.

As discussed, there is clear evidence that the environment shapes face processing in some way to be seen in the other-race effect (ORE) (Chiroro and Valentine, 1995). The ORE emerges as early as 3-months (Bar-Haim et al., 2006; Kelly et al., 2005), suggesting that even very early visual experience is highly formative of development. Similarly, there is perceptual narrowing for face species (Scott and Fava, 2013). There is direct evidence that these perceptual narrowing processes are causally related with the statistics of the visual input and the socioemotional environment (Pascalis, Scott and Kelly, 2005), meaning that factors external to the child shape face processing proficiencies. While the ethnicity and species of the faces in the environment is a rather drastic external factor, many fine-grain aspects of the external environment are likely to shape the degree of visual exposure to faces and bodies.

An important external factor is the presence of child siblings. Here, we will focus on the external variable of the presence of child siblings for two reasons. Firstly, the presence of siblings is associated positively with childhood and adulthood socioemotional skill level, such as mentalizing ability (Lo and Marr, 2022; McAlister and Peterson, 2006). Secondly, child siblings are uniquely placed to potentially affect face frequency in the visual field of a child, given that they often act as a playmate, and given that unlike adults, they are typically similar in height and so these faces are more easily accessible (Frank et al., 2013; Kretch, Franchak and Adolph, 2014). For this reason, the pathway between the presence of siblings and later enhanced socioemotional ability could plausibly be related to a greater incidence of faces in the egocentric view of children with siblings.

Adult faces tend to be out of the easily accessible field of view for walking toddlers, who must crane their necks to view them (Kretch, Franchak, and Adolph, 2014). This may mean that siblings of a similar age (and typically a similar height) increase the accessibility and thus the frequency of faces within the toddler view. It could also be that simply having more individuals present within the household increases the frequency with which faces and bodies are easily viewable and accessible. These environmental differences may then have an impact on socioemotional stimulus processing development.

Indeed, there is direct evidence that presence of a younger sibling in toddlerhood translates to later enhanced ability to process infant faces (Cassia et al., 2013); this impact of siblings on

relative processing of different face age groups has been reported for younger siblings also (Conte et al., 2019). While there is limited evidence on the relationship between sibling status and visual body processing, it is not implausible that there would be a similar relationship between sibling status and body processing.

It is also important, in understanding factors that may shape the environmental input, to discuss factors that are specific and internal to the individual child, and that are sometimes more closely related to innate factors. It could be that two individual children in the exact same environment would still be exposed to different face and body frequency as a result of the way in which their navigation of the environment shapes their input. This may take the form of shaping the input through looking behaviour and attention, or it may take the form of shaping the input through eliciting responses in other individuals. This shaping of the social environment could be acute, as infants and toddlers directly elicit responses from others, or it could be longitudinal, as internal variables shape relationships within the home and social environments.

Some of the most pertinent examples of child internal individual variables affecting the socioemotional input the toddler is exposed to can be taken from work examining how locomotor ability shapes the stimuli observed: compared to crawling toddlers, toddlers who walk have more readily available visual input from caregivers and other individuals, particularly when it comes to faces (Kretch, Franchak and Adolph, 20134). This comes about because of the relative pragmatic difficulties of accessing faces, which are often in the upper half of the visual field, from a crawling position, wherein the head is oriented down to the floor and looking

upwards requires an elaborate movement of the body from crawling to sitting. When infants are just learning to navigate the environment, they may also need to spend more time looking at their own bodies in order to coordinate difficult movements. Comparatively, faces are much more often in the view of confidently walking toddlers, along with all other features present in the upper half of the visual field, as their heads are facing forwards and they can easily look up with a head movement alone, as opposed to the complex motor actions a crawling infant may have to engage in just to access one face. This shows that as children become more independent, they actively and systematically shape their own environmental input, influenced by internal factors such as interests, knowledge, and disposition.

At a less basic level, differences in the attentive behaviours of infants and toddlers at the same developmental stage may be shaped by internal variables, and the input may be shaped in this way. Infants and toddlers selectively attend to and take in information with head and eye movements, shaping the visual input that they are exposed to. Early on, infants will respond differently to familiar or novel stimuli, showing a familiarity or novelty preference (depending on age, the degree of familiarity, etc.) (Houston-Price and Nakai, 2004). From early infancy, visual attention is moderated by temperament (McConnell and Bryson, 2005). This demonstrates that from the time that infants are capable of selectively allocating visual attention to stimuli, their own motivations, knowledge and interests (based on experience and temperament) shape the visual input that they receive.

There is also evidence that infants systematically shape their own perceptual input through wider bodily movement: for example, with increased manual motor abilities, infants can manipulate the objects that they are viewing to generate a wide range of views; they do this in a fashion which appears to optimise visual object learning (Bambach et al., 2016). This is, like gaze and head movement, shaped by prior interests, knowledge and disposition, and this kind of direct shaping of the environmental input may be impacted by motor ability. This demonstrates how as infants and toddlers become more independent, they take a more active role in shaping their perceptual input through moving and using their whole bodies to explore. The visual input is shaped in at least some part according to a variety of internal variables. So, toddlers are not passive receivers of surrounding information and may approach and interact with faces or other stimuli according to their interests.

Finally, interpersonal behaviour may shape the perceptual input, with internal variables such as needs, interests, abilities and disposition may shape the perceptual input through shaping the responses of the adults around the child. There is a great degree of literature on the bidirectional relationship between child and parent behaviours in shaping emotional and social development (e.g., Wang et al., 2021). So, two toddlers with very different temperaments may elicit very different responses from the same caregiver. These differences in response may be very acute and mediated by differences in child behaviour, or they may be more a function of the relationship that has been established over time. Caregiver responses may then increase or decrease the frequency of faces and bodies in the child view.

One specific factor that has been proposed to be related to how infants shape their own environments, and hence the type of visual audio-visual input, is temperament (Martinis et al., 2012). Temperament refers to individual differences in the ways in which infants react to the world and people around them, typically measured by parent-report and broken down into emotional reactivity, self-regulation and surgency (Garstein and Rothbart, 2003; Goldsmith, 1982; Rothbart and Bates, 2007). Reactivity is the tendency and nature of an individuals' responses to the environment, while regulation encompasses the control processes that up or down regulate these responses; surgency is akin to extraversion (Putnam et al., 2001; Rothbart, 1989). The most frequently used instruments for measuring temperament (with its associated factor structures) are the IBQ-R in infancy and the ECBQ in childhood (Enlow et al., 2016; Gartstein and Rothbart, 2003; Putnam et al., 2001, 2006, 2008).

Temperament is linked to later social skill level, particularly skill level in processing social cues. Among the fundamental factors of temperament, emotional reactivity and emotional regulation are particularly important for understanding early temperament and its impact on socioemotional development (Rothbart, 1986; Rothbart and Bates, 2007). For instance, less reactive infants tend to be less empathetic by age two (Young, Fox, and Zahn-Waxler, 1999).

Infant temperament has been shown to be relevant for predicting a wide variety of later social skills, including those related to processing social cues displayed by human faces and bodies. Particularly, emotional reactivity and emotional regulation seem to be influential in predicting socioemotional development, or otherwise related to socioemotional outcomes

(Rothbart, 1986; Rothbart and Bates, 2007). For example, less reactive 4-month-olds develop less empathic behaviour towards strangers at the age of 2-years-old (Young, Fox and Zahn-Waxler, 1999). Aspects of infant temperament are acutely associated with emotional expression processing— for example, approach and fearfulness are associated with physiological differences in processing of fearful body expressions (Rajhans et al., 2015), and the factors of both negativity and regulation are associated with differences in infant processing happy and fearful faces respectively (Martinis et al., 2012). This could mean that temperament shifts infant response to socioemotional stimuli in the short-term, which could then have a broader impact on socioemotional development and learning in the long-term.

Importantly, temperament has been associated with risk of developing socially based disorders such as ASD (Pijl et al., 2019). As previously outlined, these socially based disorders often entail deficits in processing faces, bodies and emotional expressions thereof. Given the significance of the socioemotional stimulus processing in these disorders, it could be that temperament shapes aspects of how infants and toddlers attend to and receive information from faces and bodies in the environment. For example, it could be that toddlers who tend to experience high arousal and show poorer self-regulation abilities, tend to experience high arousal in response to faces which is less likely to diminish as exposure to faces continues. This could elicit an avoidance response, such as looking away— infants are known to look away from distressing or overwhelming stimuli as a form of self-regulation (Ekas et al., 2013) -- shaping the input. There is some evidence that temperament is related to the ways in which infants explore

the environment using motor action to explore objects (Rothbart, 1988), but it is unclear how this would translate into visual exploration of the environment.

Indeed, there is some evidence that temperament may be associated with specific aspects of looking behaviour towards faces and emotional expressions. For example, the level of negativity has been shown to moderate looking behaviour with threat-related emotional expressions (Pérez-Edgar et al., 2017). Further, Rennels, Kayl and Kulhanek (2020) found that elements of regulation (orienting) and surgency, in combination with face characteristics, predicted aspects of looking behaviour in 6- to 10-month-olds. There was an interaction with infant gender in this relationship. These acute relationships between temperament and looking behaviour could translate into wider differences in the visual inputs of infants (and later toddlers) with different temperaments. There is more limited evidence for the relationship (if any) between temperament and body processing.

There is also some more indirect evidence that toddlers and children with different temperaments could shape their socioemotional environments and thereby indirectly shape the visual input that they receive. For example, toddler negative affect is associated positively with parent stress, while toddler effortful control is negatively associated with parent stress (Adreadakis et al., 2020). So, given that parents of toddlers with different temperaments experience different levels of stress, these toddlers may experience different durations and styles of interactions from these parents. The parents may interact more or less; they may take the toddler to group activities more or less frequently; they may initiate more or fewer face-to-

face interactions. Different parent ratings of toddler temperament have been directly associated with differences in toddler-directed behaviours— higher ratings of activity level are associated with higher maternal anger, while toddler fearfulness is associated with maternal worry (Bryan and Dix, 2009). This may result in a difference in the amount of time the toddler spends with adults— and their bodies and faces— in view.

So, there is indirect evidence that temperament may shape the socioemotional visual input, as temperament shapes aspects of looking behaviour and elements of the social environment. This could then explain longitudinal relationships that exist between temperament and socioemotional development. However, there has thus far been no direct examination of the relationship between temperament and the socioemotional visual input of faces and bodies. Understanding this would be essential for building models explaining the relationship between temperament and social development.

Aims and Hypotheses

The aim of this paper is to characterise how the naturalistic frequency of faces and bodies in the egocentric view of toddlers may be shaped by individual differences in both temperament and the presence of siblings. This is one essential piece of evidence for building models of how individual differences in face and body processing emerge.

We focus on toddlers because in toddlerhood, many developmental changes are still taking place and because older, mobile toddlers may more easily be able to shape their own input as their motor skills are better-established (Franchak, 2019). Walking toddlers view the environment from a unique perspective, with faces in the upper half of the visual field being accessible with only concerted motor effort (Karasik et al., 2011; Kretch et al., 2014). This analysis will allow us to understand how differences in later socioemotional processing skills could be explained by differences in input, and how these differences in input might be themselves shaped. We will focus on longitudinal temperament as measured by the IBQR (Garstein and Rothbart, 2003) and a corresponding concurrent measure of temperament on the ECBQ (Putnam et al., 2006). The former may be relevant as temperament longitudinally shapes the caregiver-child relationship, while the latter may shape the input through more directly determining the nature of interactions. Concurrent temperament may shape the frequency with which faces and bodies are seen through also directly shaping toddler interests and looking behaviour.

We will also explore the relationship between whether the toddlers have a child sibling and the presence of faces and bodies in the toddler view. We expect that the presence of child siblings in the family home will shape the toddler's visual input, largely as a result of pragmatically making faces and bodies more accessible within the toddler view and more frequent in the toddler environment. We expect this to be especially pronounced for faces, as adult faces are typically more difficult for toddlers to access. This may potentially explain

relationships that have been observed between siblings and specific aspects of social development.

We hypothesise that: 1) faces will make up a higher percentage of the visual input of children with siblings than children without siblings; 2) bodies will make up a higher percentage of the visual input of children with siblings than children without siblings; 3) aspects of longitudinal temperament will be associated with the frequency of bodies in the toddler view; 4) aspects of longitudinal temperament will be associated with the frequency of faces in the toddler view; 5) aspects of concurrent temperament will be associated with the frequency of bodies in the toddler view; 6) aspects of concurrent temperament will be associated with the frequency of faces in the toddler view

Methods

See general methods section for full details.

Data Analysis Strategy

To maximise the sample size and statistical power here, we first consider the full $N = 26$ for the role of siblings, and then separately the $N = 23$ participants with IBQR data provided at 13-months and then the $N = 24$ participants with ECBQ data provided concurrent to recording.

Results

We observed great variability in the frequency of faces and bodies. For participants with IBQR data, the mean percentage of frames containing bodies was 53.10% (SD = 6.73), with a minimum of 36.58% and a maximum of 64.61%; the mean percentage of frames containing faces was 7.27% (SD = 4.47) with a minimum of 1.86% and a maximum of 17.60%. For participants with ECBQ data, the mean percentage of frames containing bodies was 53.83% (SD = 6.78), with a minimum of 36.58% and a maximum of 64.61%; the mean percentage of frames containing faces was 7.68% (SD = 4.48), with a minimum of 1.86% and a maximum of 15.96%. The wide standard deviations here reflect how variable the percentage of frames containing faces and bodies was.

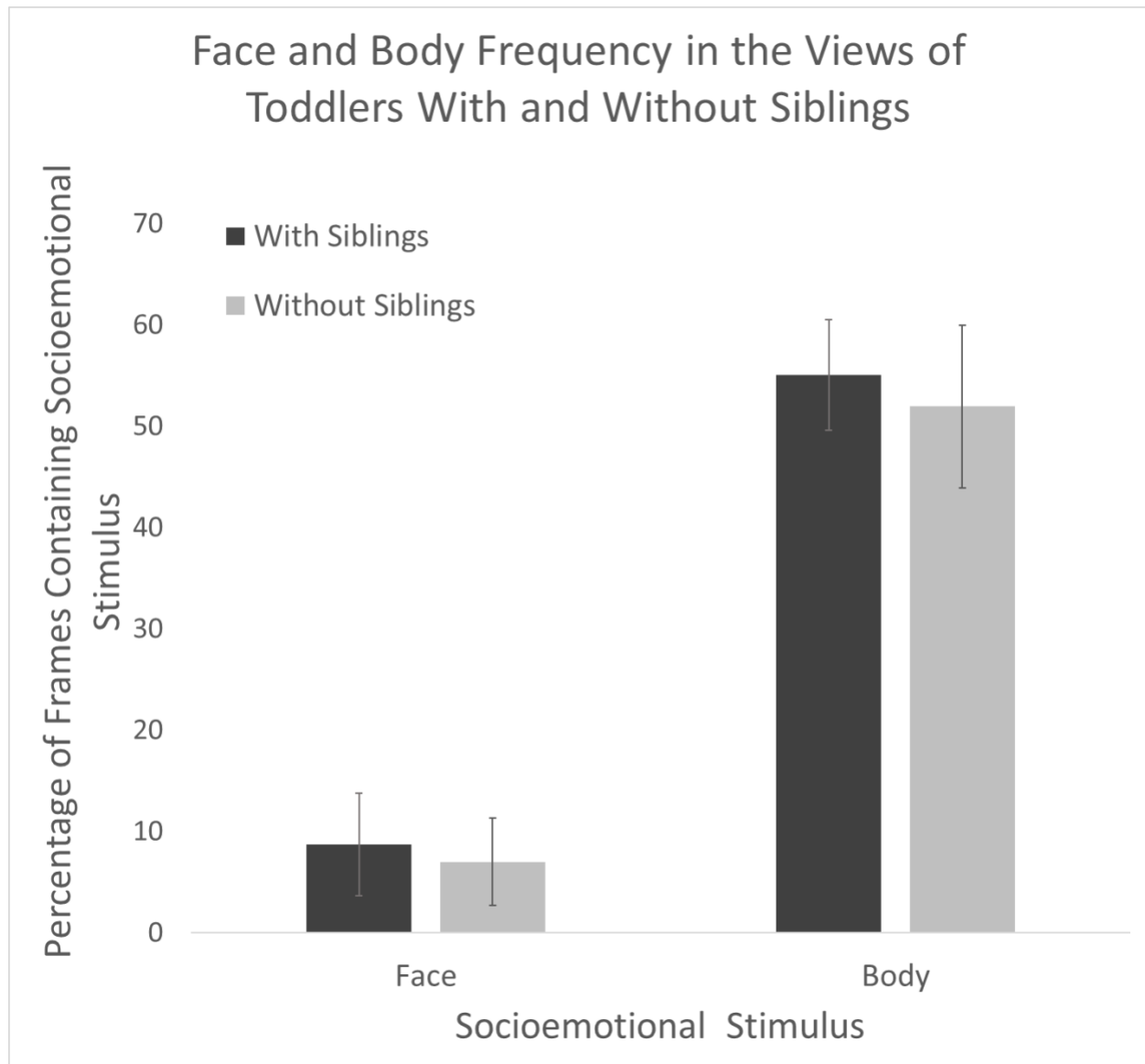
Presence of Siblings

Of the 26 participants who provided valid video data, 11 had no child siblings and 15 had at least one child sibling. The descriptive statistics suggest that the number of frames containing faces and bodies in the sibling and no sibling groups were similar (figure 1).

However, there was an uneven gender split between these groups that presented a potential confound: while 9 out of the 15 children with siblings were male, only one participant in the no-siblings group was male. To investigate how likely gender was to present a confound, gender was considered in the sibling analysis. The below chart (fig. 2) shows the mean and standard deviation face and body frequency values for male and female participants.

Figure 1

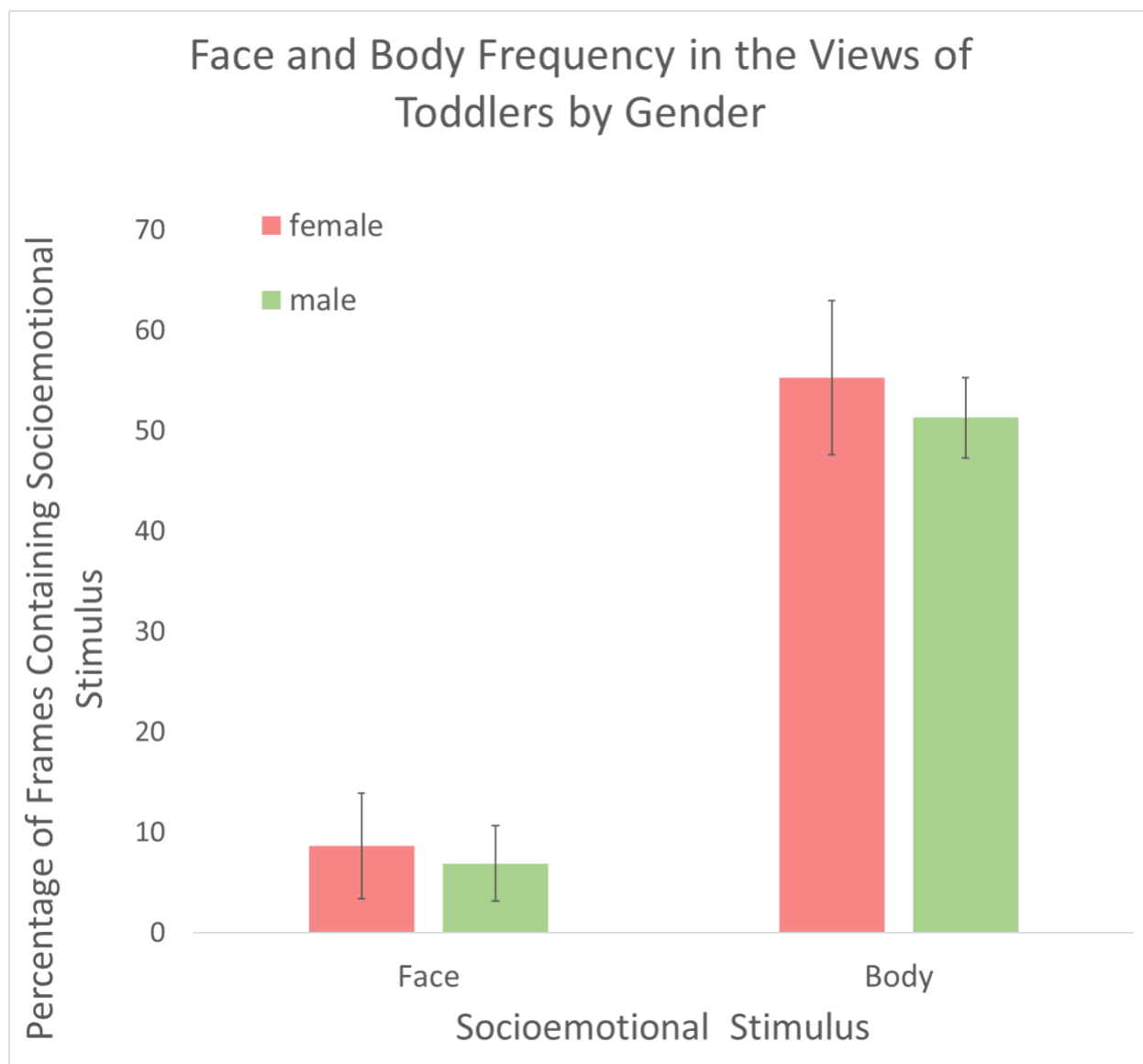
The percentage of frames containing faces and bodies by sibling status.



Note: This chart demonstrates the average and standard deviation frequency of each social stimulus by sibling status. On average, in the sibling group, 55.06% (SD = 5.48) of frames contained bodies, and 8.70% (SD = 5.07) contained faces. On average, in the no sibling group, 51.92% (SD = 8.00) of frames contained bodies, and 7.00% (SD = 4.31%) of frames contained faces.

Figure 2

The percentage of frames containing faces and bodies by gender.



Note: This chart demonstrates the frequency of each social stimulus by sibling status. On average, for female participants, 55.26% (SD = 7.67) of frames contained bodies, and 8.65% (SD = 5.28) contained faces. On average, for male participants, 51.28% (SD = 3.98) of frames contained bodies, and 6.90% (SD = 3.75) of frames contained faces.

A multivariate analysis of variance (MANOVA) was conducted to examine the effect of sibling status and gender on face and body frequency. Pillai's Trace was used to assess the multivariate effect due to unequal group sizes. The results indicated that there was no significant multivariate effect of sibling status, $V = 0.146$, $F(2,21) = 1.801$, $p=0.190$, $\eta^2 = 0.146$; or gender, $V=0.173$, $F(2,21)=2.193$, $p=0.137$, $\eta^2 = 0.173$; or their interaction, $V=0.088$, $F(2,21)=1.014$, $p=0.380$, $\eta^2 = 0.088$.

Longitudinal Relations Between Input and Temperament

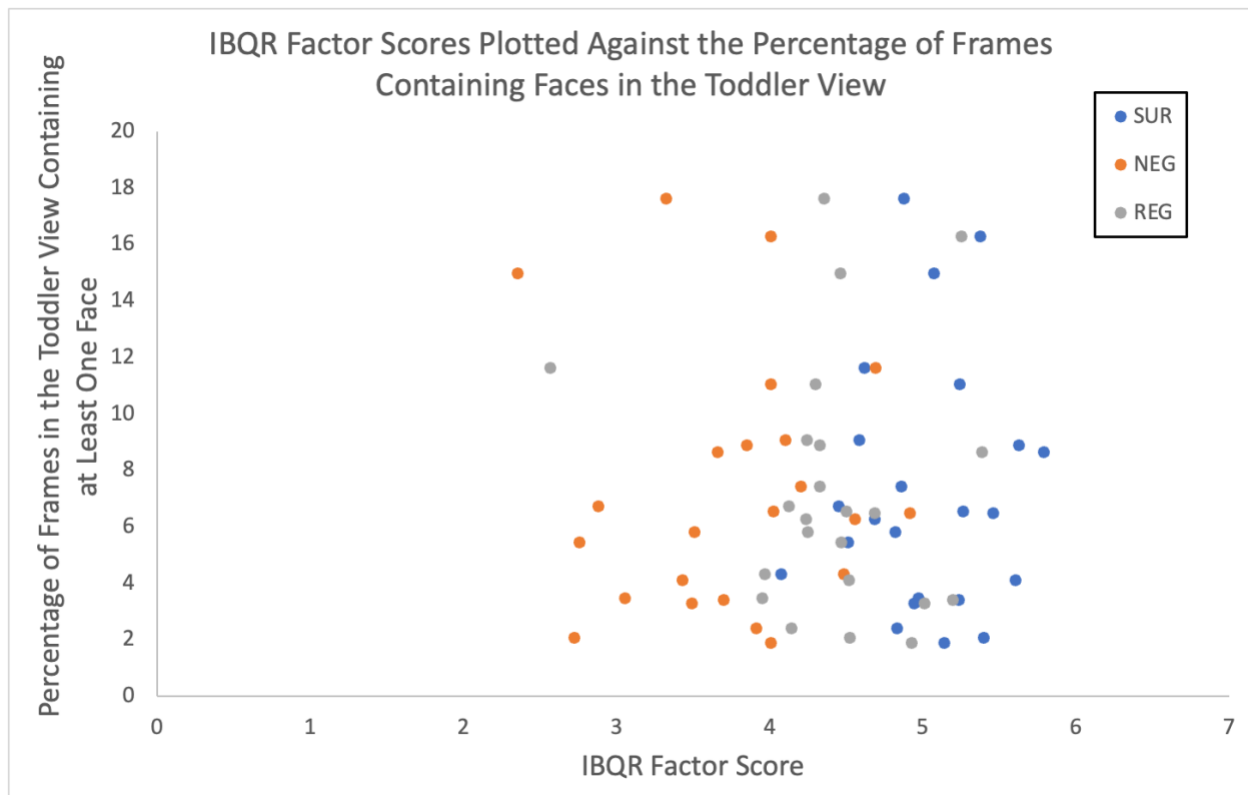
Of the 26 participants, 23 had complete IBQ-R data for analysis (14 female, 9 male).

Associations between aspects of this personality scale and the presence of faces and bodies in the toddler view was examined.

Firstly, relationships between aspects of the visual input and the three factors proposed by Garstein and Rothbart (2003) were examined-- Surgency/Extraversion, Negative Affectivity, and Orienting/Regulation. Scatter charts were created of the scores on these factors and face (fig. 3) and body (fig. 4) frequency within the toddler view. A multivariate multiple regression analysis was conducted to examine the effect of temperament factors (SUR, NEG, REG) on face and body frequency. The multivariate test results indicated that temperament did not significantly predict the combined dependent variables: surgency ($V = 0.167$, Pillai's trace, $F(2, 19) = 1.908$, $p = .176$, $\eta^2 = 0.167$); negative affectivity ($V = 0.051$, $F(2, 19) = 0.514$, $p=.606$, $\eta^2 = 0.051$); regulation ($V = 0.039$, $F(2, 19) = 0.384$, $p = .686$, $\eta^2 = 0.039$).

Figure 3

The percentage of frames containing faces as plotted against surgency, negative affectivity and regulation factor scores.



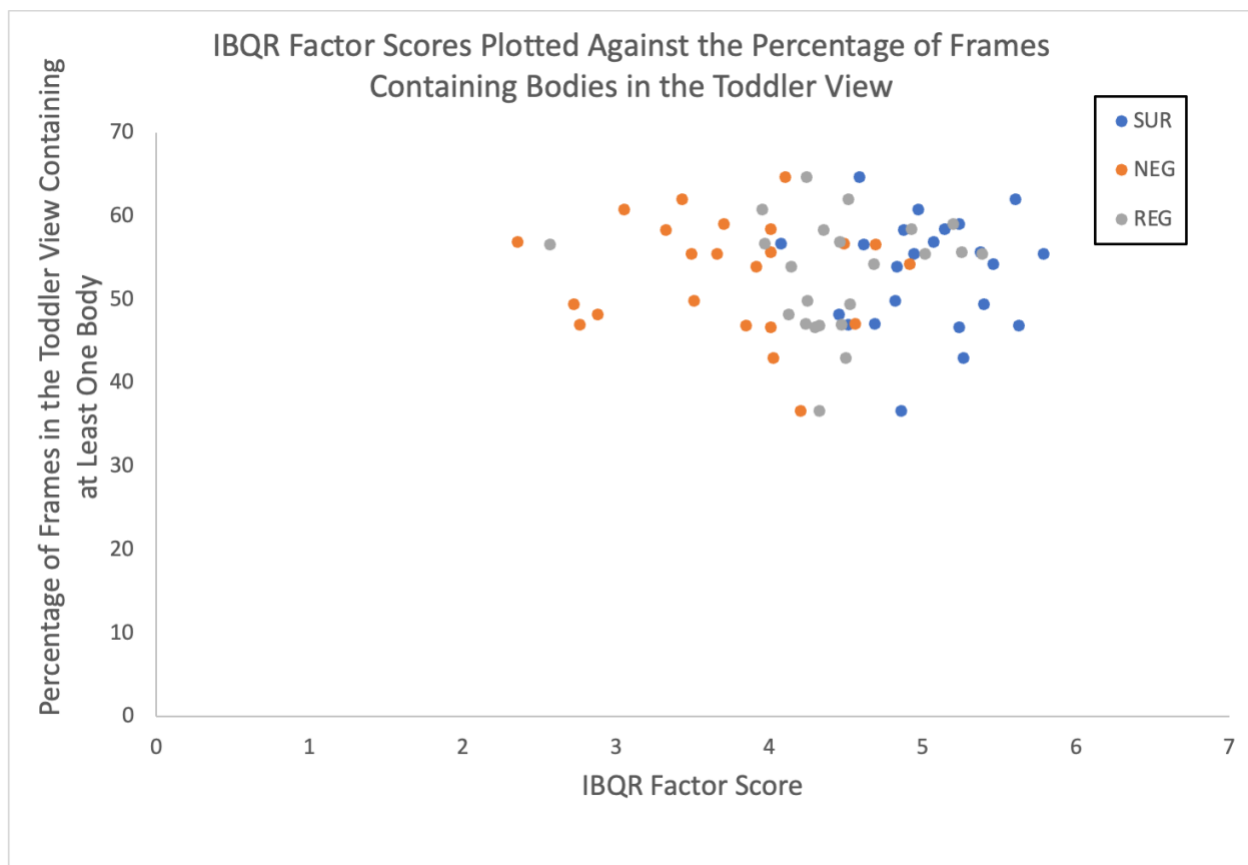
Note: This chart illustrates the relationship between scores on the three IBQR factors and the frequency of faces within the toddler view.

The univariate tests also revealed null results. While the regression model for face frequency was significant, $F(3,19) = 18.85$, $p < .001$, $\eta^2 = .739$, individual factor scores did not predict face frequency. SUR ($F(1,19) = 1.260$, $p = .275$), NEG ($F(1,19) = 0.006$, $p = .939$) and REG ($F(1,19) = 0.353$, $p = .559$) were not significant predictors of face frequency.

Similarly, the regression model for body frequency was significant overall, $F(3,19) = 325.36$, $p < .001$, $\eta^2 = .980$, but again, none of the temperament factors significantly predicted body frequency: SUR ($F(1,19) = 3.067$, $p = .095$), NEG ($F(1,19) = 1.082$, $p = .311$), REG ($F(1,19) = 0.386$, $p = .541$).

Figure 4

The percentage of frames containing bodies as plotted against surgency, negative affectivity and regulation factor scores.



Note: This chart illustrates the relationship between scores on the three IBQR factors and the frequency of faces within the toddler view.

We next examined the relationships between individual IBQR dimensions and the toddler visual input. These were exploratory, given that there was no indication of any significant or strong relationships between the three IBQR factors and face or body frequency in the toddler view.

When running a multivariate multiple regression on the individual IBQR dimensions and face/body frequency, two significant relationships emerged: between duration of orienting and body frequency ($F(1,8) = 6.895, p = .030$) and smiling and laughter and body frequency ($F(1,8) = 7.686, p = .024$). To assess the robustness of these findings, a bootstrapped analysis with 1000 samples was conducted. The bootstrapped results did not confirm the initial significant relationships. Specifically, the relationship between duration of orienting and body frequency was no longer significant (bootstrapped $p = 0.645$, 95% CI [-43.947, 29.420]), and the relationship between smiling and laughter and body frequency was also not significant (bootstrapped $p = 0.638$, 95% CI [-57.097, 36.121]).

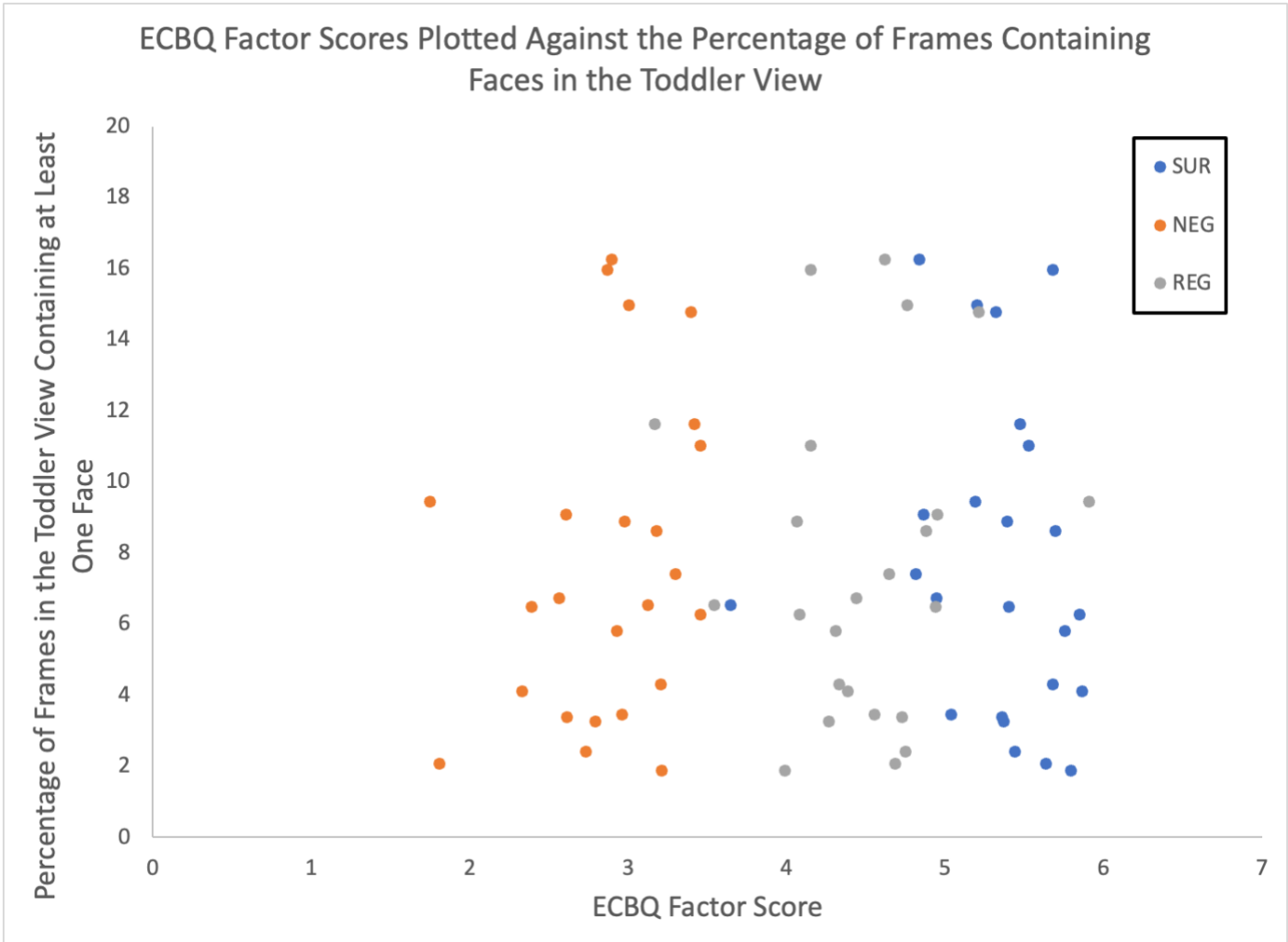
Concurrent Relations Between Input and Temperament

Of the 26 participants, 24 had complete ECBQ data gathered at around the time of recording available for analysis (15 female, 9 male). Associations between aspects of this concurrent personality scale and the presence of faces and bodies in the toddler view was examined. Firstly, relationships between the frequency of faces and bodies in the toddler view and ECBQ surgency, negativity and regulation factor scores were examined. Scatter charts were

created for the frequency of faces (fig. 5) and bodies (fig. 6) in the toddler view against these three factors.

Figure 5

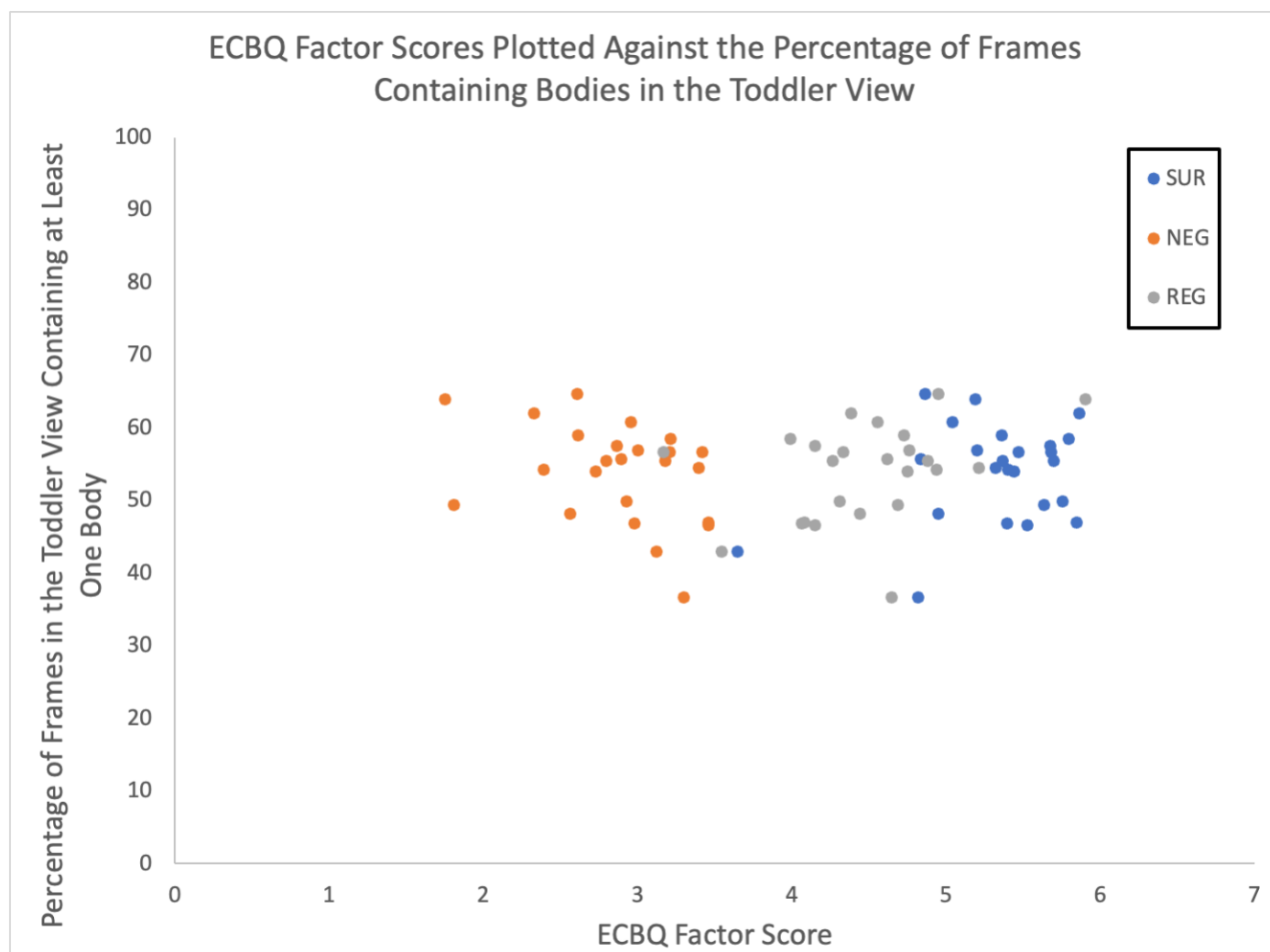
The percentage of frames containing faces as plotted against surgency, negative affectivity and regulation factor scores.



Note: This chart illustrates the relationship between scores on the three ECBQ factors and the frequency of faces within the toddler view.

Figure 6

The percentage of frames containing bodies as plotted against surgency, negative affectivity and regulation factor scores.



Note: This chart illustrates the relationship between scores on the three ECBQ factors and the frequency of bodies within the toddler view.

A multivariate multiple regression was conducted to assess the relationship between ECBQ personality factors (surgency, negative affectivity, and regulation) and the frequency of faces and bodies in the toddler view. The results indicated that the three factors were not significantly related to the combined dependent variables—surgency: Pillai's Trace $V = .134$, $F(2, 19) = 1.465$, $p = .256$, partial $\eta^2 = .134$; negative affectivity: Pillai's Trace $V = .218$, $F(2, 19) = 2.655$, $p = .096$, partial $\eta^2 = .218$; regulation: Pillai's Trace $V = .103$, $F(2, 19) = 1.093$, $p = .356$, partial $\eta^2 = .103$. Univariate analyses showed that none of the factors were significant predictors of the individual dependent variables. For face frequency, negativity had the strongest statistical relationship ($B = 4.310$, $p = .077$, see table 1), but it did not reach significance at $\alpha = .05$.

Table 1

The relationships between the three ECBQ factors and Face and Body Frequency in the Toddler view.

Relationship	B	p
Surgency / Face	-1.549	.423
Negativity / Face	4.310	.077
Regulation / Face	2.686	.187
Surgency / Body	3.814	.179
Negativity / Body	-3.742	.277
Regulation / Body	2.512	.388

Note: This table provides a summary of the results of the univariate analyses testing the relationship between surgency, negativity and regulation factors on the ECBQ and face and body frequency in the toddler view.

Given that the strongest relationship between the three factors of the ECBQ and the visual input was between negativity and face frequency, further analysis was conducted on the specific subscales within this factor. Eight dimensions load on this factor– sadness, frustration, fear, soothability (negatively loaded), discomfort, motor activation, perceptual sensitivity and shyness.

Table 2

The relationships between individual dimensions loading on the negativity factor and face frequency in the toddler view

Dimension	B	t	p
Fear	1.560	.558	.585
Sadness	1.025	.598	.559
Frustration	.053	.037	.971
Soothability	-.171	-.073	.943
Discomfort	-1.436	-.586	.566
Motor Activation	-1.653	-.466	.648
Perceptual Sensitivity	1.576	1.314	.209
Shyness	-.615	-.437	.668

Note: This table provides a summary of the results of a regression testing the relationship between different dimensions loading on negativity and the frequency of faces in the toddler view.

A multiple regression analysis was conducted to examine the relationship between face frequency and the eight dimensions that load onto the negativity factor. The overall regression

model was not significant, adjusted $R^2 = -.218$, $F(8, 15) = .486$, $p = .848$, indicating that the model did not significantly predict face frequency.

None of the individual dimensions were significantly related to face frequency (see Table 2).

Even before applying bootstrapping, none of the predictors reached statistical significance ($p > .05$). Bootstrapping the regression coefficients did not alter these results, confirming the lack of significant associations between the dimensions loading on negativity and face frequency.

These results suggest that none of the eight dimensions loading on negativity are significant predictors of face frequency in this sample.

Discussion

The aim of this research paper was to assess which internal and external individual differences may be related to (and therefore potentially shape) the wide individual differences in the frequency of faces and bodies in the view of individual toddlers as they freely explore their home environments. Toddlerhood is a time when significant changes in socioemotional development occur and changes in motor ability led to increased independence. As toddlers become more adept at navigating their environments, their ability to shape and respond to their visual input, both directly and indirectly, becomes increasingly complex. Prior research, including within chapter 1 of this thesis, suggests that walking toddlers infrequently see faces as compared to earlier ages and as compared to how often they see bodies, and that there is a great degree of variation in the statistics of the visual input (Fausey et al., 2016; Jayaraman et al., 2015, 2017; Sugden et al., 2014). We sought to examine what factors may be related to this

variation in the visual input, particularly considering the varied developmental trajectories for face and body processing.

Specifically, we examined the external individual differences of whether toddlers have a child sibling in the home and the internal individual differences of toddler temperament. We examined toddler temperament both concurrent to the time of recording and from earlier in infancy, at approximately 13-months. This is the first attempt to examine what factors, if any, may be related to or shape this socioemotional visual input.

We hypothesised that external variables, specifically having siblings, may be associated with both face and body frequency in the toddler view (hypotheses 1 and 2) for several reasons. One reason was pragmatic, with child siblings providing an additional source of face and body input, with their faces being more easily within the toddler view (Kretch, Franchak, and Adolph, 2014). Given links between having siblings and socioemotional development, including face processing (e.g., Lo and Marr, 2022; McAlister and Peterson, 2006; Cassia et al., 2013; Conte et al., 2019), a relationship between the presence of siblings and the frequency of faces and bodies in the toddler view could also provide a plausible causal mechanism through which siblings shape socioemotional development. Specifically, we hypothesised that faces and bodies would be in the view of toddlers with siblings more frequently than they are in the view of only children.

We also hypothesised that internal variables such as temperament may be associated with the frequency of faces and bodies in the toddler view for several reasons. Temperament has been linked to a wide variety of variables, with these variables often relating to socioemotional development. For example, prior research indicates links between temperament and concurrent looking behaviour at socioemotional stimuli (Rennels, Kayl, and Kulhanek, 2020) as well as parent-child relationships (Andreadakis et al., 2020; Bryan and Dix, 2009). So, temperament has the potential to shape both visual input both concurrently through looking behaviour and longitudinally through indirectly the social environment and relationships therein. Exploring relationships between temperament and socioemotional visual input could also shed light on the long-term relationships between temperament and social development that are evident in both typical and clinical populations (Pijl et al., 2019; Young, Fox, and Zahn-Waxler, 1999). Given that this is the first analysis of temperament and the naturalistic socioemotional visual input, we formulated broad hypotheses that aspects of both concurrent and longitudinal temperament would be related to both face and body frequency in the views of toddlers. We did not confirm any of our hypotheses.

Sibling Status and the Naturalistic Input of Faces and Bodies

Hypotheses 1 and 2, which suggested that toddlers with and without siblings would see different frequencies of faces and bodies, respectively, was not supported by our analysis. Given the confound of gender on presence of siblings in our sample, we included this in our model. There was no significant effect of sibling, gender or an interaction, suggesting that toddlers with siblings do not see faces or bodies more frequently than toddlers without siblings. So, perhaps

external, pragmatic constraints of the environments of toddlers are not so formative of individual differences in socioemotional stimulus (face, body) visual input.

It would seem strange that the presence of siblings in the toddler environment is unrelated to the frequency of faces and bodies, since sibling status has previously been associated with both later general socioemotional proficiency and proficiency with faces of different categories (Cassia et al., 2013; Conte et al., 2019; Lo and Marr, 2022; McAlister and Peterson, 2006). It also seems strange given that intuitively, the faces of siblings especially would be much easier for toddlers to access than the faces of caregivers. Simply being in the presence of more individuals— particularly individuals with faces at a visually more easily accessible height— would seem likely to shape face and body frequency in the toddler view.

One potential explanation for these null findings in terms of sibling differences could be the confound of gender. All but one of these participants were female. Given that female and male children may be involved in different socialisation practices (Brody, 2000; Brody, 2013; Seavey, Katz and Zalk, 1975), the gender difference could have confounded sibling differences. While we find no differences in face frequency between the male and female toddlers, the frequency of bodies in the view of female toddlers was slightly higher than that in the view of male toddlers. Further, the direction of the gender imbalance, with more females in the no sibling group, could have masked effects of both gender and sibling. For example, if female children receive more social input from parents, this could counteract the absence of siblings to interact with. While these issues may necessitate further investigation with a wider, more

gender-balanced sample, our results provide no evidence for an influence of the presence of siblings.

Another confound on the relationship between whether toddlers had a sibling and the visual input is that we had to group in participants with siblings of very diverse ages in the binary “sibling” category: toddlers with older siblings able to play with them were grouped in with toddlers whose siblings were infants. It could be that a wider sample containing many participants with siblings of a variety of ages, which would allow comparison of participants with older siblings and younger, infant siblings, would reveal that only siblings of a certain age have an impact on the statistics of the visual input.

That said, it could be that the lack of an effect of the presence of siblings here does in fact reflect the reality: that pragmatic aspects of the environment, such as the number and characteristics of people present, are not the primary factors shaping the frequency of faces and bodies in the toddler view. Firstly, it could be that external factors are not very formative at all. There is some evidence to support this— for example, Xu et al. (2011) find that toddler visual input dynamics do not vary much between solitary and with-parent interactions. This could mean that elements of the social environment such as the number of individuals present are unrelated to the statistics of the visual input.

It could also be that external factors only impact finer-grain aspects of the visual input. These finer-grain aspects could range from specific face identity and demographics through to

emotional expressions. Particularly, previous findings on increased infant face processing proficiency for children with siblings (Cassia et al., 2013) could reflect differences in the kinds of faces and bodies that toddlers with and without siblings see, and not necessarily the frequency with which they view faces and bodies overall. Prior findings on increased socioemotional proficiency observed in children with siblings could also be explained by differences in the specific information available in the faces and bodies. Specifically, it could be that toddlers with siblings do not see more faces but see a wider variety of facial expressions; or it could be that they see a wider variety of bodily expressions. Future research should examine whether sibling status is associated with differences in the nature of the faces and bodies in the toddler view—such as the average ages, the frequency of different emotional expressions, or the frequency of different motor actions.

Temperament and the Naturalistic Input of Faces and Bodies

Hypotheses 3 and 4, which suggested longitudinal temperament would be related to different frequencies of faces and bodies in the toddler view, was not supported by our analysis. Our results showed no significant relationships between longitudinal temperament and the frequencies of faces and bodies in the toddler view. So, perhaps the internal variable of longitudinal temperament is not so influential in shaping the frequency with which toddlers see faces and bodies. Hypotheses 5 and 6, which suggested that concurrent temperament would be related to different frequencies of faces and bodies, respectively, was not supported by our analysis. Our results showed no significant relationships between concurrent dimensions of

negativity and the frequencies of faces and bodies in the toddler view. So, perhaps the concurrent temperament does not influence the degree to which toddlers see faces and bodies.

The fact that neither concurrent nor longitudinal temperament were related to the frequency with which toddlers see faces and bodies is surprising, given the prior research on temperament and looking behaviour towards socio-emotional stimuli like faces (Rennels, Kayl, and Kulhanek, 2020) and parent-child relationships (Andreadakis et al., 2020; Bryan and Dix, 2009). This led us to expect concurrent temperament to directly shape the visual input, possibly through directing the activities the child was involved in or their looking behaviour, and longitudinal temperament to indirectly shape the visual input, possibly through its impact on social relationships in the family home. However, this was not the case. One possible explanation for the null result for temperament is that we had a relatively small sample size, which reduced the statistical power of the tests conducted here. In future, methodologies tested and used here can be deployed in recruiting a wider sample of toddlers.

However, it could also be that temperament is truly not related to the frequency with which toddlers see faces and bodies. Firstly, concurrent temperament may not be relevant. While some of the prior research on this topic (Rennels, Kayl, and Kulhanek, 2020; Martinos et al., 2012; Rajhans et al., 2015) found phenomena such as subtle differences in looking behaviour at faces themselves, much of it finds relationships between temperament and very specific elements of processing different facial emotional expressions. So, it might be that temperament does not affect overall face and body attending behaviour and therefore visual

input per se but affects the visual input of facial and possibly bodily emotional expressions. To this end, future studies should examine the relationships between fine-grained aspects of the visual input like facial expression and concurrent temperament.

A similar process could be reflected in longitudinal temperament. The prior research on the impact of infant temperament on relationships with caregivers (Andreadakis et al., 2020; Bryan and Dix, 2009; Levitzky and Cooper, 2000) often demonstrates a relationship between infant temperament and caregiver emotional state specifically. So again, it may not be that temperament longitudinally affects broad face and body frequency in the toddler view through shaping these relationships. It may only be that temperament longitudinally affects the specific caregiver behaviour and emotional expressions directed at the toddler. Future research should therefore examine how these specific aspects of the face and body visual input relate to longitudinal temperament.

However, it could also be that these results reflect no relationship between temperament and the socio-emotional visual input. While temperament may be related to face and body processing, as well as later social development (Martinos et al., 2012; Rajhans et al., 2015; Young, Fox and Zahn-Waxler, 1999), this relationship may not emerge through temperament shaping socioemotional input. Innate, shared maturational processes may instead be formative of both temperament and socio-emotional development. For example, face processing ability is highly heritable (Wilmer et al., 2010), as are elements of social cognition (Baron-Cohen et al., 2013) as is temperament (Cloninger et al., 2019). So, shared genetic

variance between these elements could be consistent with maturational face and body processing development as opposed to an explanation contingent on temperament shaping development through the visual environmental input.

Overall Conclusion

In conclusion, this study represents the first analysis of factors that may correlate with toddlers' socioemotional visual input, specifically examining factors associated with the frequency with which they see faces and bodies. Toddlerhood is a time of rapid developmental progression in many domains, and prior research (including our own in chapter 1) suggests wide variability in the statistics of the visual input (Fausey et al., 2016; Jayaraman et al., 2015, 2017; Sugden et al., 2014). Our results showed no significant differences in the frequency of faces and bodies in the view of toddlers with siblings as compared to the view of toddlers without siblings and no robust effect of longitudinal or concurrent temperament.

This is the first examination of its kind, and while we find null results, it could be that more specific aspects of the faces and bodies within the toddler view are related to individual differences like temperament or the presence of siblings. Further research should examine face and body frequency alongside these candidate finer-grain aspects of the visual input, such as specific emotional expressions and identities present, with wider samples. Individual differences beyond those that we have examined here may also be explanatory of individual differences in face and body visual input.

**Adult Perception of Naturalistic, Dynamic Emotional Expressions: Consensus and
Disagreement**

Emma J. Jackson, Elena Geangu

This paper has not yet been submitted for publication

Abstract

Adult perception of emotional expressions is an important topic: understanding to what degree adults show consensus in their judgements of real-world emotional expressions is essential in building accurate models of social cognition, as well as in building computer vision models able to replicate aspects of this. Prior research has focused almost exclusively on posed, static and artificial emotional expressions, which are highly unrepresentative of the kinds of judgements adults make in their everyday lives. Using a corpus of 311 short video clips extracted from naturalistic interactions between individuals, mostly family members and mostly in the home environment, we sought to establish the degree to which consensus on facial emotional expression judgements arises for these naturalistic emotional expressions, as well as the ways in which these stimuli were perceived in terms of intensity and confidence ratings, using ratings from 149 typical adults. Results suggest a high degree of ambiguity with variation by emotion valence, with positive emotional expressions overall being less ambiguous. There are differences in ambiguity, perceptions of intensity and perceived task “difficulty” (as measured by confidence ratings) by the final emotion label valence and subject age group of each clip; positive emotional expressions exhibited by adults and children and negative emotions exhibited by infants were associated with less ambiguity and higher intensity and confidence ratings. FEE rated as being more intense were less ambiguous to human raters and associated with higher confidence. This contradicts traditional models of facial emotional expression processing which often conceptualise this process as easy and unitary for typical adults.

Keywords: naturalistic, expressions, FER, dynamic, perception, social

Adult Perception of Naturalistic, Dynamic Emotional Expressions: Consensus and Disagreement

The human face provides a rich avenue for extraction of information, intentionally communicated or otherwise, for the typical adult observer (Jack and Schyns, 2017). Many models of just how precisely emotional expressions can be defined and understood have been proposed over time, and broadly these models fall into two categories: the first, tracing their origins back to the work of Darwin, propose that discrete emotional expressions are innate, universal and evolutionary adaptations defined by and conveying directly underlying internal states, with these emotional states “leaking” even if attempts are made to mask them (Ekman, 2006; Matsumoto et al., 2008). The second category encompasses contemporary models that hold a range of views but generally are united in disagreement with one or more aspects of traditional models, typically suggesting that FEE do not correspond one-to-one with internal states and that interpreting FEE is highly context-dependent (Barrett et al., 2019; Bavelas and Chovil, 1997).

Many traditional models define facial emotional expressions (FEE) as categories of facial movement defined by the underlying affective state of the person producing the emotional expression. Within these models, the FEE is taken to be solely defined and intrinsically related to the emotions and internal states experienced by the FEE producer (Ekman, 2006). Traditionally, this research utilises static, posed photographs that it is assumed can be

unambiguously and near-universally categorised according to the affective state that the captured facial poses convey.

However, challenges have been levied not only at this methodology, but at this model of emotional expressions as a whole, with suggestions that in reality emotional expressions do not univariately and unambiguously convey internal state but may be a form of potentially ambiguous nonverbal communication that is interpreted alongside contextual information (Barrett et al., 2019). While these criticisms often address the lack of ecological validity in stimuli used in emotional expressions research, relatively little research has examined what characterises categorization of real facial movements into FEE by observers—emphasising the *social* aspect of facial emotional expressions, or what is conveyed to an observer. Here we will discuss different approaches to FEE and our current work on characterising naturalistic facial movement categorization into FEE categories by typical adult observers.

Traditional Views on Facial Emotional Expressions

Over a century ago, Darwin's work on human emotional expressions (Darwin and Prodger, 1998) laid the foundation for theories suggesting that facial emotional expressions (FEE) evolved and serve adaptive functions for social animals. These early Darwinian ideas were expanded upon by Ekman (1999), who proposed the existence of universal, basic emotions, suggesting that emotional expressions are not only biologically rooted but also culturally universal. These "basic emotion theories," would suggest that facial movements directly represent internal emotional states, with internal states "leaking" onto the face even when deception is attempted (Ekman

and Friesen, 1969; Ekman, 2009). This direct and somewhat univariate relationship between internal states and FEE forms the core of basic emotion theory. Partially because of these principles, FEE are assumed to be universal and somewhat ambiguous.

Basic emotion theories would suggest the existence of discrete, categorical and universal basic emotions. Typically, there are six– anger, happiness, sadness, disgust, fear, and surprise. These “basic” emotions are associated with specific prototypes, which are exaggerated or “pure” forms of FEE. The forms of these FEE are considered to have evolved to serve specific adaptive functions, such as wide eyes to enhance threat surveillance (Darwin, 1872; Lee and Anderson, 2017).

These methodologies and assumptions have underpinned much of modern research on emotional expressions, with posed, static facial expressions, judged against “correct” answers based on predetermined scoring systems, such as the Facial Action Coding System (FACS). FACS, introduced by Ekman and Friesen in the 1970s, provides a framework for coding specific facial movements into action units that correspond to facial muscle movements (Ekman and Friesen, 1978).

Advances in recording and computer vision technologies have made it possible to test some of the long-standing assumptions of basic emotion theory in more ecologically valid contexts, including whether the six “basic” emotions proposed by Ekman (happiness, sadness, anger, fear, surprise, and disgust) (Ekman, 1999) actually are the “basic” emotions, occurring often and categorised unambiguously by typical observers.

Contemporary Challenges to Traditional Views on Facial Emotional Expressions

Contemporary views that challenge the classical understanding range from a simple acknowledgement that understanding intent or context is highly formative or essential in shaping judgments of emotional expression (Fridlund and Russell, 2006; Aviezer et al., 2017) through to a rejection of the traditional view in favour of reconceptualizing emotional expressions entirely. This shift in perspective often frames facial expressions as dynamic patterns of movement that may not always correspond to internal emotional states and that are shaped by many cultural, social, and contextual factors (Barrett et al., 2019).

One of the main critiques of traditional models is their over-reliance on artificial, posed stimuli in highly controlled environments, which may not accurately reflect how FEE are actually produced or perceived in real life. The research conducted using the static photographs of exaggerated facial expressions of Ekman, Sorenson, and Friesen (1969) may appear to demonstrate the universality and unambiguity of FEE. However, posed, artificial expressions are far more exaggerated than the subtle facial movements we may more often see in real-world interactions and therefore could be easier to categorise (Motley and Camden, 1988). The high agreement when it comes to ratings of prototypical, posed emotional expressions, could be an artefact of their artificial nature, representing stereotypes more than prototypes (Barrett et al., 2019).

FEE can often be symbolic or directly communicative, rather than abstract expressions of emotion. For example, Bavelas and Chovil (1997) note that an essential aspect of the human face is its mobility, which allows for a wide range of subtly different movements. Fine-grain differences in these movements can serve to create symbolic acts of communication beyond a direct sign of internal affective state. Many similar expressions can therefore convey very different symbolic meanings. For instance, pantomimed shock when telling a story is understood to convey affective information from *at the time of the event* rather than *at the time of retelling*, as conveyed by very subtle aspects of this dynamic facial display (along with the context). Barrett et al. (2019) argue that traditional basic emotion conceptions of FEE lack reliability (the same emotion can be expressed through many different facial movements), specificity (there is no one-to-one mapping between facial movements and emotions), and generalizability across cultures and contexts. They suggest that rather than being universal signals of internal states, FEE are more accurately conceptualised as patterns of facial movements that function as socially meaningful gestures within a specific context. These gestures are created, moderated and modified for social communication, and will be selectively deployed to convey appropriate messages. For example, humans use display rules to modify expressive behaviour and produce socially acceptable communication (Malatesta and Haviland, 1982).

There may not be discrete mappings between facial movements and the messages that they convey. Durán, Reisenzein, and Fernández-Dols (2017) review the research in this field for macaques, finding that the same facial movements can occur in many contexts and for many

purposes, which could indicate a primarily communicative function. There is also evidence for low co-occurrence between emotions and expressions in humans (Durán and Fernández-Dols, 2021). So, FEE may be highly context dependent communication rather than direct signaling of emotion (Barrett et al., 2019). Contextual cues, including body language, may also be essential for accurate interpretation of FEE (Aviezer et al., 2012; Aviezer et al., 2017). Prior knowledge is also a contextual cue that is essential in shaping how facial movements are interpreted (Palumbo and Jellema, 2013). This suggests that FEE are rarely judged in isolation but are instead processed within a broader, context-rich framework.

Given that the purpose of FEE within social species is communication and given the challenges to the traditional view outlined here, contemporary models often consider FEE at two levels: the affective, which concerns the relationship between facial movements and internal emotional states, and the socio-communicative, which focuses on how facial movements function as a form of nonverbal language. For example, the same smile might serve as both an indicator of a positive emotional state and a social signal of approachability and so may serve as an expressive and affective or social and communicative role (Fridlund and Russell, 2006). This dual perspective acknowledges the complexity of FEE, recognizing that they can convey both emotional and social information simultaneously. Understanding how typical observers within a given culture categorise and extract information from the kinds of facial poses and movements that they are typically exposed to could be informative in knowing how typical (rather than prototypical) FEE are interpreted in regular interpersonal interaction.

In summary, contemporary challenges to traditional views on FEE highlight the importance of context, naturalistic stimuli, and communicative function in understanding how facial movements are produced and perceived. Rather than being unambiguous signals of internal emotional states, FEE are dynamic, context-dependent social tools that vary across cultures and individuals. This perspective suggests that facial expressions are better understood as complex, communicative gestures shaped by both affective and socio-cultural factors, rather than as direct reflections of universal emotional experiences. Thus, considering and defining FEE at the level of observer categorisation and perception is crucial for understanding real-world FEE.

The Importance of Naturalistic Research and Stimuli

One theme that emerges consistently in discussions on the topic of basic emotion theories versus newer approaches to facial emotional expressions is that of posed and artificial versus naturalistic and dynamic stimuli. Much of the foundational basic emotion research relies heavily on artificial, posed and static FEE acted in response to specific instructions (e.g., Ekman, 1999; Ekman, 2006; Ekman, 2009; Ekman and Friesen, 1969; Ekman and Friesen, 1978; Ekman, Sorenson, and Friesen, 1969). However, even Ekman's research suggests that there are likely to be differences between artificial and posed FEE, with some AUs suggested to be more reliable indicators of true emotion (Ekman, 2003).

There have been attempts to generate spontaneous, more naturalistic and often dynamic FEE using a variety of stimuli, from smells to videoclips (e.g., Petrides et al., 2013; Pfister et al.,

2011; Simons et al., 2003). Using these databases and computer-vision based AU detection, it does seem that posed and spontaneous FEE are discriminable using several metrics, from AU intensity to the specific AUs employed (Jia et al., 2021).

However, the above databases of spontaneous FEE are not “truly” naturalistic. To produce these FEE, participants are exposed to a variety of tasks designed to elicit affective states within a lab setting, typically alone or with one non-interacting experimenter. This means that these categories of FEE may not fully reflect the kinds of facial movements that are present within the visual environments of humans as they navigate an unconstrained social environment.

The Importance of Understanding the Perception of Facial Emotional Expressions

Understanding how adults categorise real facial movements into emotional expressions is important when it comes to a wide range of academic fields and research topics. For example, understanding human perception of FEE can inform research on typical and atypical social functioning (Yi et al., 2016), and in computer vision the average human perception of real FEE can serve as the ground truth for behavioural detection models (Canedo and Neves, 2019).

Understanding the characteristics of real FEE and how they are interpreted is crucial given how central FEE production and understanding is to typical social interaction (Frith, 2009). FEE processing ability seems to be clinically significant, being associated with higher-order skills such as theory of mind (Poulin-Dubois et al., 2018), and typically developing adults are assumed to interpret FEE accurately and at speed (Vogindroukas et al., 2014; Linnankoski et al., 1994).

Deficits in these skills have been implicated in disorders from Autism to Antisocial Personality Disorder (Eack et al., 2015; Lombardo et al., 2007; Marsden et al., 2019). However, without understanding how FEE processing presents with real-world facial movements, understanding of these deficits is limited.

Understanding how adults perceive real-world FEE is essential for developing accurate computer vision and facial emotion recognition (FER) systems. AI models can rapidly code images for action units and FEE, work that could take many years for human coders (Benitez-Quiroz, Srinivasan, and Martinez, 2016). However, for any real-world applications, FER algorithms must provide accurate labels for real-world expressions. Training models on highly controlled, prototypical and static images can limit their ability to do this (Krumhuber et al., 2021). The expressions seen in naturalistic settings are likely to be more ambiguous and less prototypical than those used in traditional datasets, making the application data to real FEE likely far more challenging (Cowen and Keltner, 2020). This is particularly problematic when these algorithms are applied to truly "in-the-wild" footage, where factors such as varied face views, makeup, blur, occlusion, and poor lighting can drastically impair model accuracy even assuming the training data to be accurate (Masi et al., 2016; Wang and Deng, 2021; Karahan et al., 2016). There is indeed evidence that FER models significantly underperform on naturalistic FEE as compared to both posed FEE and human coder performance on the same stimuli (Dupré et al., 2020; Küntzler, Höfling, and Alpers, 2021). This would suggest that an overreliance on artificial data has indeed made real-world FEE challenging for these models.

As outlined, our understanding of both typical human behaviour and accurate FER model ground truths is limited given that much of the extant research on both typical and atypical FEE processing development has focused on responses to posed, artificial stimuli that are markedly different from and likely less ambiguous than naturalistic FEE (e.g., Ekman, 1999; Ekman, 2006; Ekman, 2009; Ekman and Friesen, 1969; Ekman and Friesen, 1978; Ekman, Sorenson, and Friesen, 1969). This may misrepresent the true nature of FEE production and processing, meaning that research with naturalistic stimuli is essential. So, for both understanding human behaviour and creating accurate FER systems, we must establish how humans process real-world facial movements. While processing and labelling of the prototypical lab stimuli could be a task with high consensus, even across cultures, it is likely that real-world emotional expressions are more ambiguous, less distinctly categorised and less universal.

Prior Research on Naturalistic Facial Emotional Expression Perception and Outstanding

Questions

Emotional expressions are assumed here to serve two overlapping functions: (1) expressing an underlying affective state, such as a spontaneous smile in response to a positive event, and (2) communicating information to observers, as with a polite smile directed at a stranger. A single expression, like a smile initiated by happiness but exaggerated for friendliness, can fulfil both functions, through modulation by social display rules. While subtle differences may exist and be perceived between spontaneous and socially motivated smiles (Surakka and Hietanen, 1998), most typical adult observers are likely to categorise them as different versions of the same broad emotional expression (Miller et al., 2022). Therefore, we assume that the

facial emotional expression (FEE) attributed to a video clip by typical observers from a given culture can serve as one "ground truth" within that cultural context. This means that we are focusing on FEE from the perspective of observer processing and categorisation. To understand FEE processing and categorisation, we must know how it presents in response to the kinds of facial movements and postures that people actually observe around them.

Some research has already addressed how best to characterise observer processing and categorization of FEE. One example is the work by Cowen and Keltner (2020), which examined categorical, appraisal, and free-response judgments to 1,500 highly varied static face-body photographs sourced from the internet. They identified 28 distinct emotions, from anger and fear to triumph and interest, and found that these emotions exist on a continuum. They demonstrated that emotion categories explain more variance than general appraisals like valence and arousal, with expressions being reliably recognised across contexts. The authors here took an approach that to some extent mirrors the approach that we will employ: they avoided prototypical stimuli, collecting an assortment of images rather than generating them, and took the "correct" emotional expression label for each photograph to be the consensus judgement, rather than pre-assigning labels. Though we will not use as many labels or free responses, their findings highlight the variability and subtlety in non-prototypical expression ratings, including categories less commonly studied, such as "doubt" and "pride," alongside the traditional "basic" emotions.

Not only this, but with follow-up self-report examining perceptions of the labels generated, participants reported how often they believe themselves to see emotions fitting each of the labels. Many of the secondary emotion labels were reported as being more common than some of the basic emotions like fear. Their study also demonstrated substantial variability and ambiguity in emotional perception, even considering the many FEE labels available to them. For example, 66% of the images showed less than 50% agreement between participants, and 8% of stimuli failed to reach the 25% agreement threshold required for reliable categorization. This suggests that more naturalistic stimuli, which are not prototypical, lead to more ambiguous FEE judgments than traditional research would suggest. These findings challenge the assumption that real, representative FEEs can be easily and unambiguously categorised.

Interestingly, Cowen and Keltner also explored differences in recognition rates (a measure of agreement) between certain emotions. While not all six “basic” emotion labels were examined, recognition was calculated for most of them. Anger had the highest recognition, followed by surprise, sadness, disgust, and fear. This contrasts with prior studies using highly controlled, artificial stimuli, which found happiness to be the most and fear the least easily recognized emotion as measured by accuracy (Palermo and Coltheart, 2004). It remains unclear whether such fine-grain distinctions will hold for the dynamic stimuli that we focus on. It is also unclear if recognition rates are a result of the facial movements themselves varying in ambiguity or may be at least in part the result of naturally occurring differences in intensity between typical presentations of different FEE.

While Cowen and Keltner (2020) provide the most comprehensive bottom-up examination of perceptions of varied FEEs, their study used static photographs from the internet, at least some of which were likely posed to some degree. They also do not offer much insight into other aspects of how emotions may be perceived in expressions. While this research tells us in great detail *which* FEE participants perceive to be present in sampled stimuli, it does not give us as much detail on how participants perceive these expressions, particularly in terms of ambiguity, confidence, or intensity. To address this gap, we will focus on a more limited set of FEE labels—specifically, the six basic emotions and “neutral”—to enable exploration of how participants categorise natural facial movements within these labels and the extent of their agreement in greater detail. Ensuring that these six FEE are examined will also allow for clear comparisons between our research and the “basic” emotions research that would suggest low ambiguity for these categories.

The use of dynamic FEE is particularly important for understanding real-world social communication. Prior research suggests that human interpretation of emotional expressions improves when dynamic information is available. For instance, Van Der Zant and Nelson (2021) examined affective judgments of tennis players based on both static and dynamic images. Simply put, participants were asked to make judgements as to whether the player pictured in the stimulus had just won or lost a point. Participants were more accurate when viewing dynamic, full-body stimuli, which highlights the limitations of static, closely cropped stimuli typically used in FEE research. The poor participant performance for static images and cropped

faces highlights that the static, artificial and closely cropped stimuli typically used in FEE research do not well approximate real-world affective state judgement. One potential explanation for this is that typical participants engage with and visually scan dynamic versus static images of faces in different ways (Horlin et al., 2013). Dynamic elements of a facial display may also in and of themselves convey affective information (Bould et al., 2008).

Clearly, there is limited research addressing the topic of categorisation of truly naturalistic, dynamic FEE. The question of whether these facial movements can be reliably categorised as distinct FEE and whether these categories reflect previously proposed unambiguous basic emotional expressions is an important one, and there is currently little evidence to directly address this. Partially, this is likely because acquisition of truly naturalistic, dynamic emotional expressions is arduous. There are many elements of FEE perception that are likely to be important.

Agreement, Intensity and Confidence

Here, we will focus on three key aspects of participant response to real-world FEE: agreement between participants, participant intensity ratings and participant confidence ratings. These three elements have frequently been examined and can give a comprehensive view of how clear or ambiguous an FEE is, as well as factors underpinning agreement of ambiguity. Currently, there is no comprehensive research exploring these factors for real-world FEE, nor how they may vary across different emotions or individual differences in the FEE producer/model (e.g.,

age group). Understanding these dimensions is crucial for building accurate models of how people process potentially ambiguous real-world emotional expressions.

Naturally, agreement between raters can give a good sense of how challenging, variable or ambiguous an expression or FEE category is for human perceivers. This is one of the measures presented by Cowen and Keltner (2020) as explanatory of differences between FEE categories. So, comparing agreement rates across categories of interest not only provides one metric of the ambiguity of a stimulus or group of stimuli but allows for comparison with prior research using less naturalistic facial movements.

Intensity ratings are another important factor. Prior studies using posed FEEs have shown a strong relationship between accuracy and intensity (Guo, 2012; Hess et al., 1997; Palermo and Coltheart, 2004), with more intense emotions being easier to recognize. If ambiguity or disagreement is higher for certain emotions, we must consider whether this is linked to lower perceived intensity. Further, as we will include a variety of age groups in our stimuli, and as emotional display rules vary with age (Malatesta and Haviland, 1982), we anticipate differences in intensity across age groups. This could be particularly true for negative emotions, where downregulation is socially desirable and so these display rules apply strongly. Therefore, we are likely to have a wide range in FEE intensities, especially for negative emotions, which could be explanatory of differences in agreement or ambiguity.

Confidence ratings can offer additional insight into how participants perceive their judgments. This is an important and separable variable from “accuracy” (agreement) when it comes to understanding how participants are categorising FEE. Prior research examining challenging FEE categorization tasks has demonstrated that participant confidence ratings do not always align with accuracy (Grenville and Dwyer, 2022). Therefore, confidence ratings can provide a useful, separate measure of participants’ conscious perceptions of what their own “accuracy” is likely to be. This can give a sense of whether participants perceive their judgement to be making their best-guess, or whether they feel that they are making accurate decisions. For example, low consensus with high confidence could indicate that participants are genuinely and clearly perceiving subjectively different FEE, whereas low confidence could indicate that they are randomly guessing. The relationship between confidence and intensity could also be informative-- for example, participants could perceive more intense FEE as being easier to categorise.

The Influence of Emotion Category and Stimulus Model Age Group

It is likely that there are differences in specific aspects of FEE processing between different emotion categories and different categories of model or producer. Both the emotion category and the FEE model individual differences are likely to influence participant responses. One model individual difference that we will focus on here specifically is the model age group.

There are good theoretical reasons to believe that participant response may differ by the FEE category a stimulus is assigned to. Emotion processing is not entirely domain general. For

example, the ability to recognize different FEEs develops at different stages, with neonates recognizing happiness but not fear, with the latter ability emerging later in development (Farroni et al., 2007). Different FEE categories are also detected and processed differently in adulthood-- for example, which specific FEE is exhibited on a face can moderate face detection ability (Zhang et al., 2012). Psychiatric conditions in adulthood often affect the recognition of specific FEEs rather than FEE processing as a whole (Doty et al., 2013). This would indicate that, to a certain extent, processing and detection of different FEE entails qualitatively different mechanisms. Therefore, it is important to examine how specific FEE categories, rather than emotions overall, influence participant responses. For example, one FEE category being perceived in unambiguous ways does not mean that ambiguity is also low for other FEE categories.

Similarly, there is evidence to suggest that the age group of the FEE model is another key variable in a sample of highly heterogeneous stimuli such as ours. Prior research suggests that FEE action units differ significantly across age groups (Onal Ertugrul et al., 2023), and experience with a particular age group can moderate model age-related differences in FEE recognition (Plank et al., 2022). Indeed, the very existence of instances of some discrete FEE in infancy (like clear fear and anger) has been called into question (Camras et al., 2007). This means that to accurately assess ambiguity, intensity and confidence in ratings of FEE exemplars taken from the natural visual environment, it is necessary to separate stimuli by subject age group. Despite this, age group is a factor that has not been directly examined for naturalistic FEE databases before. Many papers examining FEE processing and categorization that attempt to use more naturalistic

stimuli or methods– e.g., the aforementioned Van Der Zant and Nelson (2021) -- only include adult stimulus subjects. Even when such papers do include child or infant stimulus subjects, as is the case with the above Cowen and Keltner (2020), responses are not examined by stimulus subject age group.

So, we will examine the role of both model age group and FEE category in shaping agreement, alongside confidence and intensity ratings, generated by typical adults presented with a wide range of naturalistic facial movements. There is limited prior research on this.

Summary of Outstanding Questions

The traditional approach to FEE would suggest that many of the kinds of facial movements that are seen in everyday life can unambiguously be categorised into FEE categories, with high levels of consensus across participants (e.g., Ekman, 1999; Ekman, 2006; Ekman, 2009; Ekman and Friesen, 1969; Ekman and Friesen, 1978; Ekman, Sorenson, and Friesen, 1969). Not only this, but the traditional view would make no specific predictions made about differences in ambiguity, intensity or confidence depending on the FEE category or model age group.

However, there is little evidence supporting this idea of unambiguous FEE, and much existing research only examines artificial or posed FEE. Prior research using more open labelling and semi-naturalistic stimuli suggests that less consensus exists than previously assumed, but without in-depth examination of how intensity, confidence, and model age group influence these judgments (Cowen and Keltner, 2020).

Further, very little research uses truly naturalistic FEE, with semi-“naturalistic” clips or photographs from internet searches and films (Dhall et al., 2011) or evoked emotion in-lab (e.g., Petrides et al., 2013; Pfister et al., 2011; Simons et al., 2003) typically used. This means that all work on expressions is limited in its understanding of how adults perceive actual real-world FEE. This gap impacts on many research areas, from examination of adult social cognition to validation of FER computer vision models. This study aims to fill this gap, providing a more detailed understanding of FEE perception in real-world, dynamic facial movements.

Aims and Hypotheses

Currently, there is little research examining the kinds of FEE that humans are exposed to in their everyday lives, and there is even less research on how these emotional expressions are categorised by typical human observers. We will collect data to assess the latter question using excerpts of a wide pool of naturalistic videos rated by typical observers. We seek to address the outstanding question of just how typical adults perceive and label the kinds of facial movements that they see in everyday life. We will establish how typical adults categorise a sample of naturalistic, dynamic facial movements taken from naturalistic, mostly egocentric recordings. We will also collect confidence ratings in these assessments as well as FEE intensity ratings. We seek to examine the degree to which the six “basic” emotions and neutral are actually unambiguously perceived by typical adults in a wide variety of real-world facial movements. We also seek to assess the degree of ambiguity in this perception, and what the sources of this ambiguity might be. We are interested in how final FEE labels and video clip subject age groups may drive ambiguity, as well as differences in rating confidence and perceived intensity.

Here, ambiguous stimuli are taken to be those that demonstrate low consensus between participants. Given the work of Cowen and Keltner (2020), which demonstrated high ambiguity and disagreement even with only semi-naturalistic clips, we expect relatively low consensus. They found that 66% of clips received consensus of 50% or lower, and we expect lesser consensus for our stimuli. Also, based on Cowen and Keltner's examination of ambiguity by FEE label, we expect the highest ambiguity to be for fear and the lowest ambiguity to be for anger. Given prior research on FEE by age, and great differences in FEE exhibited by infants (Camras et al., 2007; Onal Ertugrul et al., 2023), we expect an effect of age on ambiguity, with the greatest ambiguity for infant FEE.

Consequently, we expect a similar relationship between age groups and confidence and FEE and confidence, with raters being least confident for fear and infant FEE. For intensity, we expect adults on average to exhibit more intense FEE as unlike infants and children they have fully internalised FEE display rules (Malatesta and Haviland, 1982). Partially because of these display rules, and because of the fact that these clips were recorded as children and adults interacted within typical, naturalistic environments, which are unlikely to frequently elicit strong threat-related expressions, we expect differences in intensity by FEE label, with instances of fear being less intense on average. Finally, given prior research on intensity being strongly associated with accuracy (Guo, 2012; Hess et al., 1997; Palermo and Coltheart, 2004) and the likely relationship between accuracy and confidence, we anticipate that intensity, accuracy and confidence will all be positively correlated with one another. Overall, we hypothesise that:

1) these naturalistic expressions will be ambiguous, with a relatively low level of agreement; 2) further, ambiguity will differ by FEE, being highest for fear and lowest for anger; 3) ambiguity will differ by model age, being greatest for stimuli containing infants; 4) confidence ratings will differ by video subject age group, with lower confidence ratings for infant clips on average; 5) confidence ratings will differ by FEE label, with participants being least confident for fear on average; 6) intensity ratings will differ by video subject age group, with higher intensity ratings for infant clips on average; 7) intensity ratings will differ by FEE label, with fear ratings being the least intense; 8) Intensity and confidence ratings will both be positively associated with participant agreement and each other.

Methods

Participants

158 participants were recruited using the participants recruitment system at the University of York. Four pilot participants and five participants who made errors in trial progression were excluded from analysis, leaving 149 participants with usable data (155 female, 25 male and 9 other). To take part, participants confirmed that they had primarily grown up in a Western, predominantly white country and that they had not been diagnosed with a disorder which would affect their processing of faces or facial emotional expressions. 133 of the participants were right-handed, 15 left-handed and one was ambidextrous. 129 were or had previously been Psychology students at undergraduate level or higher (20 had not). Ten were paid participants, the rest participated for course credit. Eight participants were or had previously been the primary or secondary caregiver to a baby, toddler or child (see table 1).

Table 1*Participant characteristics*

Characteristic	N	Percentage (%)
Gender		
Female	115	77.18
Male	25	16.78
Other	9	6.04
Handedness		
Right-handed	133	89.26
Left-handed	15	10.07
Ambidextrous	1	0.67
Psychology Degree		
Yes (current/past/equivalent)	139	93.29
No	10	6.71
Compensation		
Monetary voucher	10	6.71
Course credit	139	93.29
Primary/Secondary Caregiver to a Child		
Yes	8	5.37
No	141	94.63
Total	149	100

Note: This table provides a summary of the participants by demographic characteristic.

Stimuli

311 short video clips were used as stimuli. These clips were selected by six researchers from approximately 600 hours of footage. The members of the lab aimed to find clear examples of one of the six basic facial emotions or a neutral expression in short clips of good visual quality (i.e., not dark or blurry) containing only 1 face. 259 of the selected clips were recorded in the home environments of families with young children. Most of these clips were recorded with a head-mounted camera from the parent or child egocentric perspective, but some were recorded with a video camera mounted stationary within the room. The remaining 52 clips were

recorded with head-mounted cameras over the course of naturalistic social interactions in a lab room. For full details of the stimulus video clips, including preliminary FEE label, see table 2.

Table 2

Stimulus Video Clip Characteristics

Characteristic	N	Percentage (%)
Recording Environment		
Home	259	83.28
Lab	52	16.72
Age Group		
Adult	108 (95 female)	34.73
Children	160 (95 female)	51.45
Infants	43	13.83
Preliminary Researcher FEE Label		
Happiness	66	21.22
Anger	68	21.86
Disgust	22	7.07
Fear	24	7.72
Sadness	50	16.08
Surprise	56	18.01
Neutral	25	8.04

Note: This table provides a summary of the participants by demographic characteristic.

Preliminary label selections were made using existing descriptions of the six “basic” FEE (anger, happiness, sadness, disgust, fear, and surprise) and their associated action units within the research literature (Ekman, 1999), with neutral being the absence of an expression. These preliminary selections were used to guide the stimulus search and for randomization of stimuli.

Two versions of each video were used: one at full speed, and one at half speed. The inclusion of the half-speed presentation was motivated by the fact that these videos tended to be dynamic and brief, with rapid changes in recording perspective and view. This is characteristic of videos

recorded from the egocentric perspectives of individuals as they move their bodies and heads to observe and navigate the home environment. The average number of frames per video was 26.77 (minimum = 4, maximum = 323); at corresponding frame rates, this translates to an average full speed video length of approximately 1.80 seconds (min = 0.27 seconds, max = 5 seconds) and an average half speed video length of approximately 3.60 seconds (min = 0.54 seconds, max = 10 seconds). While information may be easily extractable from such brief face presentations by a person directing their own attention, viewing these egocentric videos solely at full speed can be disorienting for a third party. So, the presentation of each video twice—once at full speed, once at half speed—was intended to prevent underestimation of the information that typical individuals extract from their own egocentric views. The videos were played to participants with the sound muted. This is because, as outlined previously, while the audio environment is an important source of social information, here we were interested only in unimodal, naturalistic FEE.

Apparatus

Given that the videoclips were highly sensitive, it was not possible to use an online experiment builder for presentation of stimuli, and local experiment builders were unable to handle the volume of visual data. For this reason, the videoclips were presented in PowerPoint, and questions about the video clips were presented in Qualtrics on a separate screen.

After piloting to decide upon the maximum number of clips one participant could reasonably give ratings for within an hour, the dataset was split randomly into quarters, with each quarter

containing 78-80 sets of clips (with each set containing two versions of the same videoclip, one full speed and one half-speed). Each stimulus quarter was semi-randomized, as it was not possible to dynamically randomise clips within PowerPoint. Based on the preliminary lab ratings, concurrent presentations of clips likely to contain the same facial emotional expression was avoided.

Procedure

Set-Up, Instructions and Demographics.

Testing took place in an individual room containing a chair, desk, computer and small tablet. Participants were randomly assigned to a “group”, which was a quarter (1, 2, 3 or 4) and an order within this (a or b). These were not different experimental groups, but just different stimulus groups.

Having read the information sheet and signed the consent form, participants were asked to fill in some demographic questions on Qualtrics on a tablet while the experimenter loaded the first block on the computer screen. These questions asked the participants to confirm that they did not meet the exclusion criteria, asked them about their gender and handedness, asked them about whether they were or ever had been Psychology students at level 4 or higher, and asked them whether they had ever been a caregiver (primary or secondary) to a baby, toddler or child. They were also asked to specify which country they primarily grew up in. Visual and verbal instructions were given to participants on how to complete trials. At this point, the experimenter switched off the light and left the room, initiating the experiment. Before the

main trials, each participant completed two practice trials. In-between practice trial 1 and 2, the experimenter briefly returned to check for questions and reiterate aspects of the instructions. There were five blocks of trials; at the end of each block, the participant was instructed that they could rest as long as needed.

Trial Structure.

Each trial encompassed a few stages involving a PowerPoint display on the computer and a Qualtrics questionnaire on the tablet. This trial structure is outlined in table 3 and figure 1.

Table 3

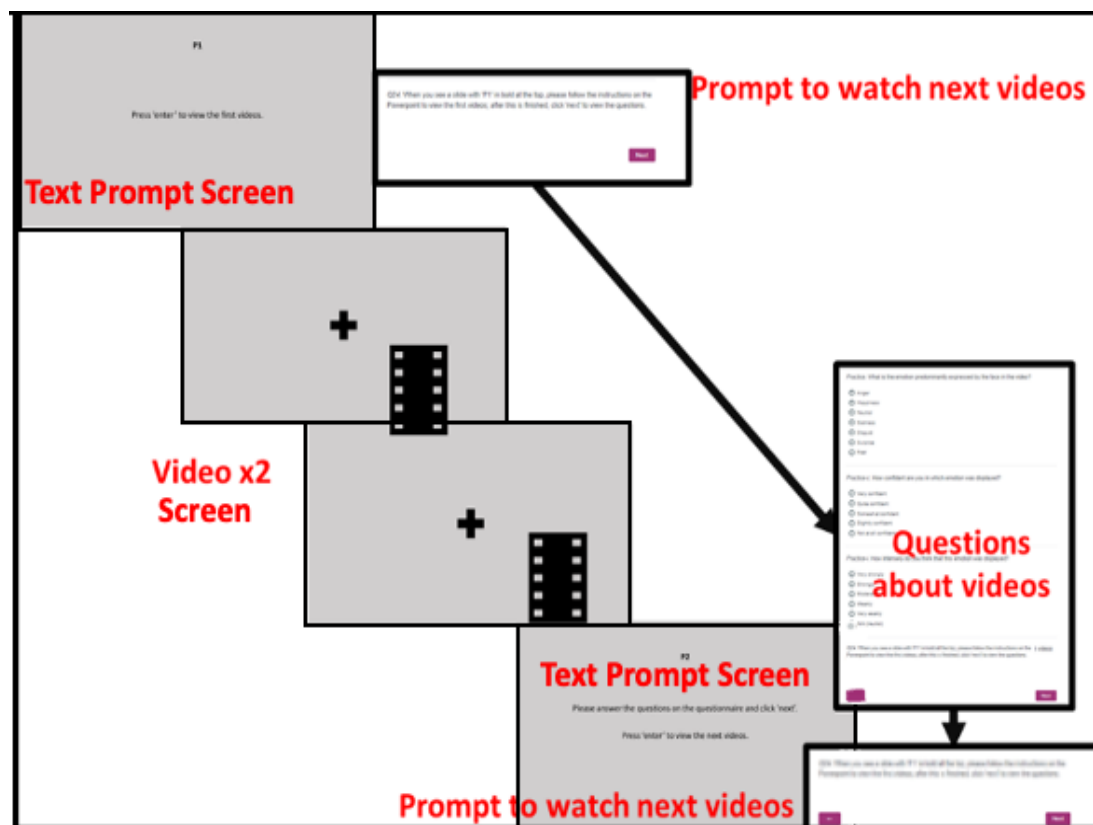
Trial Structure and Screens

Screen Category	Location	Description
Text prompt screen	PowerPoint	This instructed participants to answer the questions about the previous video set on Qualtrics before pressing 'enter' when ready to view the next videos.
Prompt to watch next videos	Qualtrics	This reminded participants to follow PowerPoint instructions to view the next set of videos.
Fixation cross 1	PowerPoint	This was presented centrally before the first video.
Video screen (full speed)	PowerPoint	This played the first video in a set (full speed) once and without sound.
Fixation cross 2	PowerPoint	This was presented centrally in between the first and second videos.
Video screen (half speed)	PowerPoint	This played the second video in a set (half speed) once and without sound.
Question screen	Qualtrics	Asked participants about the emotion they saw, their confidence, and the intensity of the expression.
Checkpoint screen	Qualtrics	This was the input screen for the checkpoint. This ensured synchronisation between PowerPoint and Qualtrics by asking participants to select a specific letter on Qualtrics.
	PowerPoint	This presented participants with a specific letter to select on Qualtrics. This ensured synchronisation between the stimulus (PowerPoint) and input (Qualtrics) platforms.

Note: This table provides a summary of the screen types presented to participants over the course of the trials and checkpoints.

Fig. 1

The overall structure of one trial within the Powerpoint and in Qualtrics.



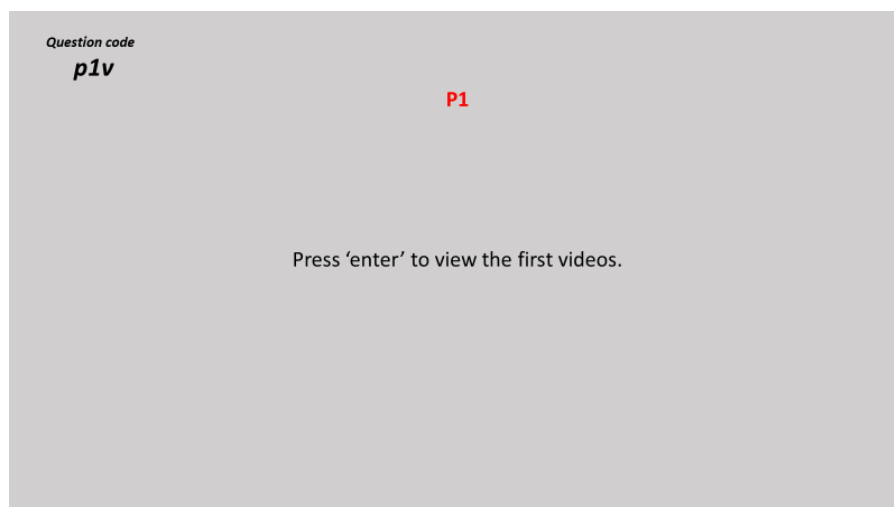
Note: This is an illustration of the very first trial. The next trial would begin with the participant pressing “enter” to view the next videos, so text prompt screens separate video sets.

At the beginning of each trial, on the PowerPoint, participants saw a text prompt screen. If this was the very first trial, it would say “Press ‘enter’ to view the first videos”; if it was any other trial, it would say “Please answer the questions on the questionnaire and click ‘next’. Press ‘enter’ to view the next videos.”

Note that the example in figure 2 is from the first practice trial, P1, and so contains extra information—there is the code P1 on the text prompt screen, and the text ‘Question code’ above the question code. The PowerPoint text prompt screen contained a question code at the top left approximately every three trials—this was to allow experimenters to visually check progress and that the PowerPoint and tablet trials were matching.

Fig. 2

The first practice trial text prompt screen in PowerPoint: this is for P1, the first practice trial.

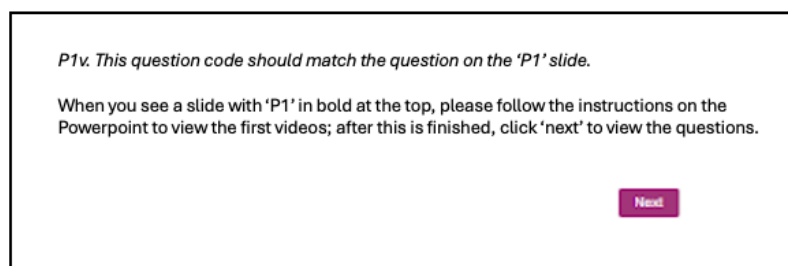


While this screen was visible on the PowerPoint, on Qualtrics, participants were presented with a text prompt screen telling them to follow the instructions on the PowerPoint to view the set of videos (figure 3). Participants then pressed ‘enter’ on the computer to play the set of videos. The PowerPoint automatically progressed through the full speed and half speed version of the set of videos, with fixation crosses in between. Participants were instructed to keep the

Qualtrics text prompt screen up while they viewed the videos, and to never press 'enter' while they could see a video or a fixation cross, as after the text prompt screen each trial progressed automatically. After seeing the set of videos, participants would then see the next text prompt screen.

Fig. 3

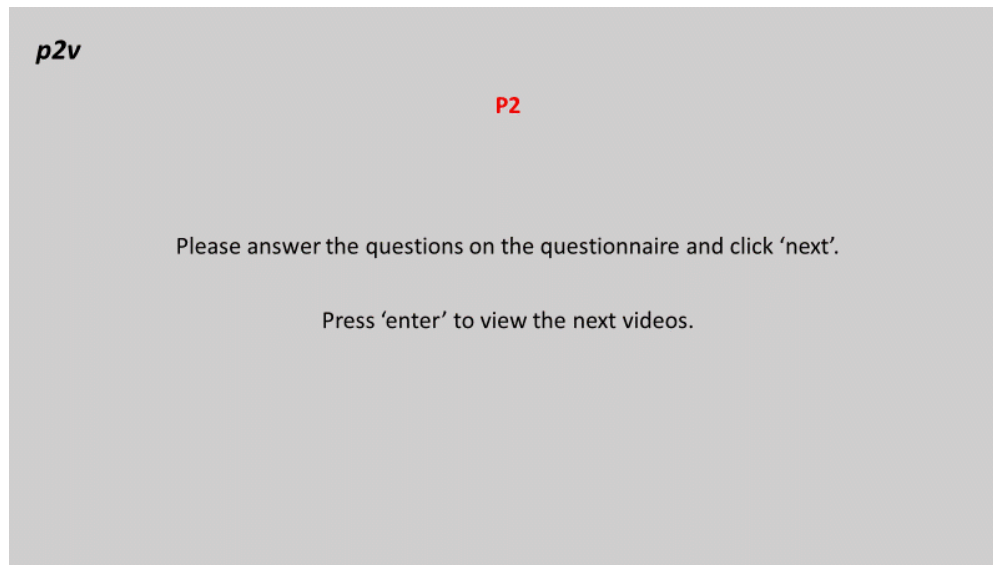
The first practice trial text prompt screen on Qualtrics.



Before pressing 'enter' to view the next videos, participants would then answer three questions about the set of videos they had just seen while a new text prompt screen was displayed on PowerPoint (see figure 4). They were asked "What is the emotion predominantly expressed by the face in the video?" with forced choice between seven options: Anger, Happiness, Neutral, Sadness, Disgust, Surprise or Fear. They were then asked two Likert scale questions: "How confident are you in which emotion was displayed?", with five options (confidence); and "How intensely do you think that this emotion was displayed?", with six options including "N/A (neutral)" (intensity). Participants were required to complete these questions and progress through on Qualtrics to the next text prompt screen before viewing the next set of videos.

Fig. 4

The text prompt screen after the videos of practice trial 1.



The experiment was split into five blocks for each participant. Within each block there were either 15 or 16 trials with a “checkpoint” at the end of each. Participants were informed that they could take breaks as needed between blocks. The whole procedure lasted an hour per participant, with an experimenter observing them and providing any assistance needed throughout.

Data Analysis Strategy

For each videoclip, there were 37-38 individual ratings; each participant rated 78-80 videoclips, which was roughly a quarter of the dataset. For each video clip, the number of ratings given for each emotional expression was calculated. The number of ratings was used to calculate the

consensus, which was the percentage of participants agreeing with the final FEE label.

Confidence-weighted labels were then calculated: each individual rating was multiplied by the corresponding confidence rating (“Very confident” = 5; “Not at all confident” = 1); this provided an overall count*confidence score for each FEE label, and the highest score was considered to be the final “winning” FEE label. The consensus/agreement was the percentage of participants that had agreed with this winning label. Intensity scores also ranged between 1 and 5 depending on the option selected; if participants forgot to select N/A for a neutral rating, this was replaced with the number corresponding median rating option (3).

Based on this data, several analyses were conducted. Firstly, the presence of the seven different categories of facial movement (anger, disgust, sadness, neutral, happiness, surprise and fear) within the 311 stimulus video clips, in addition to how this compared with our speculative labels, was examined. Presence of different FEE was broken down by stimulus model age group. Adult models were those considered to be 18 years or older, child subjects were those over the age of 24-months but younger than 18-years, and infant subjects were 24-months or younger. We then examined how the following varied by stimulus model age group and final FEE label: agreement/consensus; intensity ratings; and confidence ratings. Finally, we examined the relationships between agreement, intensity and confidence.

Results

Presence of Emotional Expressions in the Dataset

Based on confidence-weighted participant ratings, of the 311 videoclips, 220 clips (70.74%) were assigned the same FEE as the initial category allocation assigned by lab researchers.

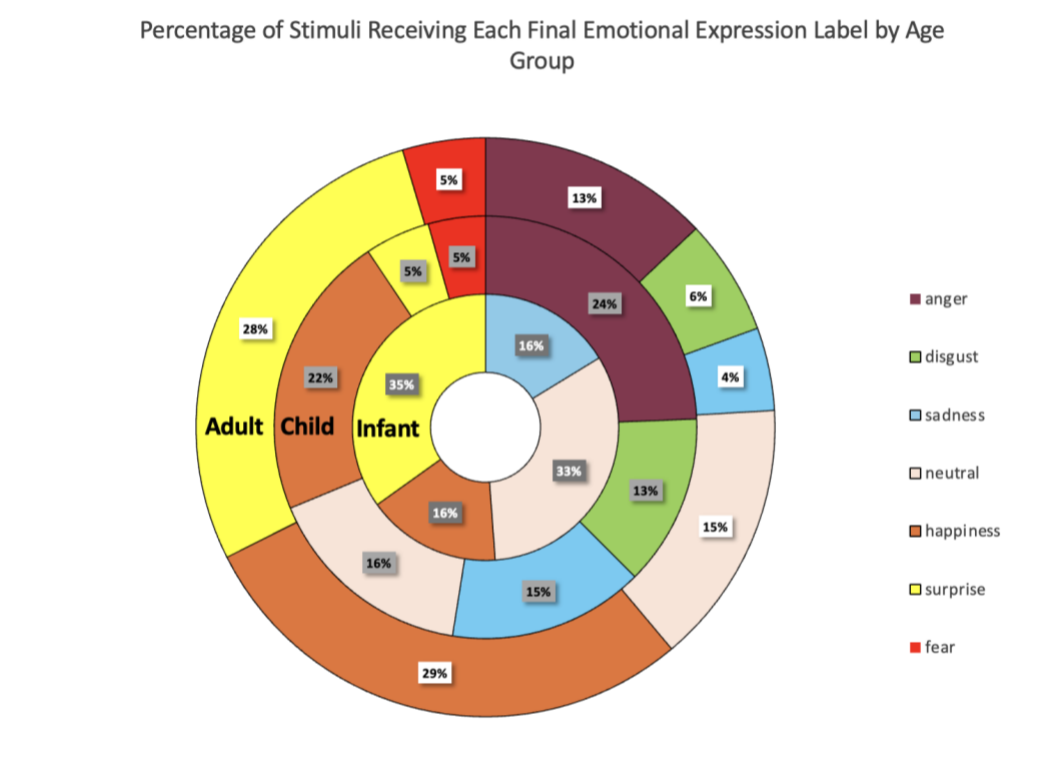
Table 4

Agreement between researcher and participant categorisation.

		Participant Confidence-weighted Rating Label						
		anger	disgust	sadness	neutral	happiness	surprise	fear
Lab Rating	anger	44	7	4	10	1	1	1
	disgust	1	15	0	2	4	0	0
	sadness	5	4	29	8	0	1	3
	neutral	0	0	1	21	1	2	0
	happiness	0	0	1	3	62	0	0
	surprise	1	0	0	9	5	41	0
	fear	2	2	1	3	0	8	8
<i>Sum</i>		<i>53</i>	<i>28</i>	<i>36</i>	<i>56</i>	<i>73</i>	<i>53</i>	<i>12</i>

Note. Each value represents the number of clips that received a given lab rating and received a given final participant FEE rating label. Correspondence was generally high, but was lower for anger, sadness and fear.

Overall, happiness was the most common and fear the least. See the pie chart below (fig. 5) for a breakdown of the percentage of clips rated as containing each emotion by age of the video model.

Fig. 5

Note. The inner ring represents clips containing infants, the middle ring represents clips containing children, and the outer ring represents clips containing adults.

In addition to the “winning” (final) FEE label, as calculated by number of ratings weighted by confidence, a “runner up” FEE label was also calculated for all clips without 100% consensus (see table 5). This is informative in understanding which emotions are most often confused for one another by participants. Happiness was most often unanimous, while other FEE were rarely or never unanimous.

Table 5

A table to compare the “winning” FEE label with the runner-up label.

Runner-Up (%)								
	anger	disgust	fear	happiness	neutral	sadness	surprise	
Winning Label (%)	anger	0	41.51	0	5.66	18.87	30.19	3.77
	disgust	35.71	0	7.14	14.29	39.29	3.57	0
	fear	25	0	0	0	8.33	33.33	3.33
	happiness	4.11	5.50	4.11	30.14	26.03	2.74	27.40
	neutral	17.86	1.79	10.71	12.50	0	26.79	30.36
	sadness	47.22	16.67	16.67	2.78	8.33	5.56	2.78
	surprise	9.43	1.89	32.08	15.09	32.08	1.89	7.54

Note. Each value represents the percentage of clips in a “winning” label row that received the runner-up FEE label denoted by the column. When these labels are the same, this indicates the percentage of clips receiving this FEE label with no runner-up (i.e., those that unanimously received the label along the left).

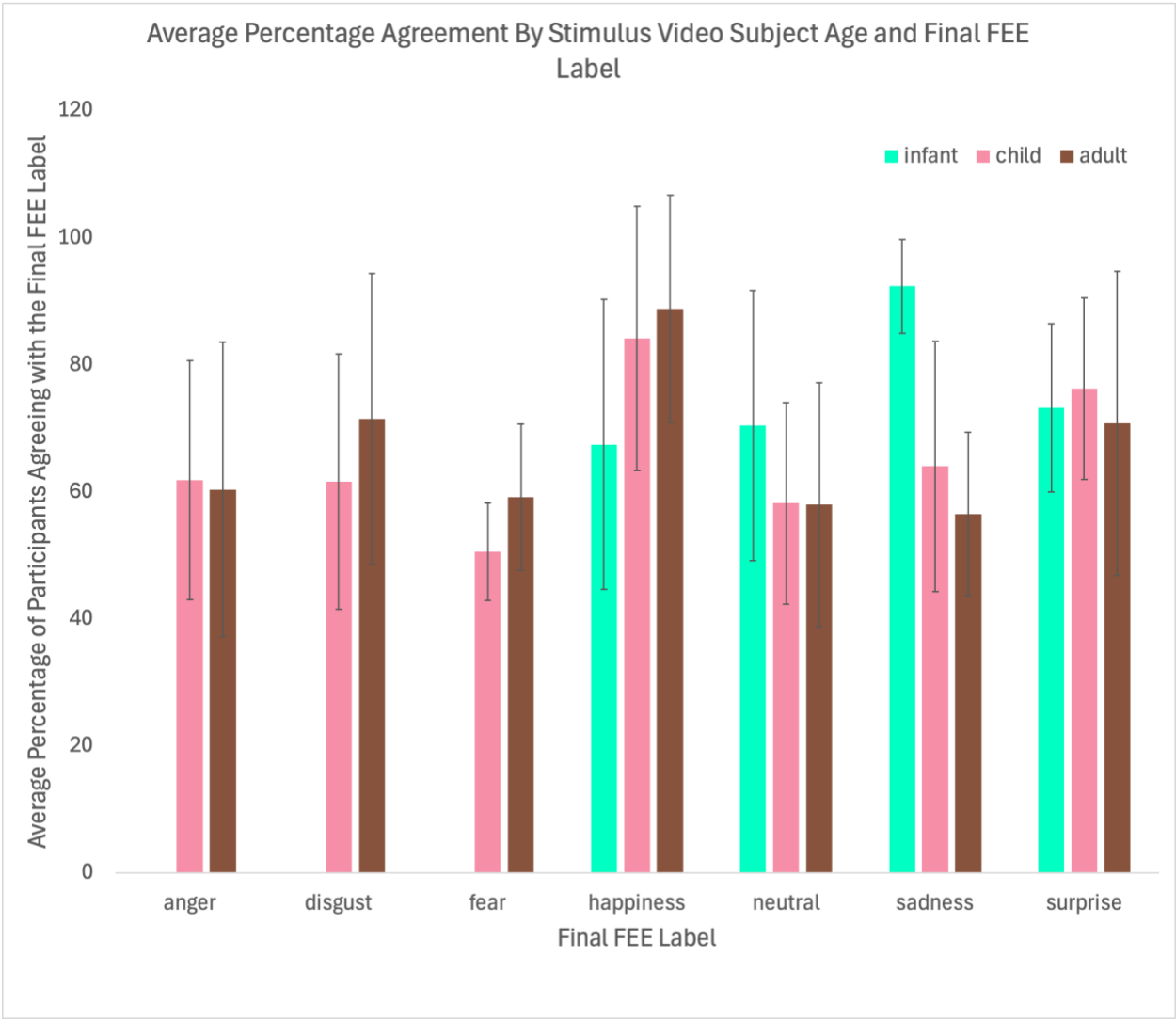
Consensus with Final Label Across Clips by Video Subject Age Group and Facial Emotional Expression

On average across video clips, the mean consensus/agreement with the final confidence-weighted FEE label was 69.41% (SD = 21.73, Min = 24.32, Max = 100). There were five video clips that, based on percentage of ratings alone before confidence weighting, had no clear consensus. One of these clips had a final confidence-weighted label of ‘happiness’, one ‘surprise’, one ‘anger’, one of ‘neutral’ and one of these clips ‘sadness’. There were 28 stimuli

that received unanimous response for one emotion. Though, only 25.7% of our stimuli failed to exceed 50% consensus with the final FEE label. In short, there was a wide range of agreement across the individual stimulus videos.

Fig. 6

Average videoclip agreement by final emotional expression label assigned to the videoclip subject and age group of the video clip subject.



Note. This bar chart represents the mean and standard deviation (error bars) agreement with the final FEE broken down along the final FEE label and stimulus video model age group.

The average degree of agreement varied a little overall by age group of the video model, being highest on average for infants and lowest on average for children. The average degree of agreement varied somewhat by overall emotional expression label assigned to the video subject, being highest on average agreement was for clips receiving a confidence-weighted label of 'happiness' and lowest on average for 'fear'. Agreement scores were then broken down by both video clip subject age and the final FEE label a clip received (see fig. 6). The average agreement was highest for stimuli containing infant subjects that received a final FEE label of sadness ($M = 92.31\%$, $SD = 7.39\%$); it was lowest for stimuli containing child subjects that received a final FEE label of fear ($M = 50.54\%$, $SD = 7.66\%$).

Given the fact that some categories of negative facial emotional expression were not observed in the infant age-group, to assess the role of age-group and emotion category together in predicting average participant agreement, the stimuli were collapsed into three broader valence categories: positive expressions (happiness, surprise); negative expressions (anger, disgust, fear and sadness); and neutral. See figure 7 for a summary of average agreement for these broader emotional expression categories. The average agreement was highest for stimuli containing infant subjects that received a final FEE label of a negative emotion ($M = 92.31\%$, $SD = 7.39\%$); it was lowest for stimuli containing adult subjects that received a final FEE label of neutral

emotion ($M = 57.93\%$, $SD = 19.21\%$). Overall, agreement was highest for infants ($M = 74.45$, $SD = 18.67$) followed by adults ($M = 71.51$, $SD = 23.47$), then children ($M = 66.63$, $SD = 20.98$).

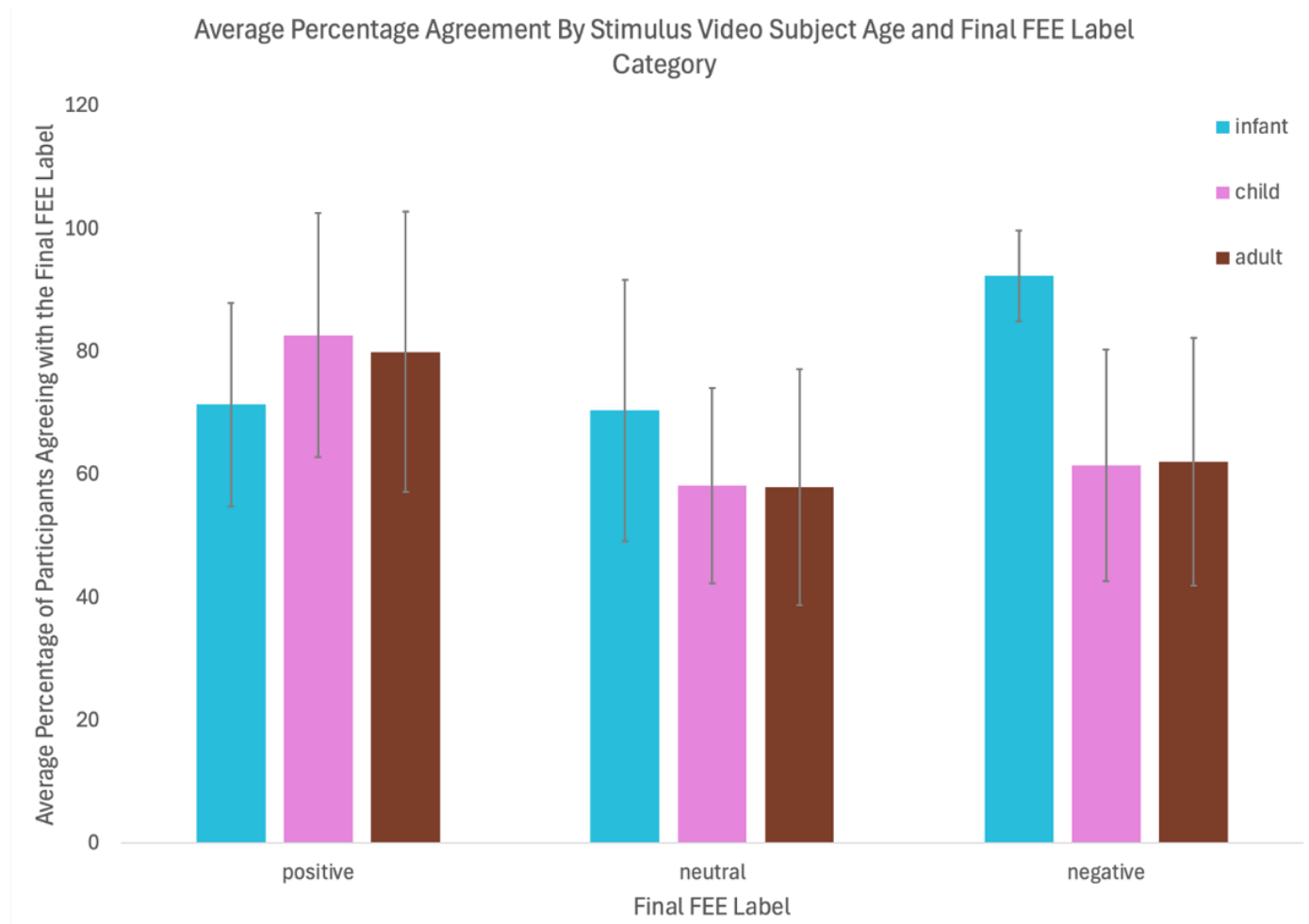
The data did not meet the assumptions for a two-way ANOVA, as Levene's test for equality of variances was violated for the present analysis, $F(8, 302) = 2.313$, $p = .020$. A GLM analysis was conducted to test the impact of video clip subject age and final confidence-weighted FEE label valence on percentage agreement with the label. The omnibus test was significant, $\chi^2(8) = 73.46$, $p < .001$, indicating a good overall model fit. Model fit indices (AIC = 2743.01, BIC = 2780.41) indicated a reasonable fit. The main effect of emotion valence was significant, $\chi^2(2) = 22.70$, $p < .001$, indicating that the valence category influenced the percentage agreement. The main effect of age was also significant, $\chi^2(2) = 9.42$, $p = .009$, suggesting that the age category of the video subjects had an impact on the agreement. However, there was a significant interaction between age and emotion category, $\chi^2(4) = 23.38$, $p < .001$, indicating that the effect of emotion on agreement varied depending on the age group of the video subject.

Parameter estimates revealed that there were significant interaction effects for infants, with the negative valence category yielding significantly higher agreement scores ($B = 38.81$, $SE = 9.38$, $p < .001$) and the neutral valence category also yielding significantly higher agreement scores ($B = 20.99$, $SE = 8.53$, $p = .014$) compared to the positive emotion category. However, no significant interaction effects were found for children or adults. Additionally, compared to the positive valence category (reference category), the negative valence category ($B = -17.85$, $SE =$

4.25, $p < .001$) and neutral valence category ($B = -21.96$, $SE = 5.41$, $p < .001$) were both associated with significantly lower agreement scores— however, this must be interpreted with caution given the interaction between age group and emotion valence category.

Fig. 7

Average videoclip agreement by final emotional expression label valence category (positive, negative, neutral) assigned to the videoclip subject and age group of the video clip subject.



Note. This bar chart represents the mean and standard deviation (error bars) agreement with the final FEE broken down along the final FEE valence category and stimulus video model age

group. There were three categories: positive (happiness/surprise); negative (anger, disgust, fear and sadness) and neutral. Error bars represent standard deviations.

Average Intensity Rating by Video Subject Age Group and Facial Emotional Expression

For analysis of intensity, the 56 stimuli that received a final label of neutral were excluded, as analysis of intensity data was unlikely to be meaningful here. When an average was taken across the remaining 255 stimuli (including ratings that disagreed with the final videoclip rating), the intensity ratings ranged from 1.61 to 4.92 with a mean intensity rating per stimulus of 3.14 (SD = 0.84).

The lowest average intensity rating for an FEE label was 2.79 (SD = 0.74) for clips given the final label fear, and the highest average was 3.44 (SD = 1.03) for clips given a final label of sadness.

Overall, average intensity ratings were very similar when broken down by age group of the video clip subject. On average, clips featuring adult subjects received an intensity rating of 3.26 (SD = 0.84); those featuring child subjects received an intensity rating of 3.02 (SD = 0.79); and those featuring infant subjects received an intensity rating of 3.32 (SD = 0.99). See table 6 for a summary of mean, standard deviation, minimum and maximum intensity ratings for each category of stimulus.

Table 6

A table to demonstrate average intensity ratings by video subject age group and final assigned FEE label.

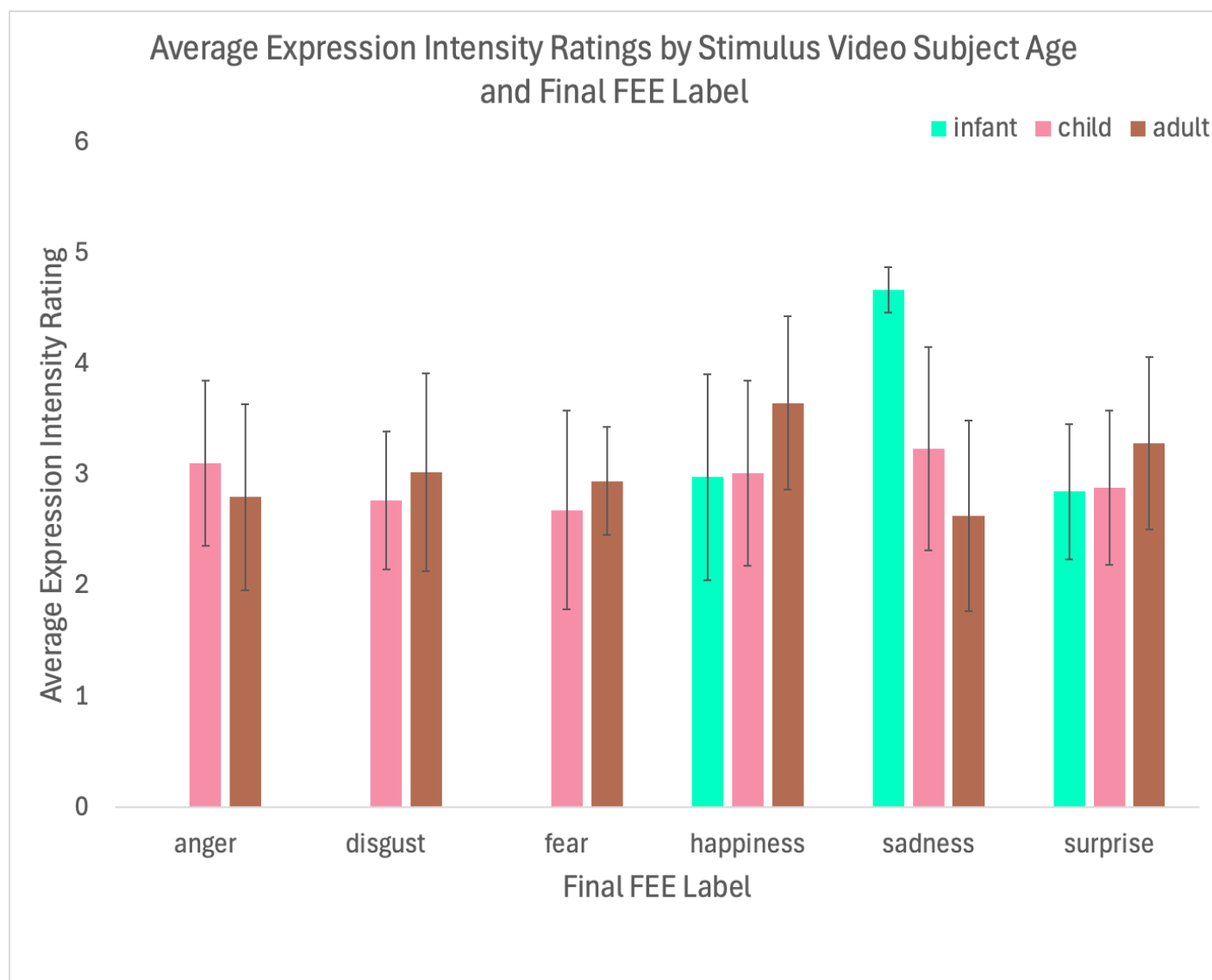
		Stimulus Characteristic							Video Subject Age Group		
		Final FEE Label									
		ang er	disg ust	fear	hap pin ess	sad nes s	sur pris e	Infant	Child	Adult	
Intensity Rating	M	3.03	2.84	2.79	3.29	3.44	3.10	3.32	3.02	3.26	
	SD	0.77	0.69	0.74	0.87	1.03	0.74	0.99	0.79	0.84	
	Min	1.62	1.70	1.61	1.61	1.68	1.73	1.62	1.61	1.62	
	Max	4.49	4.24	4.52	4.76	4.92	4.62	4.92	4.87	4.62	

Note. This table outlines overall intensity ratings by the different categories of stimuli. Figures are presented for each FEE label category and each video subject age group.

Intensity ratings were then broken down by a combination of both final FEE label and subject age (see fig. 8). The highest average intensity ratings were given to stimuli featuring infant subjects that received a final FEE label of sadness ($M = 4.68$, $SD = 0.21$). The lowest average intensity ratings were given to stimuli containing adults that received a final label of sadness ($M = 2.64$, $SD = 0.86$).

Fig. 8

A bar chart to illustrate the average expression intensity ratings by stimulus video subject age and final FEE label.



Note. The mean and standard deviation intensity ratings are broken down along two categories: which final FEE label a clip received, and the age group of the clip subject.

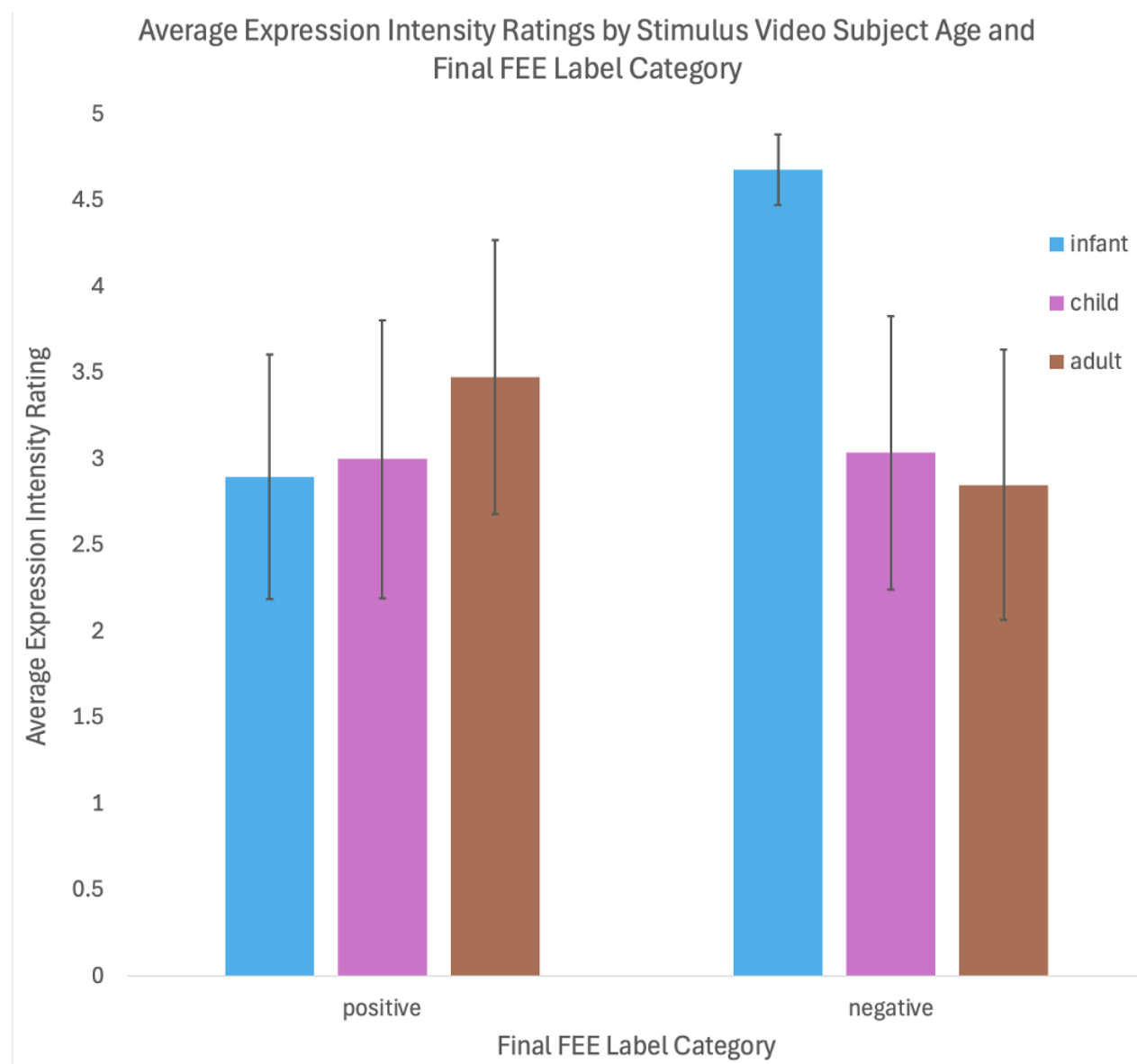
Given the fact that some categories of negative facial emotional expression were not observed in the infant age-group, to assess the role of age-group and emotion valence category together in predicting average participant intensity ratings, the stimuli were divided into two broader categories: positive expressions (happiness, surprise) and negative expressions (anger, disgust, fear and sadness). See figure 9 for a summary of average intensity ratings for these broader emotional valence categories.

The average intensity rating was highest for stimuli containing infant subjects that received a final FEE label of a negative emotion ($M = 4.68$, $SD = 0.21$); it was lowest for stimuli containing adult subjects that received a final FEE label of a negative emotion ($M = 2.85$ $SD = 0.78$).

An ANOVA was conducted to examine how intensity rating varied with the categorical independent variables of stimulus video subject age group and stimulus video final FEE valence (positive or negative). Levene's equality of variances was not violated $F(5, 249) = 1.97$, $p = .084$. There was a significant main effect of video subject age group, $F(2, 249) = 8.81$, $p < .001$, partial $\eta^2 = .066$. There was also a significant main effect of FEE category, $F(1, 249) = 8.69$, $p = .003$, partial $\eta^2 = .034$. However, there was also a significant interaction between the two: $F(2, 249) = 20.50$, $p < .001$, partial $\eta^2 = .141$.

Fig. 9

Average expression intensity ratings by final emotional expression label category (positive, negative) assigned to the videoclip subject and age group of the video clip subject.



Note. This bar chart represents the mean and standard deviation (error bars) intensity assignments broken down along the final FEE label category and stimulus video model age group. There were two categories: positive (happiness/surprise) and negative (anger, disgust, fear and sadness). Error bars represent standard deviations.

Post-hoc Bonferroni pairwise comparisons revealed several significant differences in intensity ratings. For video clips with negative emotional expressions, intensity ratings for infants ($M = 4.68$, $SD = 0.21$) were significantly higher than for children ($M = 3.03$, $SD = 0.79$, $p < .001$) and adults ($M = 2.85$, $SD = 0.78$, $p < .001$). No significant difference was found between the ratings for children and adults. For video clips with positive emotional expressions, intensity ratings for infants ($M = 2.89$, $SD = 0.71$) were significantly lower than for adults ($M = 3.47$, $SD = 0.80$, $p = .009$), but no significant difference was found between ratings for children ($M = 3.00$, $SD = 0.81$) and infants. Positive emotion clips containing adults received significantly higher intensity ratings than negative clips containing adults ($p = .007$).

Additionally, participants rated negative expressions significantly higher in intensity than positive expressions for infants ($p < .001$); they also rated them as significantly lower in intensity than positive expressions for adults ($p = .009$). However, no significant difference was found between negative and positive FEE intensity for clips with child subjects ($p = .796$).

Average Confidence Rating by Video Subject Age Group and Facial Emotional Expression

When averages were taken across all ratings for each stimulus video (regardless of whether the rating agreed with the final “correct” answer) and compared by final FEE label, the highest confidence ratings on average were given for videos that received the final label of happiness ($M = 4.01$, $SD = 0.62$); the lowest on average were given for videos that received the final label of fear ($M = 3.23$, $SD = 0.38$).

Average confidence was highest for videos containing infants ($M = 3.76$, $SD = 0.52$), followed by adults ($M = 3.64$, $SD = 0.67$), followed by children ($M = 3.49$, $SD = 0.56$). For a summary of figures by stimulus FEE and age group separately, see table 7.

Table 7

A table to demonstrate average confidence ratings by video subject age group and final assigned FEE label.

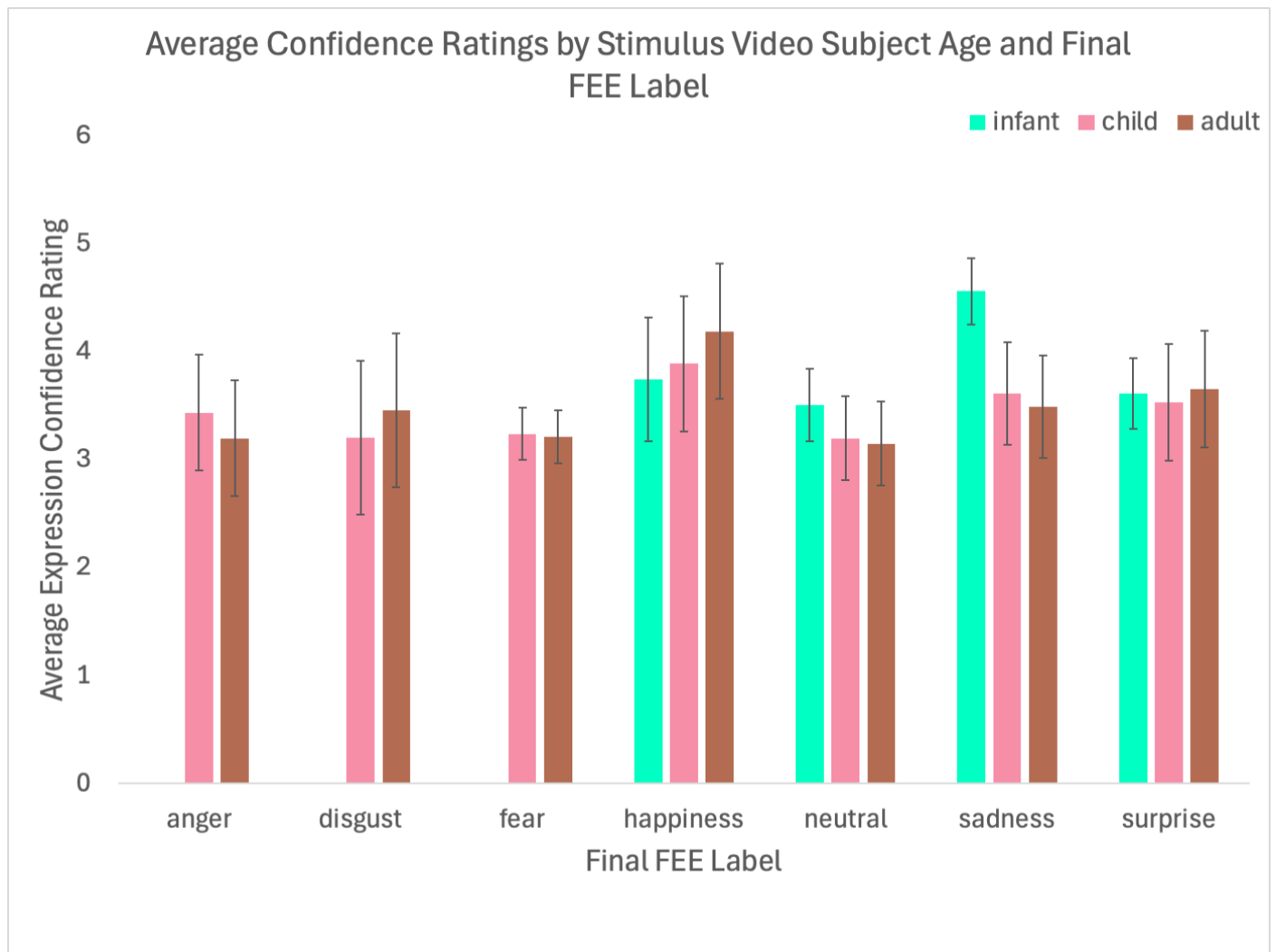
Stimulus Characteristic									Video Subject Age Group		
Conf. Rating	Final FEE Label								Infant	Child	Adult
		anger	disgust	fear	happiness	neutral	sadness	surprise			
	M	3.38	3.27	3.23	4.01	3.26	3.78	3.63			
	SD	0.54	0.51	0.38	0.62	0.39	0.64	0.46			
	Min	2.35	2.59	2.82	2.39	2.24	2.66	2.73			
	Max	4.51	4.35	4.03	4.97	4.30	5.00	4.71	5	4.81	4.97

Note. This table outlines overall confidence ratings by the different categories of stimuli. Figures are presented for each FEE label category and each video subject age group.

Confidence ratings were then broken down by both final FEE label and age group of the clip subject (see fig. 10). The highest average confidence ratings were given for clips containing infant subjects that received a final FEE label of sadness ($M = 4.56$; $SD = 0.31$), while the lowest were given for clips containing adult subjects that received a final label of neutral ($M = 3.15$; $SD = 0.39$).

Fig. 10

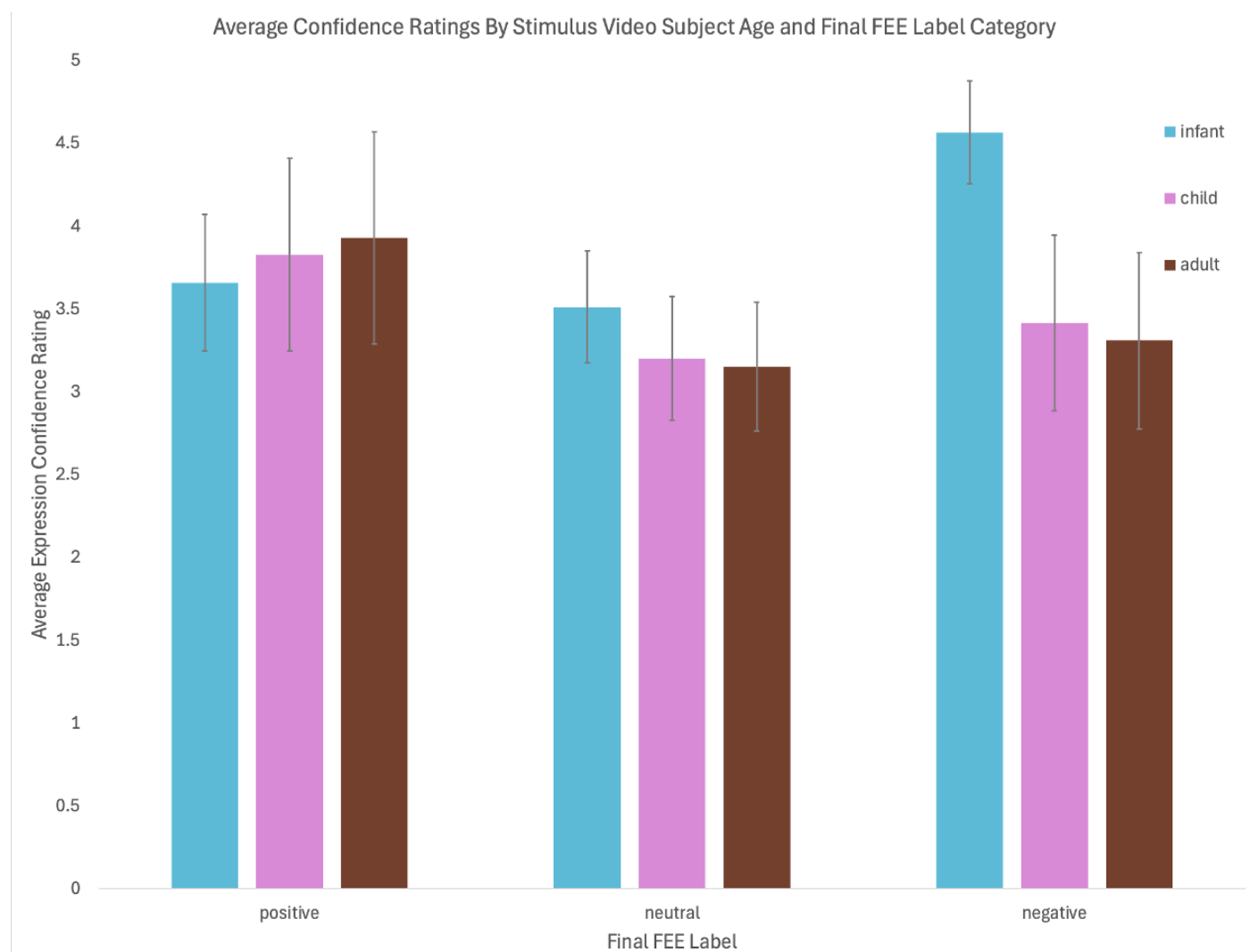
A bar chart to demonstrate average videoclip confidence rating by final emotional expression label assigned to the videoclip subject and age group of the video clip subject.



Note. This bar chart represents the mean and standard deviation confidence ratings broken down along two categories: which final FEE label a clip received, and the age group of the clip subject.

Fig. 11

Average confidence ratings by final emotional expression label valence (positive, neutral, negative) assigned to the videoclip subject and age group of the video clip subject.



Note. This bar chart represents the mean and standard deviation (error bars) participant confidence in their ratings broken down along the final FEE label valence category and stimulus video model age group. There were three valence categories: positive (happiness/surprise), neutral and negative (anger, disgust, fear and sadness). Error bars represent standard deviations.

Given the fact that some categories of negative facial emotional expression were not observed in the infant age-group, to assess the role of age-group and emotion category together in predicting average participant confidence ratings, the stimuli were divided into three broader valence categories: positive expressions (happiness, surprise), neutral expressions and negative expressions (anger, disgust, fear and sadness). See figure 11 for a summary of average confidence ratings for these broader emotional expression valence categories. The average confidence rating was highest for stimuli containing infant subjects that received a final FEE label of a negative emotion ($M = 4.56$, $SD = 0.31$); it was lowest for stimuli containing adult subjects that received a final FEE label of a neutral emotion ($M = 3.15$ $SD = 0.39$).

The data did not meet the assumptions for a two-way ANOVA, as Levene's test for equality of variances was found to be violated for the present analysis, $F(8, 302) = 3.72$, $p < .001$. A GLM analysis was conducted to test the impact of video clip subject age and final confidence-weighted FEE label valence category on average confidence in the label. The omnibus test was significant, $\chi^2(8) = 90.65$, $p < .001$, indicating a good overall model fit. Model fit indices ($AIC = 495.73$, $BIC = 496.46$) indicated a reasonable fit. The main effect of emotion category was significant, $\chi^2(2) = 34.89$, $p < .001$, indicating that the emotion category influenced the confidence ratings. The main effect of age was also significant, $\chi^2(2) = 20.28$, $p < .001$, suggesting that the age category of the video subjects had an impact on the confidence. However, there was a significant interaction between age and emotion valence, $\chi^2(4) = 35.94$, p

< .001, indicating that the effect of emotion on confidence varied depending on the age group of the video subject.

Parameter estimates revealed significant interaction effects—specifically, confidence ratings were particularly high for negative and neutral expressions in infant videos, as reflected in the parameter estimates ($B = 1.526$, $SE = 0.253$, $p < .001$ for negative; $B = 0.631$, $SE = 0.230$, $p = .006$ for neutral). However, no significant interaction effects were found for child or adult subject videos.

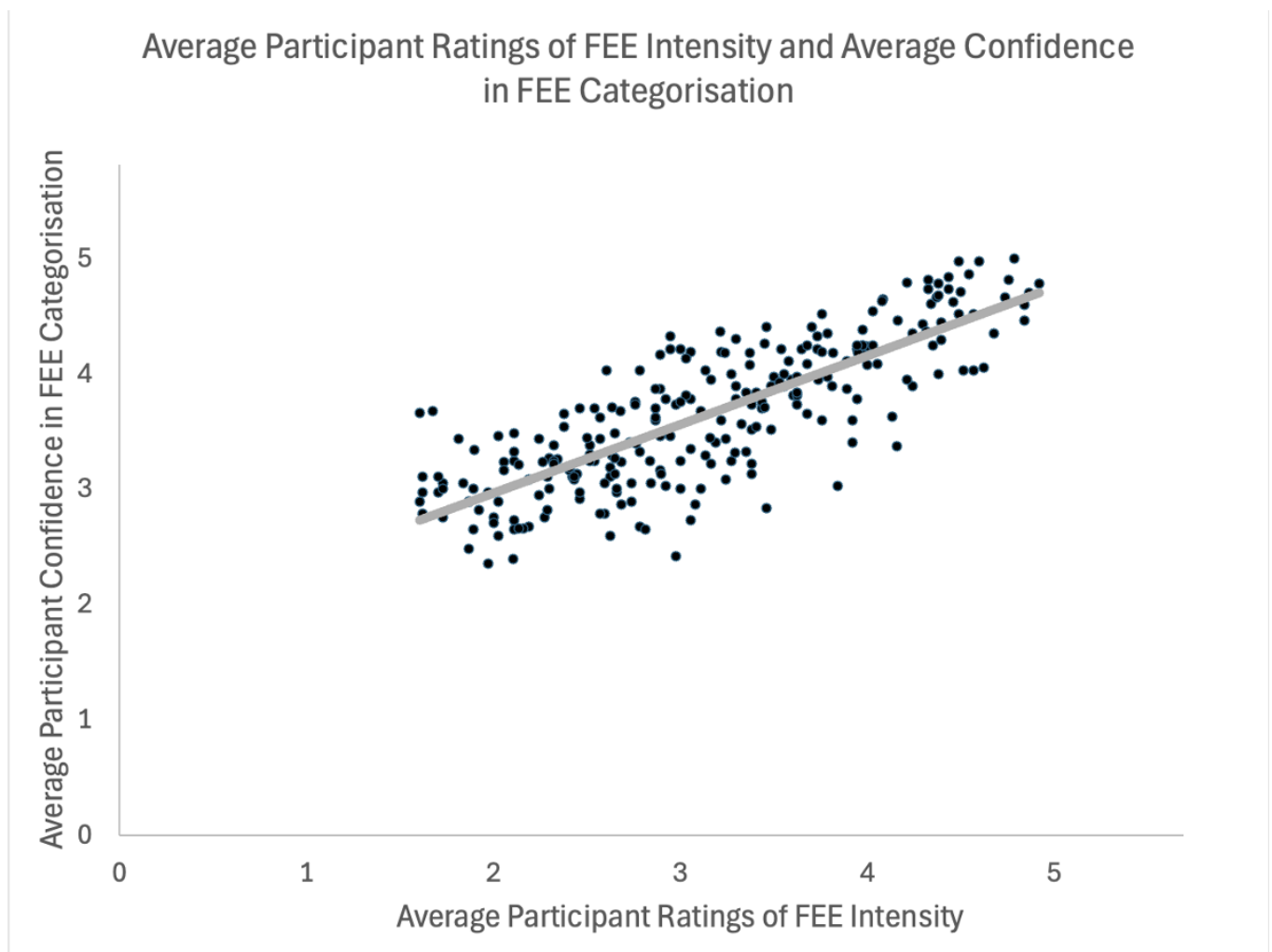
Further, as compared to the positive emotion stimuli (reference category), neutral emotion stimuli were associated with significantly lower confidence ratings overall ($B = -0.778$, $SE = 0.146$, $p < .001$), as were the negative emotion stimuli ($B = -0.620$, $SE = 0.115$, $p < .001$). Pairwise comparisons confirmed this, showing that confidence ratings were significantly lower for neutral expressions compared to both negative ($p < .001$, $MD = -0.476$) and positive expressions ($p < .001$, $MD = -0.518$). However, confidence ratings did not significantly differ between negative and positive expressions ($p = 1.000$, $MD = -0.042$). There was also a significant effect of video subject age group, with expressions exhibited by infants receiving significantly higher confidence ratings than those from both children overall ($p < .001$, $MD = 0.429$) and adults ($p < .001$, $MD = 0.448$). However, there was no significant difference in confidence ratings between expressions exhibited by children and adults ($p = 1.000$, $MD = 0.019$). These findings should be interpreted cautiously considering the above interaction effect.

The Relationship Between Average Intensity, Average Confidence and Percentage Agreement

Next, correlations between confidence ratings, intensity ratings and participant agreement were examined. This is essential in explaining which perceived elements of naturalistic facial expressions may relate to perceived and actual ‘accuracy’ (with confidence being an element of perceived accuracy and agreement our measure of actual accuracy).

Fig. 12

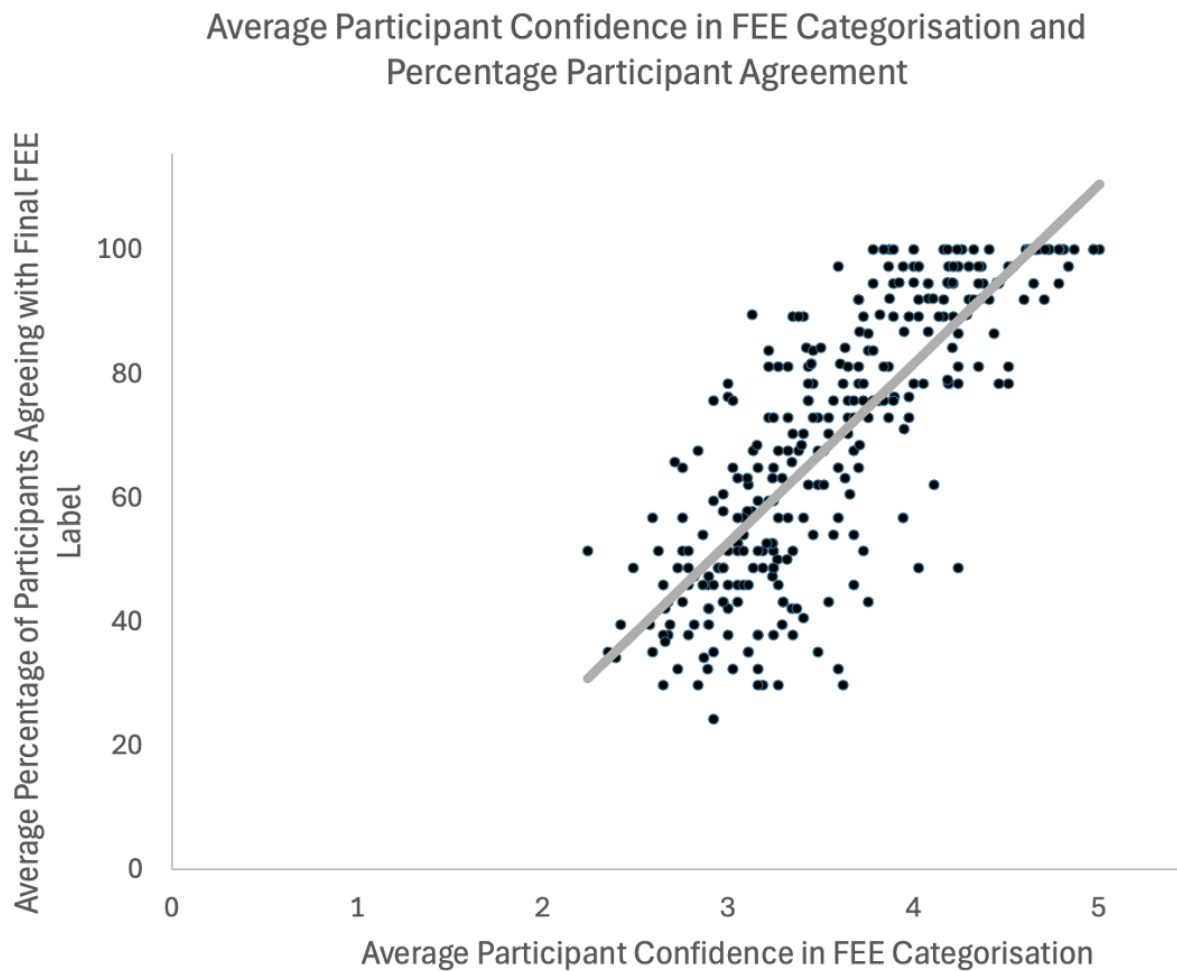
A scatterplot to demonstrate the relationship between the average confidence and intensity scores assigned to each stimulus.



For all correlations with intensity as a variable, the 56 stimuli receiving a final label of neutral were excluded. There was a significant strong positive correlation between the rated intensity of an expression and the confidence with which participants assigned it a label ($r = 0.806$, $p < 0.01$) (see figure 12).

Fig. 13

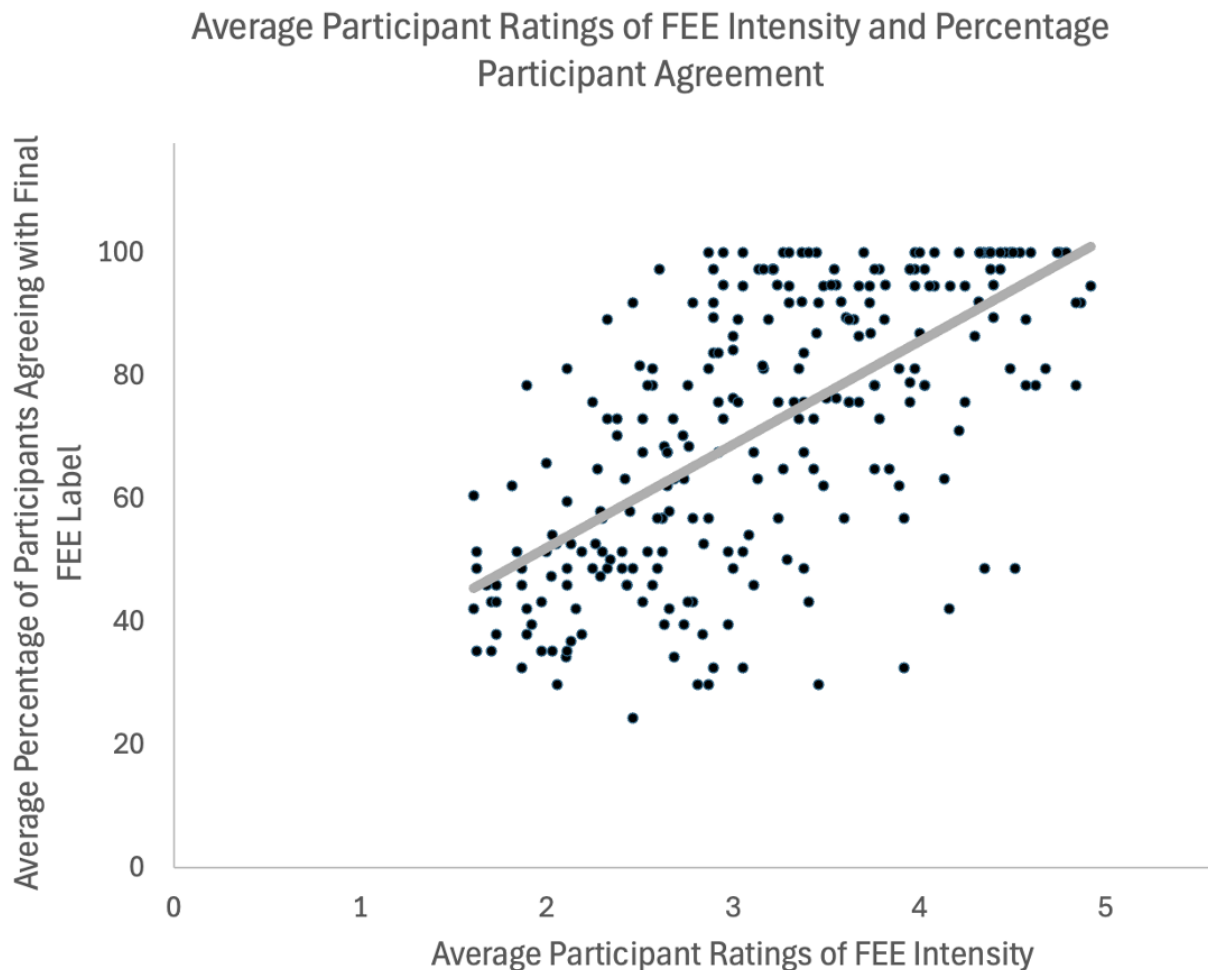
A scatterplot to demonstrate the relationship between the average confidence scores assigned to each stimulus and the percentage of participants agreeing with the final emotion label.



There was also a significant strong positive correlation between the average percentage participant agreement for a stimulus and the confidence with which participants assigned it a label ($r = 0.813, p < 0.01$). (see fig. 13).

Fig. 14

A scatterplot to demonstrate the relationship between the average intensity scores assigned to each stimulus and the percentage of participants agreeing with the final emotion label.



Finally, the relationship between average intensity score and percentage agreement for each stimulus was tested. There was a significant moderate positive correlation between the rated

intensity of an expression and the average percentage participant agreement in assigned it a label ($r = 0.653$, $p < 0.01$) (see figure 14).

Clip Length as a Potential Confound

To check for a confound of clip length for participant ratings, the correlation between clip length in frames, intensity ratings, confidence ratings and percentage agreement of a stimulus video was tested. The data were not normally distributed, so Spearman's correlations were conducted.

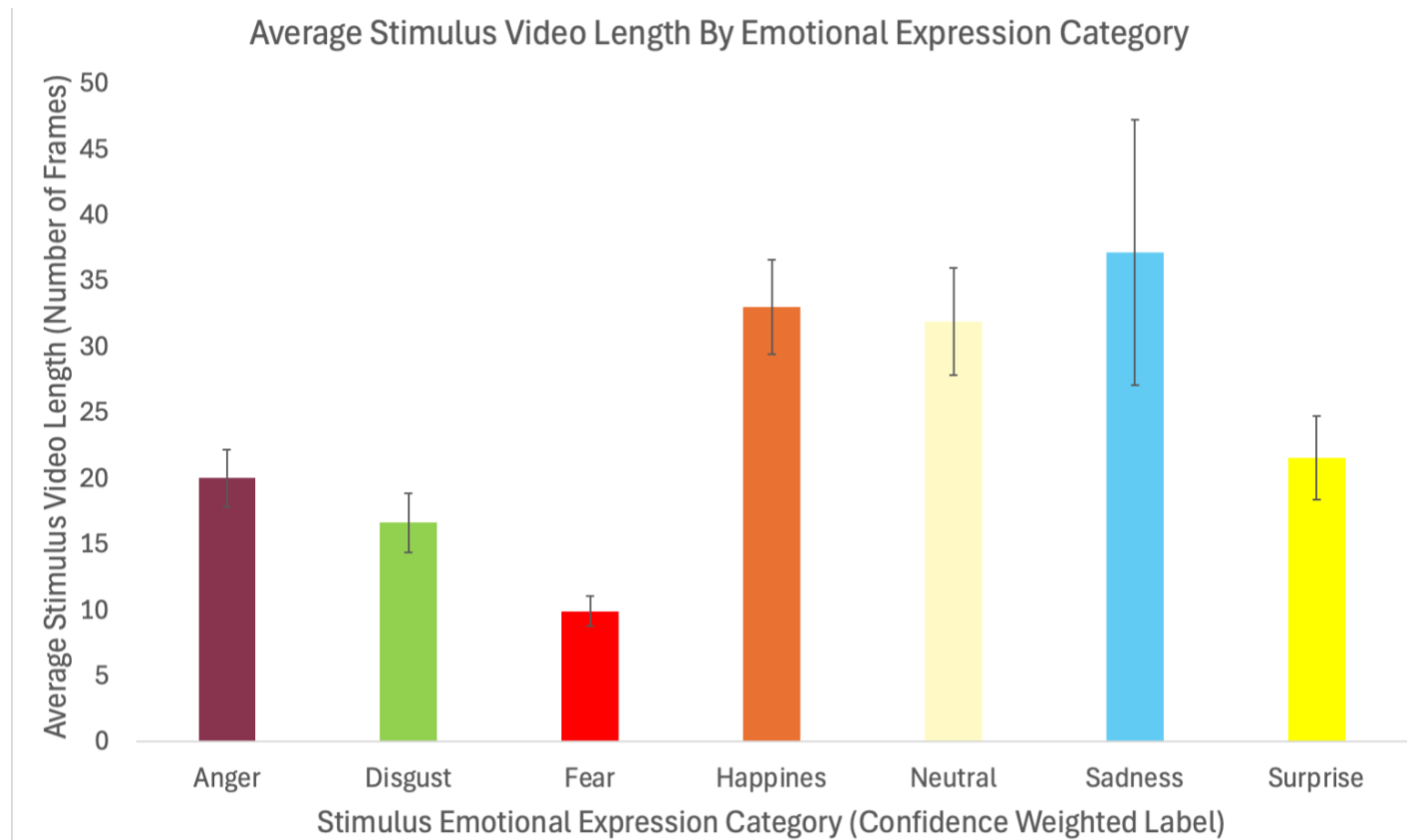
There was no significant correlation between clip length and intensity ratings ($r_s(309) = -.012$, $p = .829$). There was a moderate, positive correlation between clip length and agreement, $r_s = .23$, which was statistically significant, $t(309) = 4.10$, $p < .001$. There was a moderate, positive correlation between clip length and confidence, $r_s = .31$, which was statistically significant, $t(309) = 5.82$, $p < .001$. This could mean that participant agreement and confidence varied by stimulus length as a confound.

These relationships could potentially be explained at least in part by varying average clip lengths by final FEE label assigned to the video clip subject (see fig. 15 below). Mean clip lengths in frames varied across emotions, with the shortest average clip length in frames being fear $M = 9.92$, $SD = 3.96$, and the longest average clip length being for sadness ($M = 37.19$, $SD = 60.52$). A one-way ANOVA was conducted to examine whether there was a significant relationship between clip emotion category and clip length, $F(6, 304) = 3.21$, $p = .005$, $\eta^2 = .060$, indicating

that emotion category significantly affected clip length. Levene's test indicated a violation of the assumption of homogeneity of variances, $F(6, 304) = 6.71, p < .001$. As a result, the Games-Howell post hoc test was used to assess pairwise differences.

Fig. 15

A bar chart to demonstrate the average stimulus video length in frames by the final confidence-weighted FEE label assigned to the video clip.



Note. This bar chart represents the mean and standard error stimulus clip lengths in frames. The longest video clips were those that received a confidence-weighted label of sadness, and the shortest those that received the label of fear.

Post-hoc comparisons indicated that clips portraying happiness ($M = 33.04$, $SD = 30.57$) were significantly longer than those depicting anger ($M = 20.02$, $SD = 15.87$), $p = .045$, disgust ($M = 16.65$, $SD = 11.77$), $p = .011$, fear ($M = 9.62$, $SD = 3.96$), $p = .001$, and surprise ($M = 21.55$, $SD = 23.62$), $p = .046$. Additionally, clips portraying sadness ($M = 37.19$, $SD = 52.53$) were significantly longer than those portraying anger ($p = .031$), disgust ($p = .041$), and fear ($p = .009$). Clips portraying disgust were significantly longer than those portraying fear, $p = .037$. No other significant differences were found.

Discussion

In this paper, we sought to establish how typical adults categorise a sample of naturalistic, dynamic facial movements taken from naturalistic, mostly egocentric recordings. This is novel in that there is very limited data on adult perceptions of truly naturalistic facial emotional expressions. While the classic “basic” emotion view would suggest that the six “basic” emotions and neutral are unambiguous for typical adults, there is little research to substantiate this with real-world FEE.

We utilised the data on the ratings that 149 typical Western adults ascribed to 311 naturalistic, dynamic facial movements, with each participant rating around 78 individual stimuli and each stimulus receiving 37-38 ratings. We aimed to establish what proportion of the stimuli tested would receive each of the seven possible labels, the degree of disagreement or ambiguity that arose, and how this level of ambiguity may differ by the age group of the video clip subject and the emotional expression label valence category given to the videoclip. We were also interested

in participant ratings of FEE intensity and confidence in their own categorisation, as well as the role of clip subject age group.

Given that our stimuli were taken from real-world interactions, we could not guarantee that the “basic” emotions would occur in equal proportions within this stimulus corpus. There is very limited research on how often these FEE actually occur over the course of everyday life, though we attempted to select a varied sample of different facial movements. Researchers’ initial approximations of the expressions that we had represented in our stimulus set were largely consistent with participant categorisation. Overall, most of the stimuli in our set were categorised as containing happy FEE (23.47%), with fear being the least common (3.86%). Disgust was also relatively infrequent (9.00%), and neutral was the second most frequent stimulus category (18.01%). There was not equal representation of FEE by subject age group—for example, no infant subject stimuli were categorised as disgust, anger or fear by participants. This meant that statistical analyses of interactions between age group and FEE label were conducted on data collapsed across three valence categories: positive, neutral and negative.

We do find evidence that real-world examples of each of the “basic” FEE can be extracted; that is, anger, fear, surprise, neutral, happiness, sadness and disgust are emotion labels that can be clearly assigned by typical observers to at least some of the facial movements seen in the everyday visual environment when the task is forced choice. However, it is not clear whether we would replicate this with a more open response method such as that used by Cowen and Keltner (2020). Further, we do not find real-world examples of infant FEE being categorised as

fear, disgust or anger. While this could be a product of these emotions simply not occurring over the course of typical and functional family interactions, it could also be that infant FEE are not categorised in this way, either by our participants or adults as a whole. For example, there are rapid developmental changes in the production of FEE over infancy and early childhood (Field and Walden, 1982), so infants may not produce the same categories of facial movements. Further, parental status affects processing of and response to infant FEE (Coman, 2021), meaning that our primarily non-parent sample may not have been well-equipped for interpreting infant FEE.

Our hypothesis that there would be a relatively high degree of disagreement or ambiguity was supported. When averaged across clips, the average agreement with the consensus label was 69.41% (SD = 21.73), with a minimum of 24.32% and a maximum of 100%. This would suggest a relatively high level of ambiguity and disagreement, at least when categorising some facial movements. However, we did find a lower percentage of stimuli failing to exceed 50% consensus than Cowen and Keltner (2020), who found that 66% failed to meet this criterion as compared to our 25.7%. We would expect that, given our relatively homogenous sample (mostly young, female, childless Psychology students from Western countries) this is a relatively generous estimate of agreement. That is to say, if this research were replicated with a more representative and therefore varied sample, we might expect lower agreement (higher ambiguity).

Our hypothesis that agreement would vary by final FEE label was supported. However, the specifics of which emotions showed the highest agreement (and therefore lowest ambiguity) differed slightly from our hypothesis. While fear was the most ambiguous FEE, the least ambiguous was happiness, not anger. There were significant differences in participant disagreement by both stimulus subject age and FEE valence category (positive, neutral or negative). Stimulus subject age was associated with agreement both in isolation and in interaction with FEE valence category. For example, overall, agreement was highest for stimuli containing infants, followed by adults and then children—specifically, negative FEE valence category clips depicting infants had the highest agreement, and neutral FEE clips depicting adults had the lowest agreement. Possibly, the high agreement for infant negative stimuli was driven in part by the fact that the only negative FEE exemplars for this age group were labelled as containing sadness, with no examples of threat-related FEE identified for infants. However, overall, agreement was still higher for positive expressions. Whatever the reason, we find wide variations in agreement by the specific nature of the real-world FEE being assessed.

We therefore provide support for the idea that real-world, naturalistic emotional expressions are more ambiguous than theories taking the classical approach may predict, especially when asking participants to make judgements on short presentations of facial movements out of context. The degree of ambiguity in observer categorization varies by the category of facial movement (emotional expression). One possible explanation for the relative ambiguity of real-world FEE, particularly in the family home, could be display rules. We found the least ambiguity for negative FEE, but specifically in the infant age group. It could be that display rules increase

the ambiguity of negative FEE produced by adults, especially since this category included fear and anger for adults -- i.e., it is considered antisocial to produce the kinds of unambiguous, exaggerated presentations of these emotions that are typically used as prototypes in classical research (Matsumoto, 1990), but sadness is less constrained by these display rules, particularly in infancy.

In terms of the sources of ambiguity, when there are mixed responses to a facial movement, these mixed responses tend to follow certain patterns. Anger is most often alternatively assigned disgust; disgust neutral; fear sadness or surprise; happiness is mostly unanimous or alternatively assigned surprise; neutral is most often alternatively assigned surprise; sadness assigned anger; and surprise fear or neutral. These mixed responses could reflect one or both of two phenomena.

Firstly, these patterns of alternative response could reflect common errors or confusions, possibly based on proximity on theoretic constructs like the valence-arousal dimensions (Russell, 1989). Or it could be that these alternative responses reflect the mixtures of affective messaging that are most often actually present in facial movements. So, for example, it could be that mistaking angry faces for disgusted faces is a common error made because of similarities in the facial movements present (e.g., wrinkled nose); alternatively, it could be that subjects often display some elements of disgust when they exhibit facial movements consistent with anger. One way to examine this could be to perform analysis of the facial action units present in these naturalistic stimuli and examine how each facial action unit corresponds with the consensus participant response as well as “error” responses.

We confirmed our hypotheses that there would be a significant relationship between both FEE valence category and subject age group and participant confidence in their ratings. Negative infant expressions were associated with higher confidence, while neutral adult expressions were associated with lower confidence. For stimuli depicting adults and children, the highest confidence was for positive FEE. Infant sadness specifically was associated with higher confidence. We had anticipated that there would be lower confidence for fear and overall lower confidence for infant FEE categorisation, so specific predictions within these hypotheses were not substantiated. Average confidence was highest for infants, followed by adults, then children. However, it could also be that the absence of more “difficult” negative emotions in the sample of stimuli containing infants inflated performance for these clips, as our stimulus set did not contain infant anger or fear.

However, findings on agreement, confidence and clip FEE category were possibly confounded by stimulus clip length. Longer clips were associated with lower ambiguity and higher confidence; clips containing happiness, sadness and neutral also tended to be shorter than clips assigned to negative, threat-related FEE categories like fear and anger. It could be that greater ambiguity for these FEE arose from their more challenging, briefer presentations. However, it could be that this reflects how these FEE are exhibited in real-life interactions. There is indeed evidence to suggest that the real-world durations of FEE displays vary by emotion category, with fear being brief and sadness being more prolonged (Hoffman et al., 2010). So, rather than being a confound, this could be a feature of real-world expressions.

Overall, stimulus subject age group and final stimulus FEE label category were associated with participant intensity ratings, as anticipated. Broadly, these results followed similar patterns to confidence ratings, with intensity ratings being highest for negative infant FEE and lowest for negative adult FEE. So, this may indicate that display rules for different age groups play a part in perceived FEE intensity.

Finally, we confirmed our hypotheses that there would be significant correlations between intensity and confidence ratings, agreement and intensity ratings, and agreement and confidence ratings. This would suggest that actual FEE categorisation difficulty (as measured by the percentage of participants agreeing on a stimulus label) is significantly associated with both perceived ease of the task (confidence ratings) and the intensity with which an expression is displayed.

The fact that participant confidence and intensity ratings were related to stimulus subject age group and FEE label (to a greater degree for confidence than for intensity) could help us to explain some of the patterns seen in the ambiguity and disagreement for stimulus categorisation. Participants were most confident for infant negative FEE but were less confident in child and especially adult negative FEE; correspondingly, they rated disgust and fear as the least intense individual FEE. Given that on the basis of participant agreement, happy FEE were the least ambiguous and fearful FEE were the most ambiguous, but sadness in infants had very low disagreement rates, this would make sense: the lack of anger and fear and the presence of

highly unconstrained, intense sadness displays within the negative infant FEE category could explain higher agreement, intensity and confidence ratings for this category. This would also explain the greater ambiguity and lower intensity and confidence ratings for negative adult FEE as a category, which contains fear and anger likely to be subject to display rules and down moderated. Indeed, the correlations between intensity, confidence and agreement add further evidence to the idea that perceived FEE categorisation difficulty (confidence) is related in some way to intensity and FEE category. Participants disagree the most on FEE that they are not confident in rating and tend to show higher disagreement and lower confidence for less intense FEE.

Limitations and Future Directions

This research offers a highly novel view of adult perceptions of real-world FEE given that our stimuli were truly representative of the kinds of facial movements that can be seen in the family home environment. The data generated here were also invaluable in validating our computer vision models.

These instances of FEE were entirely naturalistic, possibly contributing to the fact that some types of FEE stimulus (particularly for the infant clip subjects with FEE such as disgust) were rare or absent. In future, it could be useful to generate some semi-naturalistic stimuli of this nature within a lab environment. One way in which to elicit a wider range of semi-naturalistic facial movements could be to conduct a range of affectively valenced activities in a lab environment (for example, tasks involving tasting of bitter substances). This could allow for collection of still

spontaneous but more varied facial movements, as it makes sense that certain FEE could be uncommon outside of certain, specific situations.

Conclusions

Our novel examination of adult perceptions of real-world FEE a demonstrates differences in ambiguity, perceptions of intensity and task “difficulty” (as measured by confidence ratings) by both final emotion label valence and subject age group; positive emotional expressions exhibited by adults and negative emotions exhibited by infants were associated with higher intensity and confidence ratings and less ambiguity. This contradicts traditional conceptions of FEE as unambiguous and processing of them as one unitary function.

Quantifying the Frequency of Emotional Expressions Within the Toddler Egocentric View

Emma J Jackson, William A.P. Smith, Elena Geangu

This paper has not yet been submitted for publication

Abstract

The development of facial emotional expression (FEE) processing is protracted over infancy and early childhood, and there is indirect evidence that the environmental input is important. However, no attempts have been made to quantify or characterise the nature of the visual environmental input at any stage of life. Toddlerhood is a crucial stage in socioemotional development, with increased independence allowing for potentially varied socioemotional input, and so characterising the degree to which toddlers see FEE could help us to understand how typical and atypical development of FEE takes place. Knowing how this compares to the visual input of adult caregivers in the same environment can help us to contextualise this in mature development, given that there is currently no direct research on the input at any age. We aimed to record the egocentric visual perspectives of toddlers and their caregivers in their home environments and quantify the degree to which these egocentric perspectives contain different FEE using computer vision. We hypothesised that while both positive and negative FEE would be present in the view of the toddlers, positive and neutral FEE would be more frequent. We also hypothesised that negative FEE would be more frequent in the caregiver view. Broadly, these hypotheses were confirmed, though the visual input of toddlers and their caregivers was markedly similar. Negative, threat-related FEE were especially infrequent in the egocentric perspectives of toddlers and their caregivers, with implications for theories of how processing of these FEE emerges. However, negative FEE were rare in both the toddler and caregiver perspectives. Future research should examine individual differences in FEE frequency in the visual input, as well as the degree to which younger infants see different FEE.

Keywords: naturalistic, emotion, toddler, egocentric, visual, computer vision

Quantifying the Frequency of Emotional Expressions Within the Toddler Egocentric View

The human face is an essential source of affective information, intentionally or unintentionally communicated, for typical adults (Jack and Schyns, 2017). Development of the ability to extract this information, especially information pertaining to affect and cognition from facial emotional expressions (FEE), is protracted; the emergence of new skills and increased proficiency with existing skills in this domain continues well into childhood and adolescence (Herba and Phillips, 2004; Leppänen and Nelson, 2006). Infancy and early childhood are periods critical for development of these skills, with rapid changes. The input of socioemotional stimuli early in development and the role that this input may play in development remains poorly understood, but the nature of the environmental input does seem to shape proficiencies in FEE processing (Safar et al., 2017). There are many potential ways that the known developmental trajectories for FEE processing could be related to or driven by the environment—for example, later emergence of processing of negative facial emotional expressions (Hoehl and Striano, 2010) has been hypothesised to be the result of negative emotional expressions being rare for younger infants and more common in the view of older infants and toddlers as they explore the environments around them (Leppänen and Nelson, 2012). However, no attempts have been made to quantify the input of emotional expressions that may contribute to the emergence and development of the ability to discriminate and extract meaningful information from faces.

The Importance of Emotional Expression Processing

Faces provide a rich avenue of information for typical human adults. With only information from the face, factors as general as the species of an individual through to factors as specific as the internal state of an individual can be extracted (Lange et al., 2022). Faces have been proposed to be a “special” class of stimulus for typical adult humans (Farah, 1996), with human faces being capable of producing a wide variety of movements that serve a crucial role in interpersonal interactions (Frith, 2009). Compared even to other primates and great apes, humans have faces adapted for communicative value, with facial features capable of emotional expressivity such as the eyes and lips more visually salient as a result of factors such as size, colour and contrast (Kret, 2015; Kret et al., 2020).

Emotional expressions, especially facial emotional expressions (FEE), have long been the subject of scientific interest, especially insofar as human FEE are “unique”—it was in 1872 that Darwin’s seminal work on human emotional expressions was first published (Darwin and Prodger, 1998), proposing an evolutionarily adaptive function for emotional expressions, especially for highly social animals like humans. While some have claimed FEE production and perception to be culturally universal, at least when it comes to some “basic” emotions (Ekman and Cordaro, 2011), there are many who fiercely contest this (Barrett et al., 2019) and there is direct evidence to speak against the universality of FEE (Jack et al., 2012). While there may be some elements that generalise cross-culturally, local variations in FEE are sufficiently different as to be conceptualised by many as different “dialects” of emotion (Elfenbein, 2013).

Facial emotional expressions have been demonstrated to serve many purposes in many contexts. Adults use emotional expressions such as pride and anger to infer social status (Shariff and Tracy, 2009; Tiedens, 2001). These expressions can be used by observers to infer behavioural intent, for example to predict approach or avoidance (Adams et al., 2006); they can also be used to shape social decision making, such as the decision to be lenient or stringent when administering social punishment (MacLin et al., 2009).

Deficits in processing emotional expressions, particularly facial emotional expressions, have been associated with both diagnostic status and impaired functioning in many clinical groups. Impaired FEE processing is common in autistic children and adults—though there have been suggestions that this could be better explained through alexithymia, which is also common in autistic individuals (Cook et al., 2013). For those diagnosed with alcoholism, reduced interpersonal functioning is specifically associated with poor FEE recognition (Kornreich et al., 2002). Further, in paranoid schizophrenia, not only is there a FEE processing deficit in those diagnosed, but in first degree relatives (Mendoza et al., 2011). While there is more limited research on the relationship between FEE processing ability and wider social functioning within the non-clinical population, there is some evidence to suggest that slower processing of FEE is associated with lower empathic skills even in neurotypical adults (Rigby et al., 2018). So, understanding factors driving FEE processing development is essential.

The Importance of Infancy and Early Childhood for Emotional Expression Processing

Cognitive development in infancy is rapid (Hunnius, 2022). This general principle extends to FEE processing development. While verbal assessment of perception and understanding is not possible in infancy, infants can be assessed using behavioural paradigms such as looking preference (Slater et al., 1985) and habituation (Cohen, 1976) as well as physiological measures such as EEG (De Haan, 2013). Specifically, for the latter, ERPs (Event-Related Potentials), which are elements of the EEG signal temporally associated with specific events (Hoehl and Wahl, 2012), are often used in research examining infant FEE processing. Data using these methodologies suggests rapid emergence of FEE processing skills.

There are very limited FEE processing skills present at birth. Within lab-based paradigms, neonates show a preference for looking at faces, or at least face-like objects (Mondloch et al., 1999); they are also able to spontaneously distinguish happy and fearful faces, though not fearful and neutral faces (Farroni et al., 2007). Neonates also exhibit a preference for happy facial expressions compared to neutral only when the happy face is accompanied by direct gaze (Rigato et al., 2011). While neonates cannot spontaneously discriminate between happiness and disgust, they can do so after habituation (Addabbo et al., 2018). So, skills are limited and restricted mostly to happy faces. However, it seems that at this early stage of development, even a limited amount of environmental input and minor changes to the specifics of a paradigm in the form of habituation can greatly influence the apparent FEE processing abilities.

This limited ability to extract affective information from faces at birth is consistent with the predictions of the CONSPEC/CONLERN model. The CONSPEC/CONLERN model, first proposed by Morton and Johnson (1991), suggests two mechanisms for face processing: CONSPEC, a subcortical mechanism that drives newborns' early preference for face-like stimuli, and CONLERN, a cortical system that develops later as infants gain experience with faces. While CONSPEC helps orient infants to faces from birth, providing an environmental visual substrate for learning, CONLERN is necessary for more complex processing. This includes the recognition of facial identity and, later, emotional expressions. This model would suggest that the ability to process FEE develops later, as cortical structures like the fusiform face area (FFA) and superior temporal sulcus (STS) mature using the environmental input provided by CONSPEC. The environmental visual input is therefore essential in understanding how FEE processing develops within the constraints of this model (Johnson et al., 2005; Slater, 2002).

By 1- to 3-months, infants demonstrate more complex FEE processing, but these abilities are largely restricted to positive FEE. There is further evidence for established processing of happy FEE, with 2-month-olds able to distinguish between happy and neutral expressions (Nelson and Horowitz, 1983). By 3 months, infants demonstrate within-category discrimination; for example, they prefer more intense smiling faces, and this preference is related to maternal sensitivity, possibly suggesting that environmental factors may influence their development (Kuchuk et al., 1986). Additionally, 3-month-olds exhibit enhanced face identity recognition for happy faces, possibly due to perceptually salient features of these expressions (Brenna et al., 2012; Turati et al., 2011). While their ability to recognise negative FEE like fear is still emerging, infants show

increased Nc (an ERP metric of attention) in response to objects paired with the gaze of fearful faces (Hoehl and Striano, 2010), although the exact mechanisms remain unclear, with alternate explanations such as a novelty response suggested (Hoehl et al., 2008). Despite this differentiation of the attention of fearful faces, EEG responses to fearful versus neutral faces themselves do not yet differ, possibly due to the involvement of subcortical structures in early face processing, as suggested by the CONSPEC/CONLERN model. These findings make clear both the rapid development of FEE processing skills and the earlier emergence of positive FEE processing.

Processing of negative FEE begins to emerge in the first half of the first year of life, but this is gradual. By 4-months-old, infants can visually discriminate between fear, anger and surprise (Serrano et al., 1992). Striano et al. (2006) also found enhanced Nc in response to angry faces with direct versus averted gaze at 4-months, demonstrating the beginnings of meaningful recognition of threat-related expressions, as anger with direct gaze would be a more salient threat. Development of the processing of negative and threat-related FEE continues to emerge slowly—by 9-months-old, in contrast to earlier ages, enhanced Nc response is observed for fearful relative to neutral faces regardless of the gaze direction of the face (Hoehl and Striano, 2010). This demonstrates that the ability to process negative FEE, particularly threat-related FEE, emerges much later than equivalent skills for happiness.

After infants are more able to interact with the environment independently, there is evidence that they can use FEE to navigate this environment. For example, Sorce et al. (1985)

demonstrated that 12-month-olds use maternal FEE to determine how to behave in an ambiguous situation (the visual cliff), with expressions of joy or interest leading most infants to cross the “cliff” and fear or anger leading few to do so. This modulation of behaviour based on observed positive or negative FEE was selectively observed in ambiguous situations. While there is evidence that vocal expressions are more influential than FEE in influencing crossing behaviour in this paradigm (Vaish and Striano, 2004), this provides evidence for meaningful understanding and use of both positive and negative FEE.

Older infants and toddlers also show an emerging capacity for making implicit assessments about the appropriateness of FEE in different contexts. By 18 months, infants can assess whether expressions are justified by environmental events, a skill not yet present at 15 months (Chiarella and Poulin-Dubois, 2013). At 20-months, toddlers expect actors to show surprise when a false belief is disproven, indicating an understanding that FEE can reflect internal cognitive states like beliefs (Scott, 2017). This marks a significant milestone, as toddlers begin clearly linking facial expressions to both external events and internal states. By 3.5 years, children show emerging categorical perceptions of emotions like happiness and sadness, a marker of mature processing, although their responses remain less consistent than those of adults (Cheal and Rutherford, 2011). By 3.5 years, children can explicitly and verbally judge emotions like happiness or sadness from faces; however, they often rely too heavily on facial cues alone and struggle to integrate contextual information (Hoffner and Badzinski, 1989).

The development of FEE processing extends beyond toddlerhood into childhood and adolescence, and changes continue to take place in the domains of speed, accuracy, and the nature of neural processing (Batty and Taylor, 2006; Herba and Phillips, 2004). While this brief outline is therefore not entirely comprehensive, it demonstrates just how many fundamental skills in FEE emerge at a rapid pace over the first three years of life. One potential driver of this would be the environment.

The Hypothesised Role of the Environment in Infancy and Early Childhood

There is much to indicate that this rapid emergence of skills depends to at least some extent on the environmental input. Many aspects of the environment are certainly essential for broad face processing development (Bayet, 2022). However, there is no direct quantification or examination of what may characterise the visual input of FEE at any age. It remains unclear how often or to what purpose typical adults actually use facial expressions, let alone how often they see them (Barrett et al., 2019).

There is evidence that at least some elements of emotional expression understanding are innate and independent of the environment. For example, congenitally blind children do still exhibit recognisable facial expressions, though these are not identical to those of sighted children, and negative expressions are less constrained by display rules (Galati et al., 2003). However, we can see associations between likely or definite differences in the environmental visual input of FEE and differences in development. Broadly speaking, atypical input seems to be

associated with atypical FEE processing, but even typical environmental variations are associated with wide differences in mature FEE processing ability.

We can see some clear evidence of disrupted socioemotional development associated with clearly inadequate environmental input with case studies of extreme neglect. Children exposed to neglectful, institutionalised care in orphanages correctly identify emotions in photos and stories significantly less frequently, with degree of impairment related to time spent in institutions (Sloutsky, 1997). Evidence of this nature may hint at a role for the environment in FEE processing emergence. However, there are many possible confounds, from preexisting genetic risk or illness through to differences in nutrition. Children placed in the institutions that are commonly examined often had pre-existing disabilities, and institutions such as these typically are holistically inadequate, with physical neglect also being common (Walker, 2011). So, there are many confounds that may explain later FEE processing differences or deficits in these groups outside of environmental differences.

There is also evidence for a potential role for the environment in atypical development when examining children of parents with psychiatric disorders. This is a very common form of potential environmental difference with around 13% of mothers experiencing postnatal depression (Leahy-Warren and McCarthy, 2007), and this is associated with atypical FEE processing in the infant. For example, infants of mothers exhibiting depressive symptoms during pregnancy and the early postpartum period show differences in emotional expression processing, with a bias in attention towards fearful faces (Kataja et al., 2018). This could be

because depressed parents provide atypical socioemotional input, as depressed individuals tend to exhibit reduced facial expressiveness (Renneberg et al., 2005). However, there is limited evidence on how psychiatric diagnoses affect the frequency of displaying specific FEE, and limited evidence on how these specific differences could impact on infant development, so explaining specifically the bias to fearful faces in infants of depressed mothers is difficult. It could also be that shared genetic risk for depression between infants and their mothers (Sullivan et al., 2000) drives this effect, as a negativity bias for negative emotions, including fear, is a well-replicated feature of this disorder (Bourke et al., 2010).

That said, there is also evidence that in more “typical” environments, with fewer potential confounds of this nature, selective deprivation of visual input may result in socioemotional stimulus processing deficits. To be specific, infants born with bilateral congenital cataracts, which preclude visual exposure to faces in the first weeks of life, later show impairments in face identity recognition when viewing angle or expression are not held constant (Geldart et al., 2002). They also demonstrate qualitatively different cognitive representations of FEE when their expression similarity judgments are analysed dimensionally. Cataract reversal patients’ FEE judgements more closely represent those of typical 7-year-olds, demonstrating pleasure and potency dimensions but lacking the arousal dimension of typical adults’ judgements; their judgments are also more highly variable than typical adults (Gao et al., 2013). Given that most cases of congenital cataracts are isolated and not associated with wider syndromic abnormalities (Bremond-Gignac et al., 2020), these differences are more likely to be the direct result of differences in the environmental input.

Finally, there is evidence that the nature of FEE processing in entirely typical individuals is shaped by the environmental visual input. This can be seen in the other-race-effect (ORE), which is the advantage for processing faces of races (or ethnicities) to which an individual is primarily exposed over infancy and childhood. Most research examining the ORE focuses on face identity processing. The ORE for face identity processing is not present at birth (Kelly et al., 2005) but is emergent by 3-months (Bar-Haim et al., 2006; Sangrigoli et al., 2004). There is much evidence that at least the face identity ORE is highly-experience dependent and particularly influenced by visual experience in infancy and childhood (Anzures et al., 2013). Crucially, the ORE extends to FEE. For example, adults are less accurate in judging at least some expressions when demonstrated by other race faces (Jiang et al., 2023), though evidence is mixed on the question of the specific expressions that this ORE is evident for (Elfenbein and Ambady, 2002). This would suggest that visual experience can shape some aspects of FEE processing proficiency.

There is some evidence that an ORE for emotional expressions can be traced back to infancy. For example, infants at 6-months-old show spontaneous visual preference for fearful faces only in own-race faces (which they are exposed to often), but not in other-race faces (which they are exposed to less often) (Safar et al., 2017). However, an ORE in FEE processing for the same emotion is not robust at this age; there is evidence that 7-month-olds effectively discriminate between same- and other-race happy and fearful expressions (Geangu et al., 2016). So, the existence of an FEE ORE that may begin its development in infancy highlights one

specific way in which the statistics of the environmental input may shape FEE development over infancy and early childhood even in typical individuals.

The Importance of Positive and Negative Emotional Expressions in the Second Year of Life

Not only is the environmental input and whether it is typical as a whole potentially significant for FEE processing development, but it could be that specific aspects of the statistics of the typical visual input are key. There appears to be a distinct difference in the timeline and social function of positive and negative emotional expressions. For example, as noted above, clear discrimination of some FEE (e.g., happiness) emerges much earlier than processing of other FEE (e.g., fear and anger).

To some extent, this distinction in processing between positive and negative, threat-related emotional expressions continues into adulthood. Typical adults are more sensitive to happy FEE than they are to other types of FEEs, detecting them at lower perceptual thresholds (Maher, Ekstrom and Chen, 2014). Though there is an attentional bias towards fearful and angry faces in childhood and adolescence, this is attenuated in typical adults (Kim et al., 2022). While display rules are somewhat culturally mediated (Safdar et al., 2009), negative FEE are more likely to be subject to these rules and therefore considered inappropriate in many circumstances (Matsumoto, 1990). So, the processes governing positive FEE and negative, threat-related FEE are dissociable.

Within the family environment, different FEE are likely to occur in different contexts and be used for different purposes. Mothers rarely direct negative emotional expressions at infants (Malatesta and Haviland, 1982). For this reason, negative FEE, especially those that are directed at younger children, may not be used frequently by adults outside of the context of transgressions (Essler and Paulus, 2021).

These are the kinds of contexts that may be likely to be present as an infant becomes a toddler and then a child. Confidently and independently mobile children are more likely to commit the types of transgressions that parents respond to with punishment (Lawrence et al., 2021). This could be related to the later emergence of negative FEE processing. These differing trajectories could reflect the changing statistics of the visual input over developmental time, with younger infants largely exposed to positive emotional expressions such as happiness, and with older infants who are more capable of committing transgressions exposed to negative FEE more frequently.

The ability to commit transgressions and the possible higher incidence of negative FEE within the context of this is only one reason that toddlerhood is a crucial point for socioemotional development and an important period to study. Understanding the visual input of emotional expressions for toddlers is especially critical due to their increasing independence and mobility, which introduces unique socioemotional learning challenges. Once toddlers become mobile, they are more likely to explore their environment autonomously, potentially encountering situations where they may explore, assess risk and transgress social or

behavioural boundaries (Tangen et al., 2022). Toddlers can seek out and independently shape their socioemotional visual input (Smith et al., 2018), so the input they receive may reflect in part their motivations and development. As a result, the emotional feedback they receive—primarily through FEE—becomes an essential source of information guiding their behaviour and development. So, understanding the statistics of the visual input can help us to understand the cues that toddlers use from others in this independent exploration.

Altogether, a comprehensive understanding of the frequency and nature of toddlers' FEE visual input, particularly in naturalistic settings, is vital for assessing how socioemotional development unfolds during this period. Especially, the frequency with which toddlers see negative versus positive FEE could help us to understand both the emergence of FEE processing and how toddlers use socioemotional information to navigate the environment.

Prior Research on the Nature of the Environment

While there is some research on the socioemotional input in terms of faces and bodies in infancy, early childhood and adulthood, there is no research objectively measuring and quantifying the naturalistic visual input with respect to facial emotional expressions at any age.

Prior research suggests that as a whole, the egocentric view in early infancy (1- to 3-months) is face-rich: around 23.2% (Jayaraman, Fausey and Smith, 2017) to 25% (Sugden et al., 2013) of scenes in the egocentric view in early infancy contain faces. This frequency declines to around 8.83% for infants 18-24 months, with the frequency of bodies remaining largely constant

throughout this period (Jayaraman, Fausey and Smith, 2017). The idea of infrequent faces in the toddler view is supported by chapter 1 of this thesis.

While there has been some research to establish the statistical characteristics of this face input—for example, Sugden and Moulson (2019) find that the most frequent faces in early infancy are those of the primary caregiver, followed by secondary caregivers—there has been no attempt to quantify the frequency with which infants, toddlers or children are exposed to or observe the various facial emotional expressions that may drive the development of the ability to extract meaningful information from this type of visual stimuli.

It is also notable that there is no prior quantification of the frequency of FEE in the view of adults. While prior research would suggest that faces are present in approximately 22.5% of the adult view (Oruc et al., 2019), and we replicate this relatively higher face visual input in our chapter 1, it is not clear how this will extend to FEE. It is possible that, at least within the home environment, adults may see negative FEE relatively frequently due to the presence of relatively unconstrained emotional displays from children within the environment.

Outstanding Questions and Hypotheses

This research paper aims to quantify the frequency with which the six primary FEE and neutral faces appear in the toddler egocentric view. This will help us to address the question of what

information exactly with respect to these FEE might be available for toddlers to learn from in the environment.

While there is prior quantification of the frequency of faces in the toddler view (chapter 1 of this thesis; Jayaraman, Fausey, and Smith, 2017), this is the first direct examination of the naturalistic input of facial emotional expressions. Given the variation in face frequency in the egocentric view with age found in prior research, it is highly likely that these results apply only to this specific age. However, as discussed, it is especially useful for us to know what the input is within the second year of life as this is a theoretically significant age wherein newly independent children may be capable of committing transgressions and eliciting social responses from those around them.

Firstly, it is important in terms of understanding to what degree environmental exposure to negative FEE may form the developmental substrate for development of FEE processing. Not only are toddlers at this age capable of seeking out and shaping their own visual input, but they are more capable of exploring the environment independently and therefore are able to commit transgressions. So, if infants, children and toddlers are exposed to negative FEE at all, we could expect them to be present in the visual input of toddlers.

Secondly, many important and complex FEE processing skills are still emergent at this age— for example, processing FEE and context together (Hoffner and Badzinski, 1989). This means that if the environment is important for FEE processing development, there is the potential for it to still be shaping the developmental changes that occur in the second year of life.

Here, we also characterise the frequency of FEE in the view of the adult caregivers of toddlers.

This gives us a point of comparison for the toddler FEE frequency, and is in and of itself valuable, given that there are no characterizations of the frequency of FEE in the adult view.

Understanding the visual input of an adult within the same home environment will help to give a reference point for the frequencies of these FEE within the toddler view.

Given that negative FEE are more often socially desirable and subject to display rules, we expect them to be relatively infrequent in the egocentric views of the toddlers, given that their interaction partners will often be the better-regulated adults within the home. However, we do not expect the converse to be true for caregivers, as the FEE of less emotionally regulated and restrained toddlers should make up a significant portion of their visual input.

We hypothesise that: 1) both positive and negative FEE will be present in the egocentric view of the toddler sample that we will examine; 2) positive FEE will be more frequent in the egocentric view relative to the frequency of negative FEE; 3) the most common FEE in the toddler egocentric view will not be negative, they will be positive or neutral; 4) the least common FEE in the toddler view will be anger, fear and disgust; 5) Negative, threat-related FEE (anger, fear and disgust) will be more common in the caregiver view than the toddler view.

Methods

See general methods section for full details.

Data Analysis Strategy

To maximise the sample size and statistical power here, we first consider $N = 26$ participants who provide data from the toddler perspective, and then the $N = 25$ participants who provide footage from the caregiver perspective and then the $N = 24$ participants with data from both perspectives. Each frame was coded for the presence of faces and facial emotional expressions using the modified Retinaface/Arcface combination model as outlined in the general methods section. The presence of different emotional expressions as a proportion of just faces detected was examined for the toddler and caregiver perspectives separately, before the presence of these expressions as a proportion of total recorded frames was compared for these two perspectives.

Given prior research, including our own, which indicates significant differences in face frequency between toddler and caregiver views (7.67% (SD = 4.49) and 13.81% (SD = 6.92) respectively in this sample, see chapter 1), we did not consider that meaningful comparison could be conducted without taking face frequency into account. For toddler-caregiver comparisons, we therefore compared the overall frequency of FEE as a percentage of frames, not faces. This comparison, focused on this subset of participants with both perspectives, allows us to analyse real differences in FEE visual input accounting for differences in input that result from quite disparate face viewing frequencies. This is because similar percentages of FEE relative to faces could reflect substantial differences in visual FEE exposure when viewed as a percentage of total frames. Thus, analysing FEE as a percentage of frames offers a more accurate understanding of real-world visual input.

Results

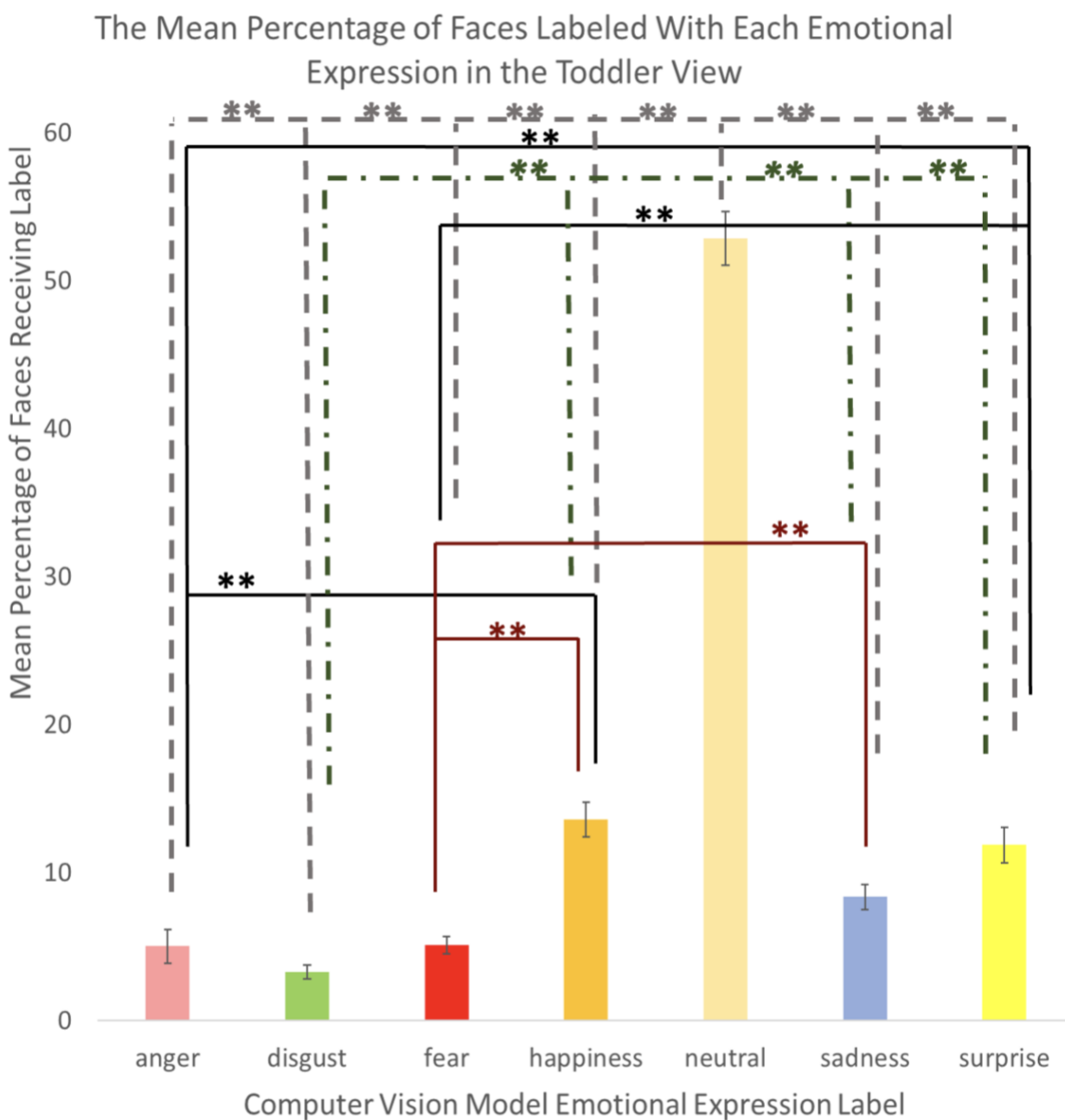
Presence of Emotional Expressions in the Toddler View

First, we examined the toddler visual input focusing on the 26 families that successfully recorded footage from at least the toddler perspective. When examining the percentage of frames containing a face receiving each FEE label from our computer vision models, the most common label for face frames was neutral ($M = 52.84\%$, $SD = 9.34\%$). This was followed by happiness ($M = 13.58\%$, $SD = 5.94\%$); then surprise ($M = 11.86\%$, $SD = 6.14\%$); then sadness ($M = 8.34\%$, $SD = 4.32\%$); then fear ($M = 5.10\%$, $SD = 3.02\%$); with anger ($M = 5.01\%$, $SD = 5.84\%$) and disgust ($M = 3.28\%$, $SD = 2.50\%$) being the least common (see figure 1).

A one-way repeated measures ANOVA was conducted to compare the frequency of the different FEE in the toddler view. Mauchly's test of sphericity indicated that the assumption of sphericity had been violated, $\chi^2(20) = 68.759$ $p < .001$, and so a Greenhouse-Geisser correction was used ($\epsilon = .535$). There was a statistically significant difference in the percentage of face frames between the different FEE, $F(3.32, 80.32) = 207.22$, $p < .001$. Post-hoc paired t-tests using a Bonferroni corrected $\alpha = .0024$ were conducted, and pairs of FEE frequencies that are significantly different are marked on figure 1.

Fig. 1

A bar chart demonstrating the mean and standard error percentage of faces in the toddler view assigned each FEE label by our Retinaface/Arcface model.



Note. Significance lines represent paired t-tests at the corrected $\alpha = .0024$. Dashed grey lines indicate differences between neutral and other FEE frequencies; dash-dot green lines indicate

differences for disgust; solid red lines for fear; and solid black lines for all other significant differences.

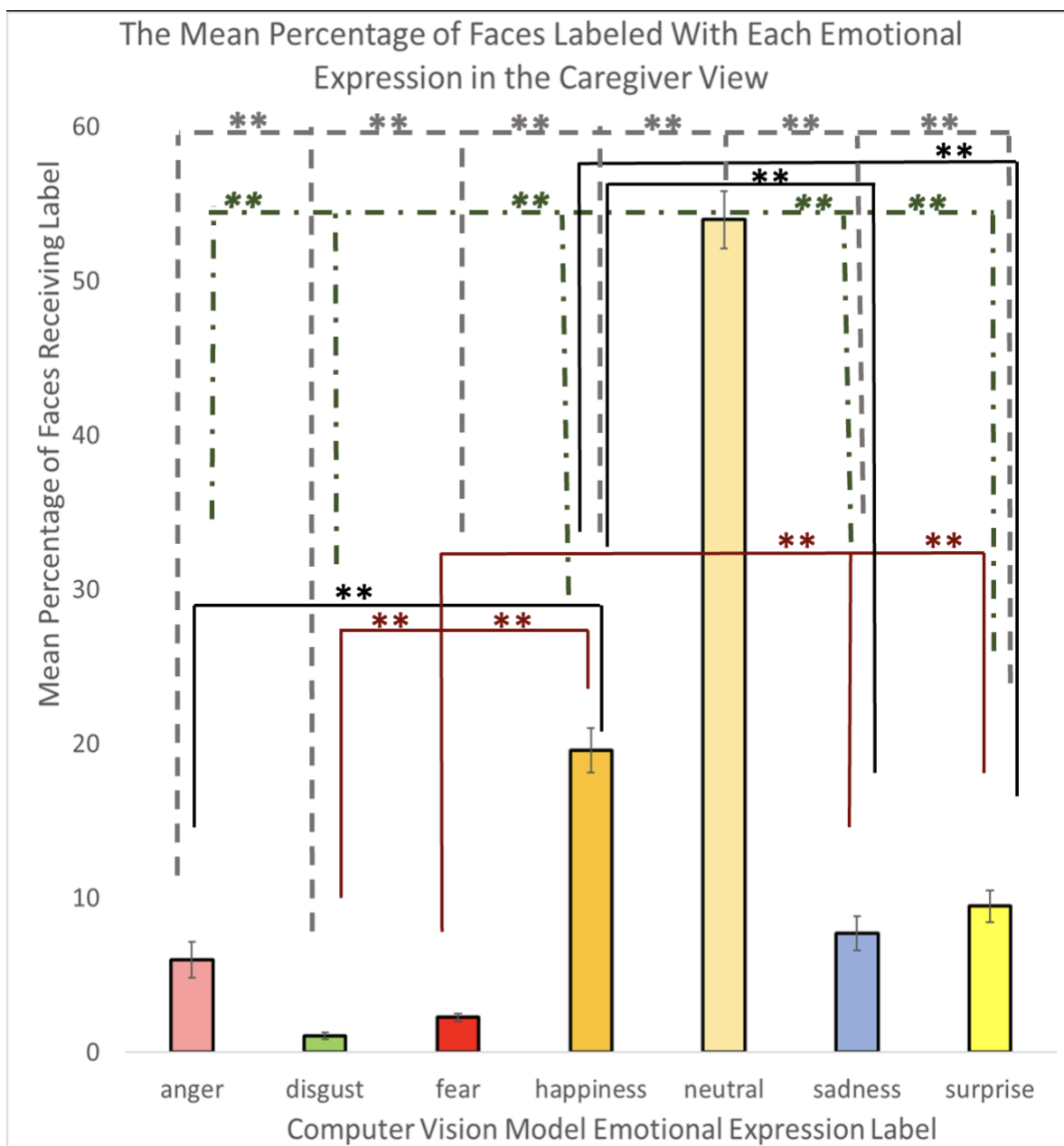
Presence of Emotional Expressions in the Caregiver View

Next, we examined the adult visual input focusing on the 25 families that successfully recorded footage from at least the caregiver perspective. For these 25 families, the most common label for face frames in the adult view was neutral ($M = 53.97\%$, $SD = 9.31\%$). This was followed by happiness ($M = 19.58\%$, $SD = 7.20\%$); then surprise ($M = 9.47\%$, $SD = 5.17\%$); then sadness ($M = 7.70\%$, $SD = 5.55\%$); then anger ($M = 5.98\%$, $SD = 5.83\%$); with fear ($M = 2.26\%$, $SD = 1.32\%$) and disgust ($M = 1.04\%$, $SD = 1.07\%$) being the least common (see figure 2).

A one-way repeated measures ANOVA was conducted to compare the frequency of the different FEE in the caregiver view. Mauchly's test of sphericity indicated that the assumption of sphericity had been violated, $\chi^2(20) = 136.420$ $p < .001$, and so a Greenhouse-Geisser correction was used ($\epsilon = .511$). There was a statistically significant difference in the percentage of face frames between the different FEE, $F(3.07, 73.56) = 220.97$, $p < .001$. Post-hoc paired t-tests using a Bonferroni corrected $\alpha = .0024$ were conducted, and pairs of FEE frequencies that are significantly different are marked on figure 2.

Fig. 2

A bar chart demonstrating the mean and standard error percentage of faces in the caregiver view assigned each FEE label by our Retinaface/Arcface model.



Note. Significance lines represent paired t-tests at the corrected $\alpha = .0024$. Dashed grey lines indicate differences between neutral and other FEE frequencies; dash-dot green lines indicate

differences for disgust; solid red lines for fear; and solid black lines for all other significant differences.

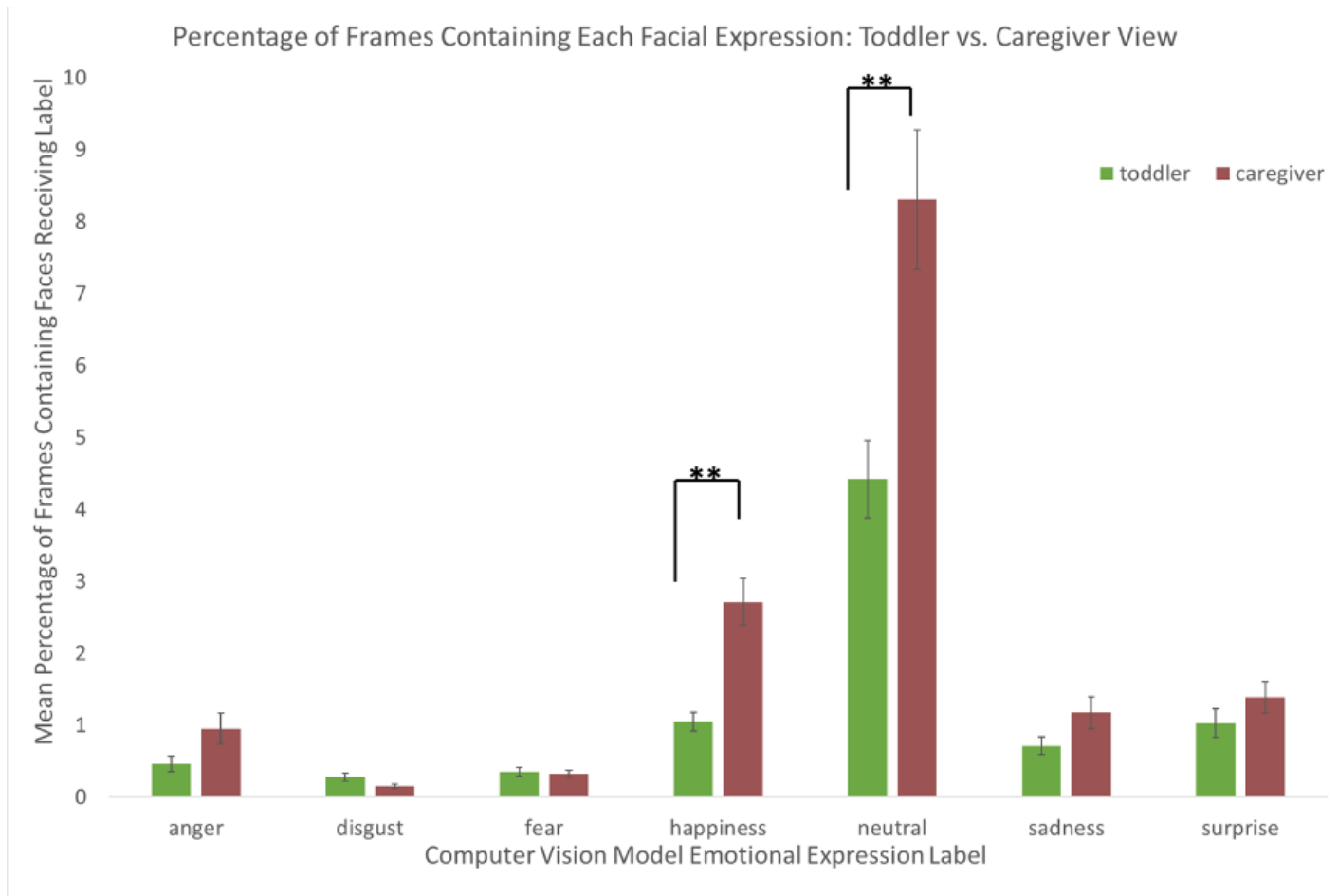
Comparing the Toddler and Caregiver View in the Sample Subset with Both Perspectives

We then sought to compare these frequencies directly in the subsample of 24 families who recorded both perspectives. We examined FEE as a percentage of frames rather than faces. First, we calculated the mean and standard deviation percentages of frames containing each FEE within this subsample. Figure 3 illustrates these figures for the 24 families who recorded both perspectives.

A two-way repeated measures ANOVA was conducted to compare the frequency of the different FEE in the different perspectives (toddler versus caregiver). Mauchly's test of sphericity indicated that the assumption of sphericity had been violated, $\chi^2(20) = 155.410$, $p < .001$, and so a Greenhouse-Geisser correlation was used ($\epsilon = .271$). There was a statistically significant interaction between FEE label and perspective $F(1.626, 37.404) = 18.227$, $p < .001$.

Fig. 3

A bar chart demonstrating the mean and standard error percentage of frames in the toddler and caregiver views assigned each FEE label by our Retinaface/Arcface model.



Note. A bar chart to illustrate the percentage of frames in the toddler and caregiver views (green and brown respectively) containing faces assigned each FEE as averaged across participants; this data is taken from only participant families that provided both toddler and caregiver perspective footage. Significant differences at $<.001$ are indicated with a **.

Table 1

A table to demonstrate results pairwise comparisons of the percentage of frames containing different emotional expressions in the toddler view.

<i>Emotion</i>	Mean difference (Standard Error) Percentage of Frames						
	Anger	Disgust	Fear	Happy	Neutral	Sad	Surprise
Anger	-	-	-	-	-	-	-
Disgust	.18 (.11)	-	-	-	-	-	-
Fear	.11 (.12)	-.072 (0.65)	-	-	-	-	-
Happy	-.59 (.15)*	-.77 (.11)**	-.70 (.11)**	-	-	-	-
Neutral	-3.96 (.49)**	-4.14 (.52)**	-4.07 (.51)**	-3.37 (.44)**	-	-	-
Sad	-.25 (.11)	-.44 (.13)	-.36 (.10)*	.34 (.12)	3.71 (.45)**	-	-
Surprise	-.56 (.21)	-.75 (.18)*	-.67 (.18)*	.025 (.15)	3.40 (.43)**	-.31 (.19)	-

Note. This table outlines pairwise comparisons of the percentage of frames in the toddler view labelled as containing each FEE by our computer vision model. Differences are relative to the column header. * indicates significance at the $p < .05$ level, while ** indicates significance at the $p < .001$ level.

Post-hoc ANOVAs were then conducted for each perspective to confirm the main effect of the emotional expression category within this sample. First, a repeated measures ANOVA was

conducted to examine differences in the frequency of emotional expressions (anger, disgust, fear, happy, neutral, sad, and surprise) from the toddler's perspective. Mauchly's test of sphericity indicated that the assumption of sphericity was not met, $\chi^2(20) = 194.21, p < .001$. Thus, the Greenhouse-Geisser correction was applied ($\epsilon = .298$). The analysis revealed a significant main effect of emotion on the frequency of expressions, $F(1.43, 32.86) = 54.21, p < .001$, with a large effect size ($\eta^2 = .776$). See table 1 for post-hoc pairwise comparisons using Bonferroni adjustment.

Next, a repeated measures ANOVA was conducted to examine differences in the frequency of emotional expressions (anger, disgust, fear, happy, neutral, sad, and surprise) from the caregiver's perspective. Mauchly's test of sphericity indicated that the assumption of sphericity was not met, $\chi^2(20) = 210.35, p < .001$. Thus, the Greenhouse-Geisser correction was applied ($\epsilon = .238$). The analysis revealed a significant main effect of emotion on the frequency of expressions, $F(1.43, 32.90) = 57.89, p < .001$, with a large effect size ($\eta^2 = .809$). See table 2 for post-hoc pairwise comparisons using Bonferroni adjustment.

Table 2

A table to demonstrate results pairwise comparisons of the percentage of frames containing different emotional expressions in the caregiver view.

<i>Emotion</i>	Mean difference (Standard Error) Percentage of Frames						
	Anger	Disgust	Fear	Happy	Neutral	Sad	Surprise
Anger	-	-	-	-	-	-	-
Disgust	.80 (.20)*	-	-	-	-	-	-
Fear	.63 (.21)	-.17 (0.05)*	-	-	-	-	-
Happy	-1.76 (.38)*	-2.56 (.32)**	-2.39 (.31)**	-	-	-	-
Neutral	-7.35 (.94)**	-8.15 (.96)**	-7.98 (.94)**	-5.60 (.78)**	-	-	-
Sad	-.22 (.29)	-1.02 (.22)*	-.85 (.21)*	1.54 (.37)*	7.13 (.88)**	-	-
Surprise	-.43 (.28)	-1.23 (.21)**	-1.06 (.21)**	1.33 (.24)**	6.92 (.86)**	-.21 (.29)	-

Note. This table outlines pairwise comparisons of the percentage of frames in the caregiver view labelled as containing each FEE by our computer vision model. Differences are relative to the column header. * indicates significance at the $p < .05$ level, while ** indicates significance at the $p < .001$ level.

Paired-samples t-tests were conducted to compare the frequency of each facial emotional expression (FEE) between the toddler and caregiver perspectives. To account for multiple comparisons, a Bonferroni correction was applied, adjusting the significance threshold

to $\alpha = 0.007$ ($0.05/7$). After this adjustment, significant differences were found for happiness, with more frequent occurrences in the caregiver view ($M = 2.710$, $SD = 1.591$) compared to the toddler view ($M = 1.048$, $SD = 0.628$), $t(23) = -5.855$, $p < .001$. There was also a significant difference in the frequency of neutral faces, which were more common in the caregiver view ($M = 8.304$, $SD = 4.734$) than in the toddler view ($M = 4.419$, $SD = 2.613$), $t(23) = -4.658$, $p < .001$. Other comparisons were not significant. See table 3 for a full summary of post-hoc t-tests; significant differences are also marked on figure 3 above.

Table 3

A table to demonstrate results of paired-samples t-tests comparing the percentage of frames receiving each FEE label in toddler and caregiver views for each family.

FEE Label	Toddler View M% (SD)	Caregiver View M% (SD)	Statistic t-value	p-value	Significant?
Anger	0.46 (0.54)	0.95 (1.07)	-2.00	.057	No
Disgust	0.28 (0.28)	0.15 (0.17)	1.94	.033	No
Fear	0.35 (0.30)	0.32 (0.23)	0.36	.362	No
Sadness	0.71 (0.61)	1.17 (1.08)	-2.14	.022	No
Surprise	1.02 (0.98)	1.38 (1.08)	-1.93	.033	No
Happiness	1.05 (0.63)	2.71 (1.59)	-5.86	< .001	Yes
Neutral	4.42 (2.61)	8.30 (4.73)	-4.66	< .001	Yes

Note. This table outlines overall confidence ratings by the different categories of stimuli. Figures are presented for each FEE label category and each video subject age group. The adjusted significance threshold is $\alpha = 0.007$.

Discussion

The primary aim of this study was to quantify the degree to which the six “basic” facial FEEs—happiness, sadness, anger, fear, disgust, and surprise— and neutral are present within the toddler and caregiver input. We examined the visual input of toddlers within the second year of life and their caregivers. We analysed egocentric view videos coded for the presence of FEE using a computer vision model. The degree to which humans of any age see FEE has not been quantified before. By also examining FEE visual input in caregivers, we intended to provide a comparison within the same environment.

We first hypothesised that both positive and negative FEEs would be present in toddlers' visual input, as toddlers (unlike younger infants) are newly able to independently explore, meaning they can commit transgressions. Caregivers may then selectively employ negative FEE in response to these transgressions (Dahl et al., 2014). We then hypothesised that positive and neutral FEEs would be more frequent than negative FEEs in the toddler view. Finally, we hypothesised that negative FEEs—particularly anger, fear, and disgust—would be less frequent in the toddler view than the caregiver view, as caregivers are more able to moderate their emotional responses than toddlers. This may then mean that caregivers encounter more negative FEEs than toddlers, as a significant portion of their visual input would be toddlers who lack mature emotional regulation skills. These hypotheses would be relevant in light of the relatively late emergence of the same FEE processing skills for negative as compared to positive FEE— as outlined, from birth infants distinguish happy faces (Farroni et al., 2007), but equivalent

skills for fearful and angry faces do not begin to emerge until 4-months (Serrano et al., 1992). Some have theorised that this could be because of relatively sparse input of negative FEE, especially early on in development.

Findings and Analysis

First, we examined the 26 participants who recorded at least the toddler perspective. As anticipated, while both positive and negative FEE were present in the egocentric view of the toddler sample, positive FEE were more frequent in the egocentric view than negative FEE, with the most common FEE in the toddler egocentric view being neutral, followed by happiness. Together happy and neutral faces made up approximately 66% of the faces in the toddler view. The least common FEE in the toddler view were anger, fear and disgust, which are negative, threat-related FEE. Together, these made up only approximately 13% of faces in the toddler view. This confirmed both hypothesis 1 and hypothesis 2.

We then examined 25 participants who recorded at least the caregiver perspective. We calculated the overall percentage of faces in the caregiver view that received each FEE label. The pattern of results was markedly similar, with approximately 73% of faces on average being happy or neutral, and anger, fear and disgust only making up approximately 9% of the faces overall.

Finally, we directly compared the frequency of each FEE in the subset of 24 participants who recorded both the toddler and caregiver views. Given the marked and potentially

developmentally meaningful differences in the frequency with which toddler and caregivers in the same environment contain faces (approximately 7.98% of toddler-perspective frames and 13.98% of caregiver perspective frames contain faces, see chapter 1), we sought to compare the actual frequency of these FEE not standardised for face frequency. To this end, we calculated and compared the percentage of frames in the corresponding toddler and caregiver views overall that contained a face receiving each FEE label. While happy and neutral faces overall were present in a higher percentage of caregiver-view frames than toddler-view frames, there was no significant difference between the frequencies of any other FEE in the toddler and caregiver views. The findings on happiness and neutral faces could have been driven by the overall frequency of faces.

The Infrequency of Negative FEE

The marked infrequency of negative FEE overall, especially those that are threat-related, and even within the caregiver view, is interesting for many reasons. Firstly, we expected the caregiver view to contain much more varied and often negatively valenced facial expressions as a result of the presence of the toddler within the home. This may not have been the case for several reasons. Firstly, it could be that toddlers use a more limited range of negative FEE, such as sadness— which made up approximately 8% of faces in the caregiver view but was not significantly more frequent overtime in the toddler view— and actually use overt negative FEE more infrequently than we might have predicted. Toddlers do engage in some emotional regulation strategies (Grolnick et al., 1996) and may have been relatively relaxed and calm within the home environment.

The infrequency of negative FEE in the toddler view is also interesting. Theoretically, toddlers (because of their ability to commit transgressions) would be more likely than infants at an earlier time point to see these negative FEE—and so, there is no reason to believe or expect that younger infants would see these facial poses any more frequently. This poses the question—what do infants and toddler learn from these very rare environmental instances of fear, anger and disgust? And to what degree is development of skills for processing emotions driven in part or whole by maturational processes? Given that the ability to recognise negative, threat-related FEE starts to emerge at 4-months (Serrano et al., 1992) with developments throughout the first year of life (Hoehl and Striano, 2010) and into toddlerhood (Xie et al., 2021), some hypothesised that this later emergence was driven in part by increasing frequency of negative FEE in view as infants grow older and become toddlers. However, it remains unclear just how such infrequent presentations in toddlerhood (which likely reflect even less frequent negative FEE earlier in life) drive this development, if at all. The infrequency of negative FEE even within the view of toddlers may explain why processing of negative FEE remains more challenging even into toddlerhood (Bayet et al., 2018).

Individual Variation and Future Research

There was notable individual variation and relatively wide standard deviations in the frequencies of FEE we examined here. If the environmental visual input does indeed drive in some way socioemotional development, including development of the ability to process social signals such as FEE, this variation may be significant. Deficits and differences in processing and

understanding different FEE have been implicated in many developmental disorders (Wingenbach et al., 2017; Fairchild et al., 2009). These deficits have been particularly widely evidenced in ASD and conduct disorder, being evident from early childhood (Wong et al., 2012; Martin-Key et al., 2018).

These deficits seem to especially impact processing of negative and threat-related FEE. Specifically, there is evidence that individuals with ASD show atypical processing of fear as early as preschool (Dawson et al., 2004). There is also evidence for deficits in anger, sadness and disgust processing in adults with ASD, while happy processing is relatively intact (Farran et al., 2011; Wallace et al., 2008). Conduct disorder has also been associated with specific fear processing deficits in preschool (White et al., 2016). It could be that variation in the statistics of the visual socioemotional input, especially for the very infrequent negative and threat-related FEE, is a driver of differences in development. Alternatively, it could also be that learning from infrequent presentations of negative FEE is more challenging for those already predisposed to developing these disorders. Future research should attempt to extract what factors may be related to or underpin individual differences in this input. Understanding how the visual input with respect to FEE might vary with other individual differences (for example, temperament, genetic risk status) could help us to build an understanding of how individual differences in FEE, as well as pathologies, emerge over developmental time.

It would also be important to establish how the visual statistics of FEE differ for younger infants, as well as perhaps for adults in different environments. When comparing FEE as a percentage of

face presentations, the toddler and caregiver views seem similar. This stands in contrast with simple face frequency in the view of toddlers and their caregivers, with toddlers seeing markedly fewer faces than either their caregivers or younger children. It could be that this is specific to this comparison-- i.e., toddlers have reached a developmental stage wherein their visual input with respect to FEE as a proportion of faces is very similar to that of adults in the same environment, but other age groups would be very different—or, it could be that individuals in the same environment see very similar frequencies of FEE regardless of age, and that this is something that doesn't change much with developmental time. Given how rare negative FEE are for toddlers, and how caregivers are more likely to moderate their expressions when directed at very young infants, we might expect younger infants to almost never see fear, anger and disgust. Knowing how the frequency of FEE compares for toddlers and their adult caregivers gives us some context for understanding what these frequencies mean, but to understand the full picture, we would need to know the nature of this visual input for young infants and perhaps for adults in different environments.

Methodology and Limitations

As with all naturalistic developmental research, pragmatic constraints such as the high participant demands and the time-consuming data coding process limited the number of participants were able to recruit and, consequently, the statistical power here. While the toddler-perspective (N = 26) and caregiver perspective (N = 25) data here may represent hundreds of hours of footage, the number of individuals is limited.

However, this sample size does represent one of the most in-depth examinations of the visual input at any one developmental age. Not only is the quantification of the FEE in the naturalistic view at any age novel, but a specific exploration of the visual input of toddlers is also novel— all prior research together that includes the toddler view represents only a fraction of this sample size. Further, given the slightly increased statistical power conferred by our repeated-measures design, this data can still give us a good first understanding of the degree to which toddlers and their caregivers see different FEE.

The sample size and wide corpus of videos analysed was only possible through the novel computer vision coding methods used here. Clearly, given the rich data often employed in the study of natural video data, manual coding is typically inefficient and often inaccurate. We demonstrate the utility of novel technologies, including computer vision models, in allowing greater quantities of rich visual data to be collected, coded and analysed for many variables, including facial emotional expressions.

Conclusions

In conclusion, this is an exciting first look at the visual worlds of toddlers with respect to the facial emotional expressions that they see. Unsurprisingly, most faces in the toddler view are neutral or happy, and negative emotional expressions, especially those that are threat-related, are rare. This is not too dissimilar to the visual input of their caregivers, though negative, threat-related FEE may be slightly more common in the adult view. This infrequency, and the likelihood of younger infants seeing negative FEE even less, could go some way to

explaining the special challenge that negative FEE processing poses, especially for younger children and clinical populations. However, it does leave how exactly toddlers and younger infants learn from such infrequent exemplars unclear. It could be that the salience of infrequent negative FEE makes them especially informative, or that maturational processes play a great role in the development of FEE processing, especially negative FEE. Going forwards, it will therefore be important to understand whether this pattern of FEE frequency holds for younger infants and for adults in different environments, as well as whether individual differences in FEE visual input may be significant in shaping developmental trajectories. Replication with a wider sample may also be possible in future, in part because of the automated coding we have used in this research.

Factors Affecting the Frequency of Emotional Expressions Within the Toddler Egocentric View

Emma J Jackson, William A.P. Smith, Elena Geangu

This paper has not yet been submitted for publication

Abstract

The development of facial emotional expression (FEE) processing seems to depend at least to some degree on the early environmental input. The degree to which infants and toddlers see different categories of face appears to affect proficiency at extracting information (including FEE information) from each type of face. Further, there is diversity in mature ability to extract FEE from faces, as well as diversity in the visual input received by different individuals. However, no attempts have been made to quantify or characterise the degree to which visual FEE input may vary between individuals, or which factors may shape this variation. This may be especially notable in toddlerhood, at which point individuals are more able to shape their visual input through their newly developed motor independence and social processing development is rapid. We recorded the egocentric visual perspectives of toddlers in their home environments and coded the degree to which these egocentric perspectives contained different FEE using an emotional expression detection algorithm. We hypothesised that the frequency of FEE would be related to both external factors (the presence of siblings) and internal factors (toddler temperament). There was a relationship between sibling status and FEE frequency, with toddlers without siblings seeing more happy faces. Further, negativity, specifically falling reactivity, was longitudinally related to FEE frequency in the toddler view, with higher negativity and lower falling reactivity associated with a greater percentage of faces in the toddler view being neutral. To further understand these relationships, future research should explore these effects with a wider sample and look to understand causal relationships, particularly within clinical populations.

Keywords: expressions, emotions, development, individual differences

Factors Affecting the Frequency of Emotional Expressions Within the Toddler Egocentric View

The human face provides rich information on the intentions, cognitions and affective states of others for typically developing adults, in part from FEE (Facial Emotional Expressions) (Jack and Schyns, 2017). There is diversity in ability to interpret these FEE—for example, in clinical populations, there are deficits observed in autistic individuals (Cook et al., 2013), alcoholic individuals (Kornreich et al., 2002) and those with paranoid schizophrenia (Mendoza et al., 2011). Beyond this, there is a significant range of FEE recognition ability within the typical population (Scherer and Scherer, 2011). There is evidence from typical and clinical samples that variations in FEE processing ability are associated with differences in social functioning (Trevisan and Birmingham, 2016) and empathy (Besel and Yuille, 2010). The origins and nature of this diversity in FEE processing ability remain unclear but understanding precisely how typical individuals come to be able to interpret FEE could be key.

The development of FEE processing primarily takes place in early childhood, and this input does seem to shape proficiencies in FEE processing—for example, the ethnicities of the faces that people see most often, especially early in development, determines the proficiencies with which they process FEE exhibited by these different categories of face (Elfenbein and Ambady, 2002; Safar et al., 2017). However, no attempts have been made to record diversity in the visual input of FEE at any age and how this is related to (and possibly shaped by) other individual differences. While currently there is no published research that directly measures the degree to which people of any age see FEE, our research on this topic would suggest that there

is great variation in the degree to which individuals are exposed to these FEE (see chapter 4 of this thesis). For example, we found that across participants, the average percentage of faces in the toddler view receiving the FEE label of happiness was $M = 13.58\%$, $SD = 5.94\%$; while for anger this was $M = 5.01\%$, $SD = 5.84\%$. This would suggest highly variable FEE frequencies, possibly particularly for negative FEE such as anger. Currently, there has been no direct assessment of possible sources of this variation.

Further, there is research to suggest that temperament (Fox and Henderson, 1999; Young, Fox, and Zahn-Waxler, 1999) and the presence of siblings (Sang and Nelson, 2017; Zhang et al., 2021) may be important factors in socioemotional development and FEE processing. So, we sought to examine whether these factors were formative in shaping the variation in the FEE visual input within the toddler view.

Diversity in Facial Emotional Expression Processing Ability and its Implications

Not all individuals within an age group show the same proficiency in FEE recognition and understanding (Cook et al., 2013; Kornreich et al., 2002; Scherer and Schere, 2011). The origins of this variable ability are unclear, but it appears to be important for functioning. Here, we outline what FEE processing “ability” might be, before outlining what the range of ability might be within the wider population and what the significance of this variation might be.

The question of FEE processing “accuracy” or ability is a complicated one, with many differing approaches taken. Some, taking a “basic” emotions view, consider that an FEE reflects

primarily or entirely an internal state, and so accuracy should be inferred based on top-down labels that represent the “true” internal state of the model (e.g., Ekman, 2014; Ekman and Friesen, 1971). Others take a more constructivist approach, conceiving an FEE as the result of a mixture of internal state and social signalling, and perhaps considering a “correct” interpretation to only be possible with accompanying contextual information (Barrett, 2011; Barrett et al., 2019). However, typically the research we will outline here takes performance and accuracy to be the speed with which an individual categorises an FEE either based on a prototypical label (if these are standardised examples, such as those produced by Ekman and colleagues) or based on a consensus response. Often, the appearance of prototypical examples is guided by FACS (Ekman and Rosenberg, 1997). A range of performance has been recorded on such tasks. It seems that performance for FEE recognition tasks often diverges into dissociable skills for positive and negative emotions– an impairment or increased ability in one does not necessarily confer an impairment in the other (Schlegel et al., 2012; Suzuki et al., 2010).

Firstly, at one end of the performance spectrum, there is enhanced performance. There are those within the population whose performance on FEE tasks is superior or highly superior: face identity super-recognizers outperform controls on FEE recognition tasks (Ali and Davis, 2020), including on more challenging tasks wherein aspects of the face are occluded (Noyes et al., 2021). While not well or clinically defined, this enhanced ability seems to be significant for functioning. In typical populations, there is a positive correlation between the accuracy of identifying basic FEE and levels of empathy and empathic concern (Besel and Yuille, 2010). This association suggests that proficient FEE processing is crucial for social functioning and the

development of empathetic relationships. So, if elements of the environment enhance FEE processing accuracy, then optimising environmental learning opportunities for infants and toddlers could optimise mature social cognition and functioning.

Then, there are deficits in FEE recognition that have been noted as components of wider disorders. Deficits in processing emotional expressions, particularly facial emotional expressions, have been observed in Autistic people (Cook et al., 2013), subgroups of those diagnosed with alcoholism (Kornreich et al., 2002) and paranoid schizophrenics, as well as their first-degree relatives (Mendoza et al., 2011). Particularly, a deficit in recognising and interpreting FEE is so marked in and characteristic of Autism that it has been proposed as a distinguishing or diagnostic factor (Loth et al., 2018). The deficit in FEE processing for Autistic individuals presents as particular difficulty with negative and threat-related facial emotional expressions (Wallace et al., 2008) and is a good predictor of everyday social functioning (Trevisan and Birmingham, 2016). This demonstrates that not only do deficits in FEE recognition ability exist, but that this diversity may be crucial in understanding the core of socio-emotional disorders.

There also may not quite be deficits in FEE processing that present as standalone or near-standalone impairments. Unlike the specific facial identity recognition in prosopagnosia (Damasio, 1985), there is no known and named disorder that selectively and solely impairs FEE processing and recognition. While prosopagnosics are typically considered to be specifically impaired in facial identity recognition, there is no known equivalent for FEE processing. Though

there is some evidence that at least some individuals with prosopagnosia show deficits in some aspects of FEE processing, especially for negative FEE (Fiset et al., 2017; Stephan et al., 2006), for many prosopagnosics this ability is intact (de Gelder et al., 2003). This could imply that while FEE processing is difficult to extract from wider face processing, it can be impaired alongside it absent other, wider or syndromic deficits.

Further, there are seemingly some “typical range” low performers. For example, older adults show reduced FEE processing ability, particularly for fear, sadness and happiness (Horning et al., 2012). Typical adults also show a range of performances on more challenging facial emotion recognition tasks, with men generally not performing as well (Kessels et al., 2014). Lower general cognitive ability, lower theory of mind and higher neuroticism are also associated with lower scores on facial emotion recognition tasks (Andric et al., 2016; Lee et al., 2014). So, healthy individuals can demonstrate a range of facial emotion recognition abilities, including on the lower end of the typical distribution. These typical range variations have been associated with a variety of professional and social outcomes (Jaksic and Schlegel, 2020; Schreckenbach et al., 2018). This would imply that these variations are meaningful.

Individual differences in FEE recognition ability can be traced back developmentally to at least preschool (Feldman et al., 1993), but their specific developmental origins and the mechanisms through which they may come about is unclear. This existence of diversity in FEE processing implies either one or both of the following statements is true: inherent differences at birth (for example, genetic or hormonal) result in differences in FEE processing ability (in line

with the theories of Nelson and de Haan, 1997); or that the differences in the environmental input result in differences in FEE processing ability (in line with the broad theories of development proposed by Meadows, 1996, for example); or that an interaction between inherent differences at birth and the environmental input results in differences in FEE processing ability.

Arguments to this effect have indeed been before based on a wide variety of evidence suggesting a role for both genes and environmental input in the development of FEE processing diversity. Facial emotion recognition, including deficits in it associated with disorders, appears to be at least in part genetically mediated (Bediou et al., 2007; Skuse et al., 2014). Further, the CONSPEC/CONLERN mode would suggest that the environmental input is essential for developing more sophisticated face processing skills such as the ability to decode facial emotional expressions, and that this is in part the function of the newborn face bias (Johnson, 2005). This line of argument has been extended to a genome-environment interaction for Autism, with suggestion that largely genetically mediated disruptions to early face looking behaviour disrupt development of socioemotional skills (including FEE processing skills) through disrupting the visual input of faces (Zhou et al., 2024). There is also evidence for a gene-environment interaction in the development of this skill in healthy individuals (Hartling et al., 2019). However, the specific role the environmental visual input may play in shaping later FEE processing proficiency remains unclear.

Prior Research on the Importance of Variation in the FEE Input

Evidence that the nature of FEE processing depends on and varies to at least some extent with the environmental input can be seen in several different sources and phenomena. Here, we will discuss the implications of evidence taken from the FEE ORE and atypical visual input. This research suggests indirectly that the environmental input of FEE may be significant in shaping later FEE processing proficiency.

The clearest indirect example of evidence for the importance of early socioemotional visual input can be seen in the other-race effect (ORE). This is the phenomenon of impaired face processing for ethnicities or races with which a person is not perceptually familiar, and it extends to FEE recognition (Elfenbein and Ambady, 2002; Yan et al., 2016). There is research to indicate that people tend to have enhanced accuracy in judging FEE within their familiar cultural groups, typically their ethnic cultural in-groups (Elfenbein and Ambady, 2003). The ORE is emergent by 3-months for facial identity recognition, and stable by 9-months, with infants processing faces from ethnicities that they have and haven't been exposed to differently. (Bar-Haim et al., 2006; Kelly et al., 2005; Kelly et al., 2007; Sangrigoli et al., 2004). An unstable FEE ORE for FEE can be observed as early as 6-months, as infants at 6-months-old who have primarily been exposed to own-race faces show spontaneous preference for fearful faces only in own-race faces (Safar et al., 2017). There is also evidence that, between 5- and 9-months, qualitative differences emerge in processing same- and other-race facial and vocal emotional expression congruence (Vogel et al., 2012). However, infants as young as 7 months can still discriminate between same- and other-race happy and fearful expressions (Geangu et al.,

2016), indicating that the ORE in FEE processing is not robust at this early age. Further, there is mixed evidence on when and if a distinct FEE ORE emerges in childhood (Segal et al., 2019). So, while the identity ORE certainly emerges as a result of the environmental visual input in infancy, the evidence is more limited and mixed on whether and when which elements of an FEE ORE emerge over infancy and early childhood. However, at least some elements of this FEE perceptual narrowing seem to be emerging around the same time that face identity processing perceptual narrowing takes place. This likely indicates that at least some elements of FEE visual input at this point are important for emergence of individual differences in proficiency.

We can also see the role of the environmental input in shaping individual differences in FEE processing when examining the impact of impoverished or atypical environmental input. An example of this can be the children of clinical populations. For example, infants of mothers exhibiting depressive symptoms during pregnancy and the early postpartum period, who may exhibit reduced facial expressiveness (Renneberg et al., 2005), show differences in emotional expression processing, with a bias towards fearful faces (Kataja et al., 2018). Maternal depression predicts FEE recognition into childhood (Székely et al., 2014). This is in absence of an impaired FEE processing in children of parents with mental illnesses as a whole compared to children of parents without mental illnesses as a whole (Werkmann et al., 2024). While the effect of parental depressions could possibly be the result of shared genetic and environmental risk of mental illness between parent and child (Uher, 2014), as a wide range of mental illnesses are associated with differences in FEE recognition (Collin et al., 2013), a component of this causal pathway to later mental illness could even be through atypical visual input provided by

parents with depression. This would suggest that the visual input underpins later variation in FEE processing proficiency.

Another example of highly atypical input being associated with atypical FEE processing can be seen in neglected or abused children. Neglected children might see fewer FEE overall or a different distribution of FEE, and abused children, whose caregivers report higher levels of anger (Plate et al., 2019), might experience atypical visual exposure to negative FEE. Indeed, while the evidence on neglect and abuse is somewhat mixed, it generally indicates that abused and neglected children go on to show differences in FEE processing, particularly for negative emotions including anger (Doretto and Scivoletto, 2018; Pollak and Tolley-Shell, 2003). There is some indication that severe abuse or neglect can be associated with differences that are marked enough as to be deficits—severely environmentally deprived and institutionalised children correctly identify emotions in photos and stories significantly less frequently than children exposed to typical environmental input (Sloutsky, 1997). However, even in these groups, evidence is a little mixed, with some research finding only mild deficits largely restricted to happy expressions in children raised in these same environments (Moulsen et al., 2015). That said, these kinds of highly atypical environmental inputs do seem to be associated with differences in FEE processing. So, profound differences in the input may lead to differences in later FEE processing ability, even if these differences are milder than we might have expected.

While this evidence from atypical environments such as the infants of parents with mental illnesses or abusive parents could be explained through confounds such as shared

genetic variance or comorbid disorders, confounds of this nature are less likely for research examining infants born with bilateral congenital cataracts. Infants with congenital cataracts are otherwise typical but are precluded from visually accessing the environment. These infants, deprived of face input in the first weeks or months of life, go on to show subtly but notably qualitatively different cognitive representations of FEE even in adulthood, with the dimension of arousal not represented in the same way that it is in typical adults (Gao et al., 2013). So, atypical FEE input in infancy alone can translate into later differences in FEE processing decades later, even when deprivation is brief and confounds unlikely.

The broad relationship between variation in the degree to which FEE are present (or likely to be present) in the views of infants and young children and variation in later FEE input would suggest that the statistics of the visual input are important in shaping the individual differences in later functioning. This would mean that understanding variation in the FEE visual input as well as the potential sources of this variation would be essential in building models of both typical and atypical socioemotional development. However, the evidence that does exist for a relationship between the degree to which infants and young children see different FEE and socioemotional development is indirect, and there is no direct prior quantification of the frequency of different FEE in the visual field.

The Importance of Variation in Visual Input in Toddlerhood

We will be focusing on factors predicting variation in the FEE frequency within the toddler view for several reasons. Firstly, toddlerhood is a key time of rapid socioemotional

development (Brownell et al., 2007). This is also the age at which individual variation in socioemotional processing skills can start to become apparent to parents and observers. For example, differences in social functioning are apparent in toddlers who go on to be diagnosed with Autism by 24-months, and it is in toddlerhood that diagnosis of this condition becomes possible (Landa et al., 2007). Even in typical toddlers, clear variation in social cognition is apparent and related to family variables and toddler experience (Bulgarelli and Molina, 2016). This demonstrates how divergent developmental trajectories become in this period.

Another factor that makes toddlerhood unique is the emerging independence of the individual. Toddlers, in part because of their emerging motor and verbal skills, are more able to and more interested in asserting their will and exploring independently, often in line with their desires and temperament (Colson and Dwarkin, 1997). This also means that their individual differences, including temperament, will have more of a role in shaping their social environments and relationships with their parents (Andreadakis et al., 2020). There is indeed evidence that toddlers do directly shape their own inputs according to their motivations (Bambach et al., 2018; Xu et al., 2011). So, toddlers are more able to shape their own inputs than infants, often in very direct ways, and often intentionally shape their own environments and inputs. Toddlers may shape the degree to which they see different FEE either by directly seeking out certain stimuli, concurrently eliciting responses from others, or indirectly through the relationships that they build with others.

Overall, we have focused on toddlerhood as a key point in development wherein individual differences in socioemotional processing are already clear, but development is still rapidly ongoing and possibly shaped by the increasing independence of infants' exploration. This makes toddlerhood a likely candidate for a clear role of internal and external individual differences in shaping FEE visual input.

Prior Research Characterising the Visual Input of Facial Emotional Expressions

There is very limited prior direct research on the environmental input of infants, children and adults with respect to facial emotional expressions. Though, there is research on the frequency with which infants, toddlers and adults see faces. There are frequent faces in the visual input at birth and in adulthood (approximately 25% of the visual input), but frequency declines to less than 10% of the visual input in toddlerhood (Jayaraman, Fausey and Smith 2015; Jayaraman, Fausey and Smith 2017; Oruc et al., 2019). To some extent, there is also evidence for the characteristics of the faces that infants see, with these faces tending to be a smaller number of familiar individuals (Sugden et al., 2019). In chapters 1 and 2, we provided further evidence for these infrequent faces as compared to bodies in the toddler view and demonstrated that this did not generalise to caregivers in the same environment, also finding that temperament and the presence of child siblings was not associated with individual differences in the frequencies with which toddlers see faces and bodies.

In chapter 4 of this thesis, we presented the first analysis of the visual statistics of the toddlers' and caregivers' visual inputs with respect to FEE. We found that most of the faces

toddlers and their caregivers are exposed to are neutral or happy, and that negative FEE, especially fear, are quite rare, especially when considered as a portion of overall frames as opposed to a percentage of frames containing faces. However, there was significant variation in the degree to which toddlers and their caregivers saw these different FEE, necessitating exploration of the degree of variation and the factors that may explain it. It could be that these individual differences in the input are important in shaping later socioemotional stimulus processing or even general social cognitive abilities. So, the first step is to understand the correlates of this variation. This is an important stepping stone in understanding just how variation in early FEE input may come about. There are several factors related to variation in FEE processing ability at older ages that may be informative in understanding the variation in the early visual input. Here, we will focus on the presence of siblings and toddler temperament.

The Role of Siblings

Family structure, particularly the presence of child siblings, is a key environmental factor that may influence individual differences in Facial Emotional Expression (FEE) processing. There are two primary reasons for this: the relationship between sibling status and general socioemotional development, and the pragmatic aspects of environmental visual accessibility.

The presence of siblings plays a significant role in socioemotional development. While emotional intelligence in adolescence and adulthood appears unaffected by sibling presence (Rodrigues and Abhyankar, 2012), childhood experiences tell a different story. The presence and gender of siblings have been associated with improved perspective-taking abilities, a crucial

component of cognitive empathy (Sang and Nelson, 2017). Additionally, the quality of sibling relationships, rather than merely their presence, is linked to a tendency towards prosocial behaviour (Smorti and Ponti, 2018). Furthermore, children with siblings tend to outperform only children on theory-of-mind tasks in early childhood (Zhang et al., 2021). This indicates that in some way, siblings enrich socioemotional development.

One way that siblings could enrich the visual input of FEE could be through providing an extra or qualitatively different source of FEE visual input. Given that children have not yet developed the emotional regulation skills of adults (Campos et al., 1989; Grolnick et al., 1999) or fully internalised display rules (Saarni, 1979), it is likely that the expressivity of child siblings differs from that of adult caregivers. So, perhaps they express more dysregulated and varied emotions that lead to more negative FEE to be present within the visual input.

Another reason to believe that FEE input is related to sibling status is based on practical considerations. Child siblings provide a diverse range of peer interactions within the home environment (Howe and Recchia, 2005). These interactions are particularly influential for toddlers because siblings are often at a similar height, making their faces more visually accessible compared to those of standing adults (Frank et al., 2013; Kretch, Franchak, and Adolph, 2014). Toddlers need to actively view the FEE of much taller adults, and looking away from distressing stimuli is a key self-regulation method (Rodriguez et al., 2005). This could mean that toddlers are less prone to intentionally access possibly distressing negative FEE on adult faces given the already difficult nature of this task. Child siblings, however, being incidentally

within eyeline, may provide a richer source of these FEE as they are visually hard to avoid.

Toddlers may also more often see siblings when not engaging in intentional interaction.

Together, this may mean that child sibling faces provide a more representative source of FEE than the difficulty to visually access adult faces.

The Role of Temperament

An important internal factor that could shape individuals' visual input is temperament in toddlers and infants, or personality in adults. Temperament refers to individual differences, at least partially innate, that encompass differences in action tendencies over a range of different domains that can broadly be separated into reactivity and regulation (Rothbart and Bates, 2007). Simply put, reactivity is the tendency of an individual to react in different situations, and regulation is their ability and tendency to up- or down-regulate these reactions (Rothbart et al., 2004). Though no direct research currently measures how much FEE individuals see across different ages, here we will outline the existing research that indicates temperament to be critical for individual differences in socioemotional development, including individual differences in FEE processing development.

A wide range of temperamental differences are implicated in social disorders such as Autism Spectrum Disorder (ASD), being associated with diagnosis or future diagnosis. For example, inhibition and aspects of surgency are linked to early autism signs in infancy (Chetcuti et al., 2020). Toddlers with ASD exhibit differences in regulation, surgency and differences in negative reactivity (Macari et al., 2017). Very early differences in perceptual sensitivity and

negative reactivity are also linked to later ASD symptoms (Narvekar et al., 2022). While there are a range of deficits and profiles in different individuals with ASD (Lord et al., 2009), the centrality of social processing, and particularly FEE processing, to ASD may indicate a role for temperament in these skill sets.

Indeed, it seems that early temperament predicts social functioning even in typical individuals, indicating perhaps a specific relationship. Dimensions of early temperament as measured in infancy are related to later social proficiency and development in toddlerhood, such as empathy and theory of mind (MacGowan et al., 2022; Mark et al., 2002; Mink et al., 2014; Wagers and Kiel, 2019; Young et al., 1999). It could be that temperament and socioemotional development are linked in a variety of ways— for example, more regulated infants could simply be better able to engage in necessary social cognitions, given that dysregulation could impair cognitive performance (Labouvie-Vief et al., 2009). It could also be that underlying differences in factors like serotonin transporter genes influences both social proficiency and regulation (Canli and Lesch, 2007). However, given that, as outlined above, there are likely relationships between aspects of the socioemotional environment and development, one possible explanation is that temperament might influence the environmental input. This could be either through a process of temperament determining reactions elicited from those around the infant or toddler, or through a process of infants and toddlers of different temperaments directly seeking out and attending to different aspects of the environment.

There are also direct relationships between key elements of temperament and the nature of FEE processing. There are concurrent relationships between aspects of temperament and the nature of concurrent FEE processing as early as infancy. For instance, negative emotionality is associated with the neural response to happy faces in 7-month-olds (Ravicz et al., 2015). Negative emotionality and self-regulation relate to infants' attentional responses to happy and fearful faces (Martinos et al., 2012). Further, negative emotionality and attentional control influence eye fixation behaviour for neutral versus emotional faces (Fu et al., 2020). This evidence suggests that infants with different temperaments respond differently to FEE at both physiological and behavioural levels, with altered looking behaviour.

These differences in FEE processing extend to concurrent relationships between temperament and FEE processing ability in toddlerhood. For example, in preschoolers, higher scores on the dimension of shyness are concurrently and longitudinally associated with FEE recognition scores, specifically for negative FEE (Strand et al., 2008). This would suggest that temperament could either shape or be otherwise related to the specific elements of socioemotional processing that are necessary for accurately categorising FEE.

One way in which these differences could come about is through temperament shaping looking behaviour. Infants and toddlers are not necessarily passive recipients of information but can actively influence their environmental input, including their visual environmental input (Smith et al., 2018). As discussed above, the relationship between temperament and FEE

processing initially presents as differences in attention and looking behaviour in infancy, evolving into differences in recognition accuracy by toddlerhood. One possibility is that the variations in looking behaviour shape the FEE input and thereby shape later processing ability.

Another way in which temperament could influence the visual statistic of the FEE input is through shaping the social responses of those around the toddler. Toddler negativity is associated with parental stress (Andreadakis et al., 2020), and variations in toddler temperament predict parenting child-rearing attitudes (Katainen et al., 1997). While given the genetically mediated nature of temperament (Saudino, 2005) it is plausible that elements of shared genetic variance explain the relationship between toddler temperament and parental affect and behaviour, it is likely that toddler temperament at some level shapes the parent-child relationship. This may then mean that the frequency with which caregivers direct different FEE at the toddler varies with toddler temperament. At this level, the relationship may be longitudinal.

However, toddlers of different temperaments may also be related to different social responses concurrently because of specific behavioural differences. High negativity and low regulation are associated with behavioural problems in children and toddlers (Carrasco et al., 2012). Given that parents selectively employ elements of emotional expressivity in response to child transgressions (Dahl et al., 2014), this may then mean that children with high negativity and low regulation then see more negative FEE and fewer positive FEE.

Based on this, we expect aspects of temperament to be significantly related to the frequencies with which toddlers see different FEE both concurrently and longitudinally. While this is a partially exploratory study, being the first paper on this topic, we would expect negativity to be particularly significant in shaping the frequency with which toddlers see different FEE, given that it is associated with many aspects of FEE processing, general socioemotional processing and parental attitudes and affect. Regulation is also likely to be significant given its association with the behavioural issues that are likely to shape the parent-child relationship longitudinally and with FEE processing concurrently. It seems likely that other, specific dimensions of temperamental factors are significant in shaping the FEE input. One that could be important given its relationship with FEE identification accuracy is shyness in toddlers.

Outstanding Questions and Hypotheses

This study aims to explore the degree to which the presence of siblings and toddler temperament are associated with the relative frequency with which the six primary FEE and neutral faces appear in the toddler egocentric view.

This builds on our previous work on this dataset, which was the first attempt to quantify the typical visual input of FEE for any age group. As with the frequency of faces in the toddler and adult view, there was notable variation in the degree to which participants saw each emotional expression. Whether this is random or might be related to other factors that we know are significantly related to socioemotional development and proficiency is an outstanding question here.

Beyond temperament and sibling status, some examination of gender differences in the frequency with which toddlers see different FEE will be necessary as a confound of sibling status. While in the sibling group 40% of the participants were female, 91% of the participants without siblings were female. There is much evidence for a gender difference in FEE processing proficiency that emerges over infancy and childhood (Chaplin et al., 2005; McClure, 2000), so the effect of this confound must be considered when examining the frequency of different FEE in the views of toddlers with and without siblings.

While as entirely novel data the analysis will be partially exploratory, we hypothesise that: 1) toddlers with child siblings will see more negative, threat-related FEE than toddlers without siblings, as a result of these less regulated interaction partners (Saarni, 1979) being present and having faces easily within the toddler view (Kretch et al., 2014); 2) aspects of early temperament as measured using the IBQR will be longitudinally related to the FEE that toddlers see, broadly surgency, regulation and negativity; 3) aspects of concurrent temperament as measured using the ECBQ will be related to the FEE that toddlers see, broadly surgency, regulation and negativity; 4) Some specific dimensions of temperament will be associated with FEE frequency in the toddler view, including concurrent shyness.

Methods

See general methods section for full details.

Data Analysis Strategy

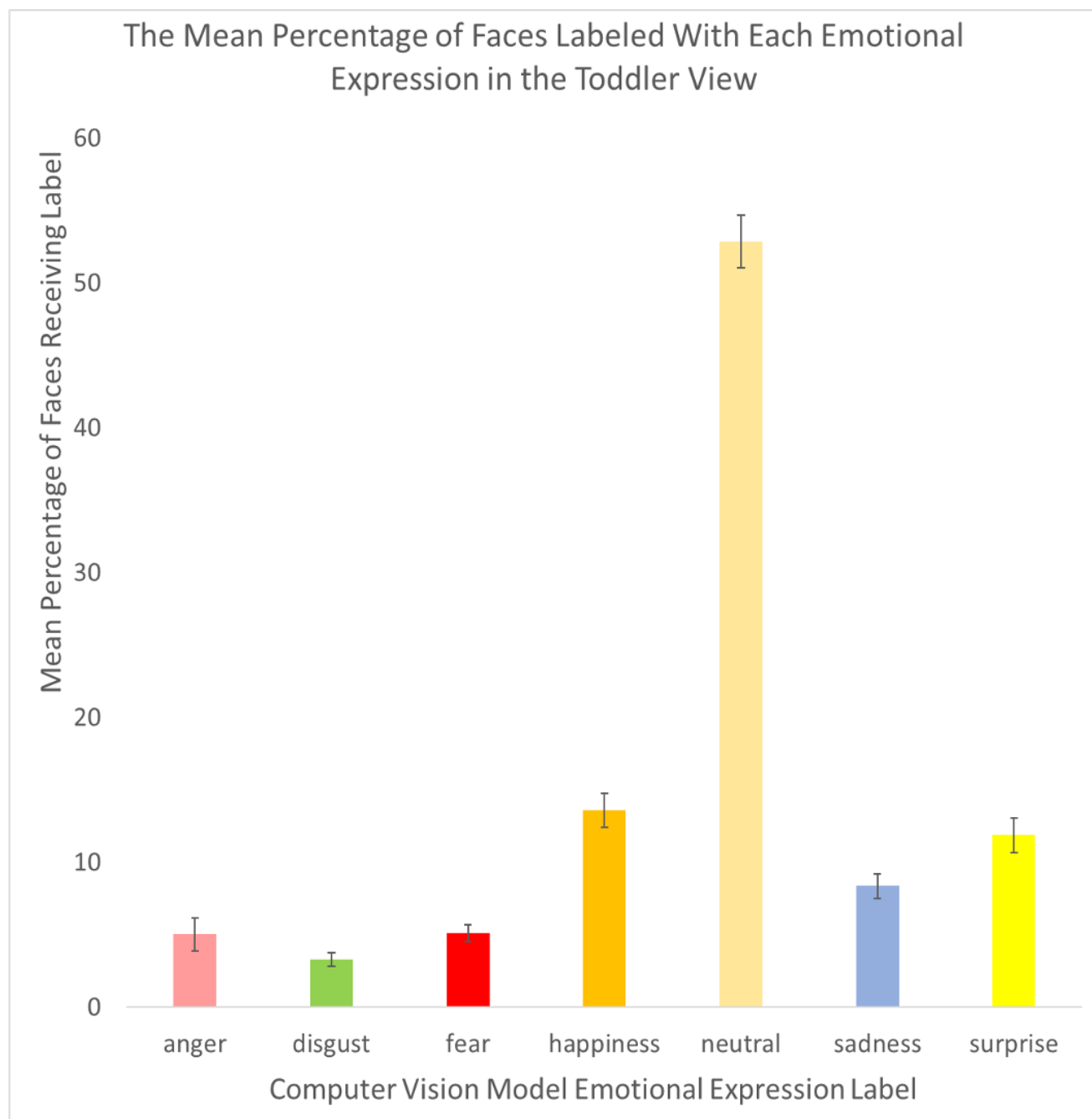
To maximise the sample size and statistical power here, we first consider the full $N = 26$ for the role of siblings, and then separately the $N = 23$ participants with IBQR data provided at 13-months and then the $N = 24$ participants with ECBQ data provided concurrent to recording.

Results***Overall FEE Frequency and Variability***

For the 26 families that successfully recorded footage from the toddler perspective, the most common label for face frames was neutral ($M = 52.84\%$, $SD = 9.34\%$). This was followed by happiness ($M = 13.58\%$, $SD = 5.94\%$); then surprise ($M = 11.86\%$, $SD = 6.14\%$); then sadness ($M = 8.34\%$, $SD = 4.32\%$); then fear ($M = 5.10\%$, $SD = 3.02\%$); with anger ($M = 5.01\%$, $SD = 5.84\%$) and disgust ($M = 3.28\%$, $SD = 2.50\%$) being the least common (see figure 1). While a full exploration of these frequencies will not be presented here-- see chapter 4 for full analysis of FEE frequency in the toddler view and its comparison with the FEE frequency in the caregiver view—the wide standard deviations in FEE frequency are of note here.

Fig. 1

A bar chart demonstrating the mean and standard error percentage of faces in the toddler view assigned each FEE label by our Retinaface/Arcface model.



Note. There was wide variation in the percentage of faces assigned each FEE label, but this is especially true for anger relative to its mean frequency.

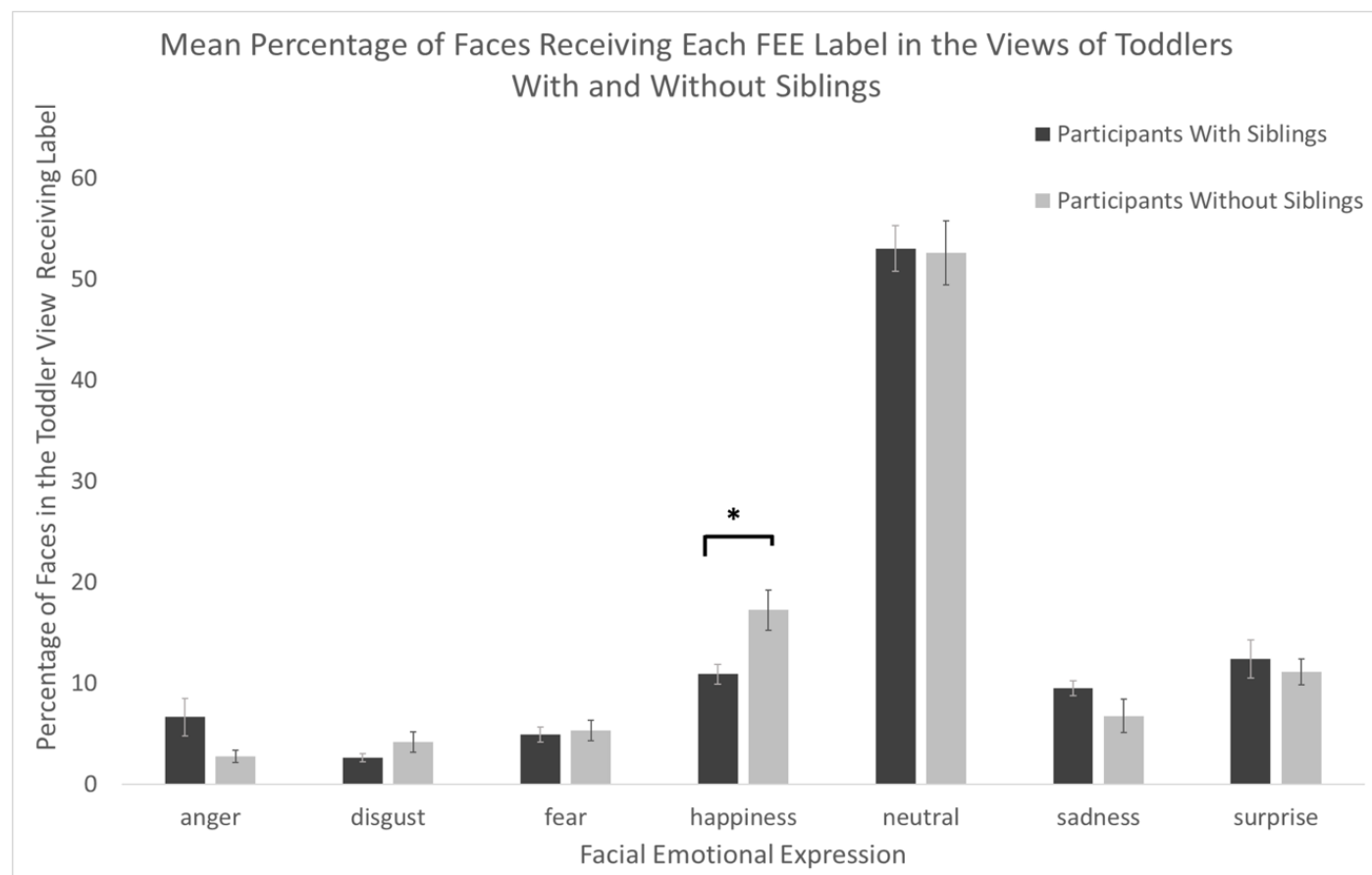
Presence of Siblings and FEE Frequency

Of the 26 participants who provided valid video data, 11 had no child siblings and 15 had at least one child sibling in the home. The mean and standard deviation frequencies of the faces receiving each emotional expression label within the views of toddlers with and without siblings was broadly similar, with overlapping standard deviations, but some FEE frequencies seemed to differ somewhat– for example, for participants with siblings, the average percentage of faces in the toddler view receiving a label of anger was $M = 6.64\%$, $SD = 7.16\%$; for happiness, this was $M = 10.89\%$, $SD = 3.77\%$. Equivalent figures for participants without siblings were anger $M = 2.77\%$, $SD = 2.01\%$; and happiness $M = 17.23\%$, $SD = 6.54\%$; neutral. For an illustration of these numbers, see figure 2.

To investigate these possible differences in FEE frequency in the view of toddlers with and without siblings, a MANOVA was conducted with an independent variable of sibling status and the dependent variables being the frequencies of the different FEE. Pillai's Trace was used to assess the multivariate effect due to unequal group sizes. The results indicated that there was a significant multivariate main effect of sibling status on the combined dependent variables, $V = 0.457$, $F(6, 19) = 2.667$, $p = 0.048$, $\eta^2 = 0.457$.

Fig. 2

The mean and standard error percentage of faces in the toddler view assigned each FEE label split by toddlers with and without siblings



Note. Error bars indicate the standard errors, and significant differences are marked with a ‘*’.

Separate follow-up univariate ANOVAs were conducted to explore the effects on each dependent variable. There was a significant effect of sibling status on happiness frequency, $F(1, 24) = 9.774$, $p = .005$, partial $\eta^2 = .289$, with a significantly higher percentage of happy faces in the views of toddlers without siblings. The following variables did not reach statistical significance: anger, $F(1, 24) = 3.011$, $p = .096$, partial $\eta^2 = .111$; disgust, $F(1, 24) = 2.671$, $p = .115$,

partial $\eta^2 = .100$; fear, $F(1, 24) = 0.118$, $p = .734$, partial $\eta^2 = .005$; neutral, $F(1, 24) = 0.013$, $p = .909$, partial $\eta^2 = .001$; sadness, $F(1, 24) = 2.736$, $p = .111$, partial $\eta^2 = .102$; and surprise, $F(1, 24) = 0.271$, $p = .607$, partial $\eta^2 = .011$.

Gender and FEE Frequency

While there did appear to be some relationship between sibling status and FEE frequency, gender posed a significant confound. Given the very unequal group sizes, a combined MANOVA was not possible. So, gender was examined separately.

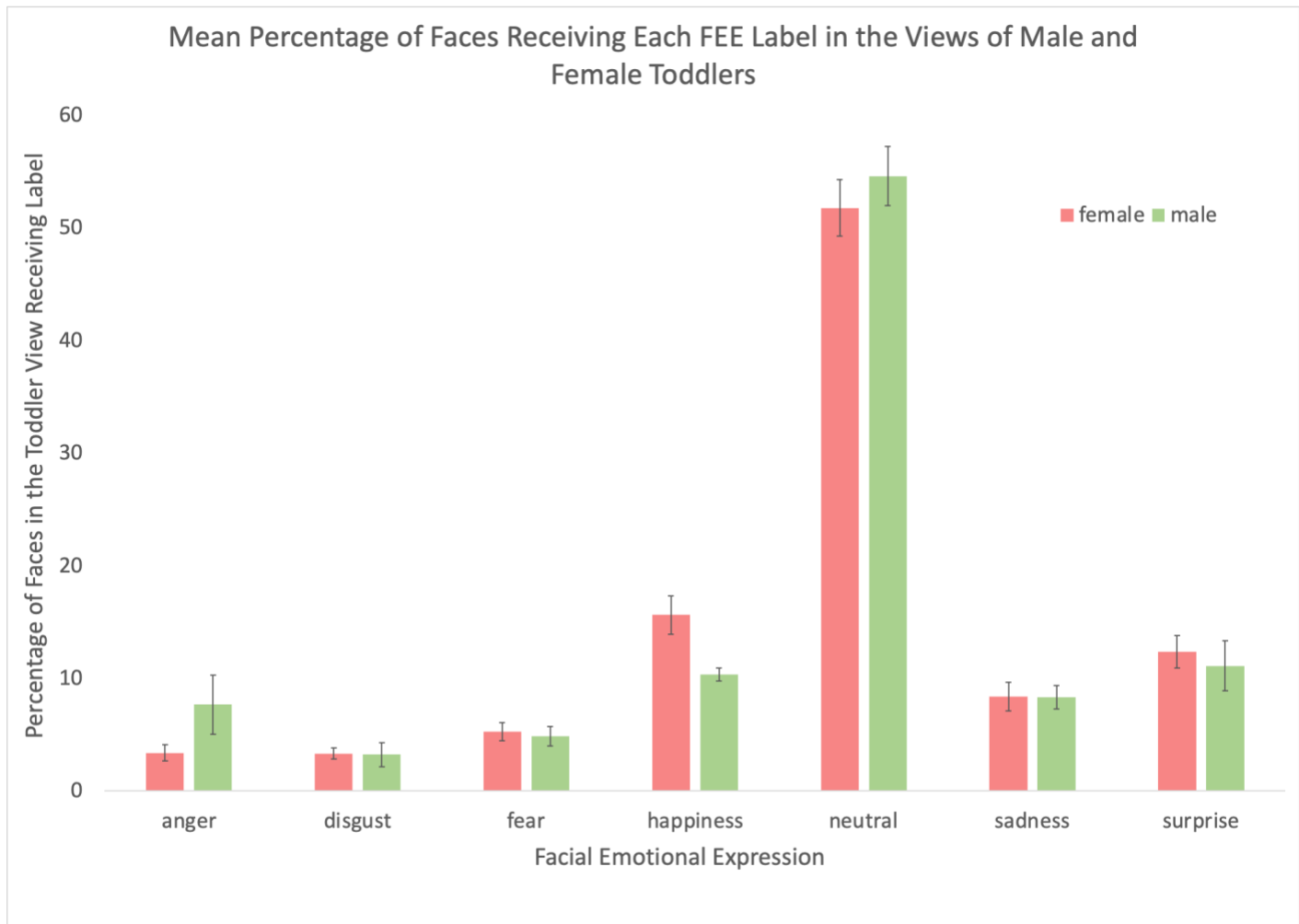
The percentage of faces receiving each FEE label in the toddler view appeared to be broadly similar by gender with overlapping error bars, but there were some FEE frequencies that appeared to be different by gender on the basis of descriptive statistics. For example, for female participants, the average percentage of faces in the toddler view receiving the FEE label of anger was $M = 3.36\%$, $SD = 2.85\%$; for happiness, this was $M = 15.61\%$, $SD = 6.74\%$. Equivalent figures for male participants were anger $M = 7.64\%$, $SD = 8.29\%$; for happiness, this was $M = 10.32\%$, $SD = 1.80\%$. For a visual representation of these figures, see figure 3.

To investigate these possible differences in FEE frequency in the view of male and female toddlers, a MANOVA was conducted with an independent variable of gender and the dependent variables being the frequencies of the different FEE in the toddler view. Pillai's Trace was used to assess the multivariate effect due to unequal group sizes. The results indicated that there was

no significant multivariate main effect of gender on the combined dependent variables, $V = 0.291$, $F(6, 19) = 1.299$, $p = 0.305$, $\eta^2 = 0.291$.

Fig. 3

The mean and standard error percentage of faces in the toddler view assigned each FEE label split by male and female toddlers.



Note. Error bars indicate the standard errors, and significant differences are marked with a '*’.

Longitudinal IBQR Temperament and FEE Frequency

Of the 26 participants, 23 had complete IBQ-R data for analysis (14 female, 9 male); this data was gathered at 13-months as part of a longitudinal research project. Associations between aspects of this personality scale and the presence of faces and bodies in the toddler view was examined.

Firstly, relationships between FEE frequency in the toddler view and the three factors proposed by Garstein and Rothbart (2003) were examined-- Surgency/Extraversion, Negative Affectivity, and Orienting/Regulation. Scatter charts were created for the frequency of each of the FEE in the toddler view and factors scores for surgency (fig. 4), negativity (fig. 5) and regulation (fig. 6).

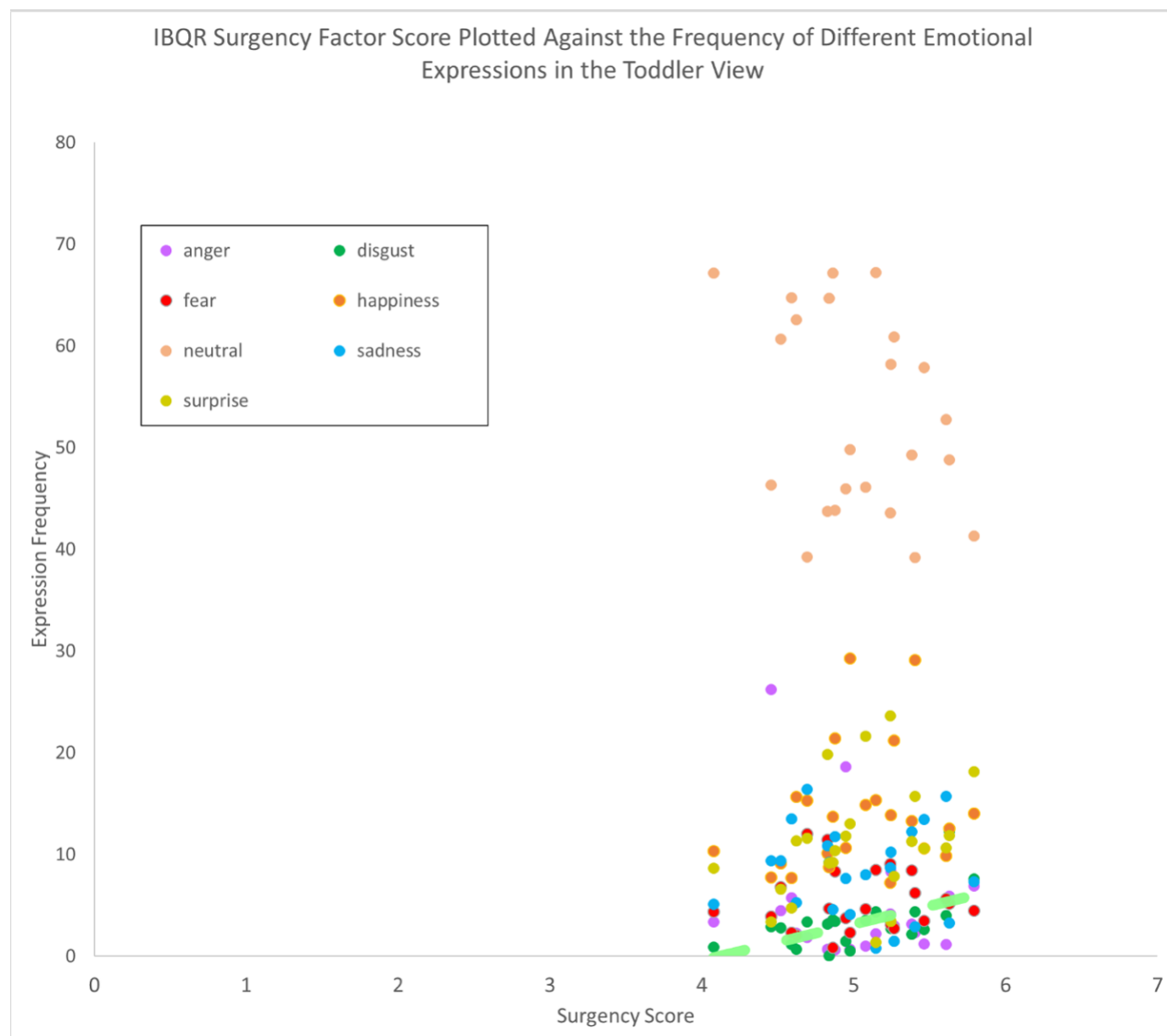
A multivariate multiple regression analysis was conducted to examine the effect of temperament factors (SUR, NEG, REG) on the frequency of the seven emotional expressions in the toddler view. The multivariate test results indicated that surgency did significantly predict the combined dependent variables (Pillai's trace $V = 0.595$, $F(6, 14) = 3.428$, $p = .027$, $\eta^2 = 0.595$). Negative affectivity and regulation did not show significant multivariate effects ($V = 0.325$, $F(6, 14) = 1.124$, $p = .398$, $\eta^2 = 0.325$; and $V = 0.343$, $F(6, 14) = 1.226$, $p = .350$, $\eta^2 = 0.344$ respectively).

Univariate analyses showed that surgency was a significant predictor of disgust FEE frequency ($F(3,19) = 7.59, p = .013, \eta^2 = 0.281$), with higher surgency being associated with a higher frequency of disgust (see figure 4), and negativity was a significant predictor of neutral FEE frequency, with higher negativity scores associated with a higher frequency of neutral expressions (see figure 5) ($F(3,19) = 4.94, p = .039, \eta^2 = .206$). No other significant effects were observed. These results suggest that higher levels of surgency are associated with increased frequencies of disgust in the toddler view, and higher levels of negativity are associated with increased frequencies of neutral expressions in the toddler view.

Given the small sample size and the multiple regressions employed, these results were checked with bootstrap analysis. After conducting a bootstrapping analysis with 1000 samples, the relationship between surgency and disgust was no longer significant (bootstrap $p = .174$). The 95% bootstrap confidence interval for the parameter estimate ranged from $-.501$ to $.626$, suggesting substantial variability and including zero, indicating no significant effect. However, the relationship between negativity and neutral remained significant (bootstrap $p = .035$). The 95% bootstrap confidence interval for the parameter estimate ranged from $.457$ to 5.568 , not including zero and suggesting a significant effect.

Figure 4

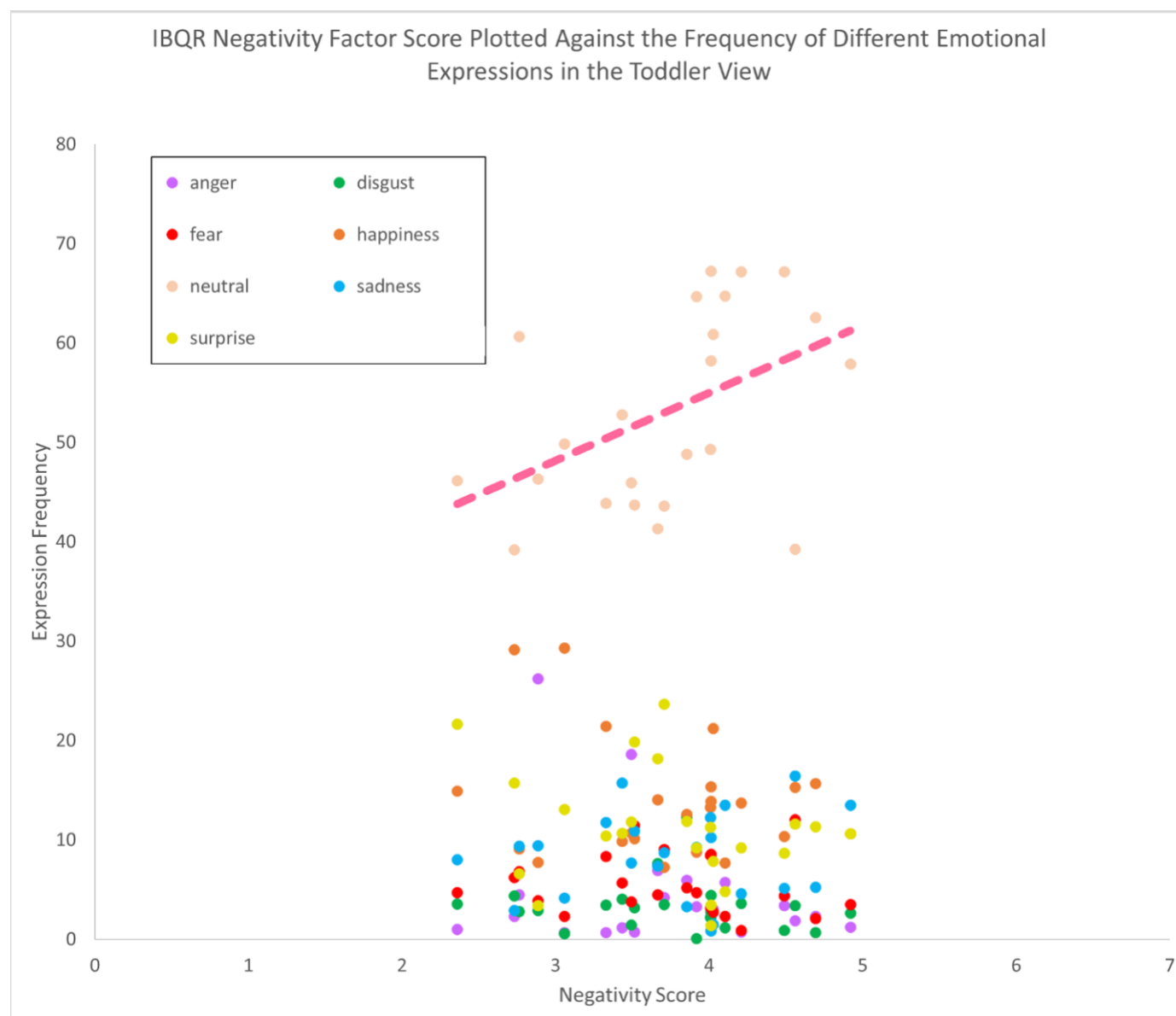
The percentage of faces receiving each FEE label as plotted against surgency factor scores



Note: This chart demonstrates the relationship between the frequency of different FEE in the toddler view and surgency factor scores. The dashed pale green line indicates the relationship between surgency and the frequency of disgust FEE, which was not significant after bootstrapping.

Figure 5

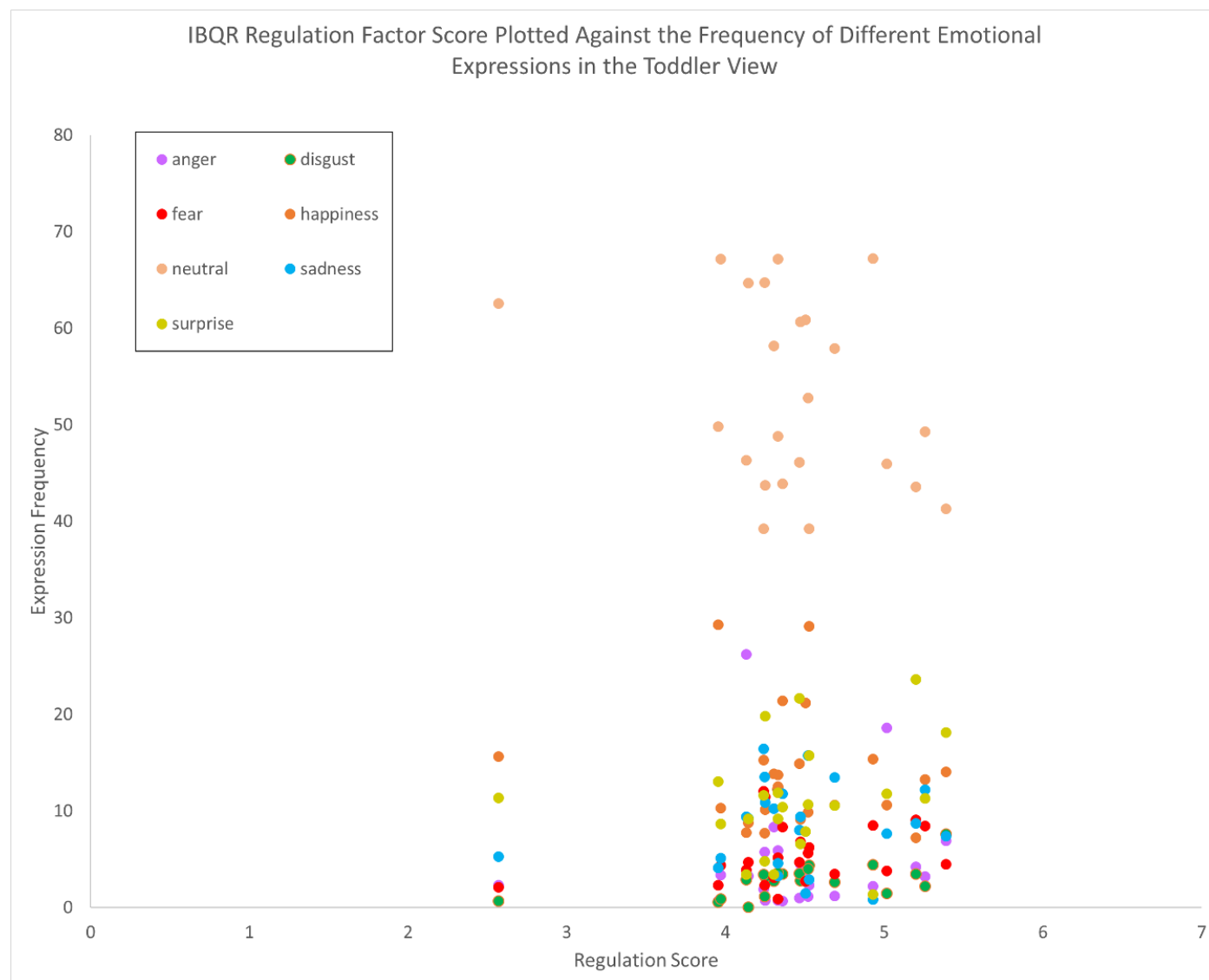
The percentage of faces receiving each FEE label as plotted against negativity factor scores.



Note: This chart demonstrates the relationships between the frequency of different FEE in the toddler view and negativity factor scores. The dashed pink line represents the relationship between negativity and the frequency of neutral FEE, which remained significant after bootstrapping.

Figure 6

The percentage of faces receiving each FEE label as plotted against regulation factor scores.



Note: This chart demonstrates the relationship between the frequency of different FEE in the toddler view and regulation factor scores. There were no significant relationships here.

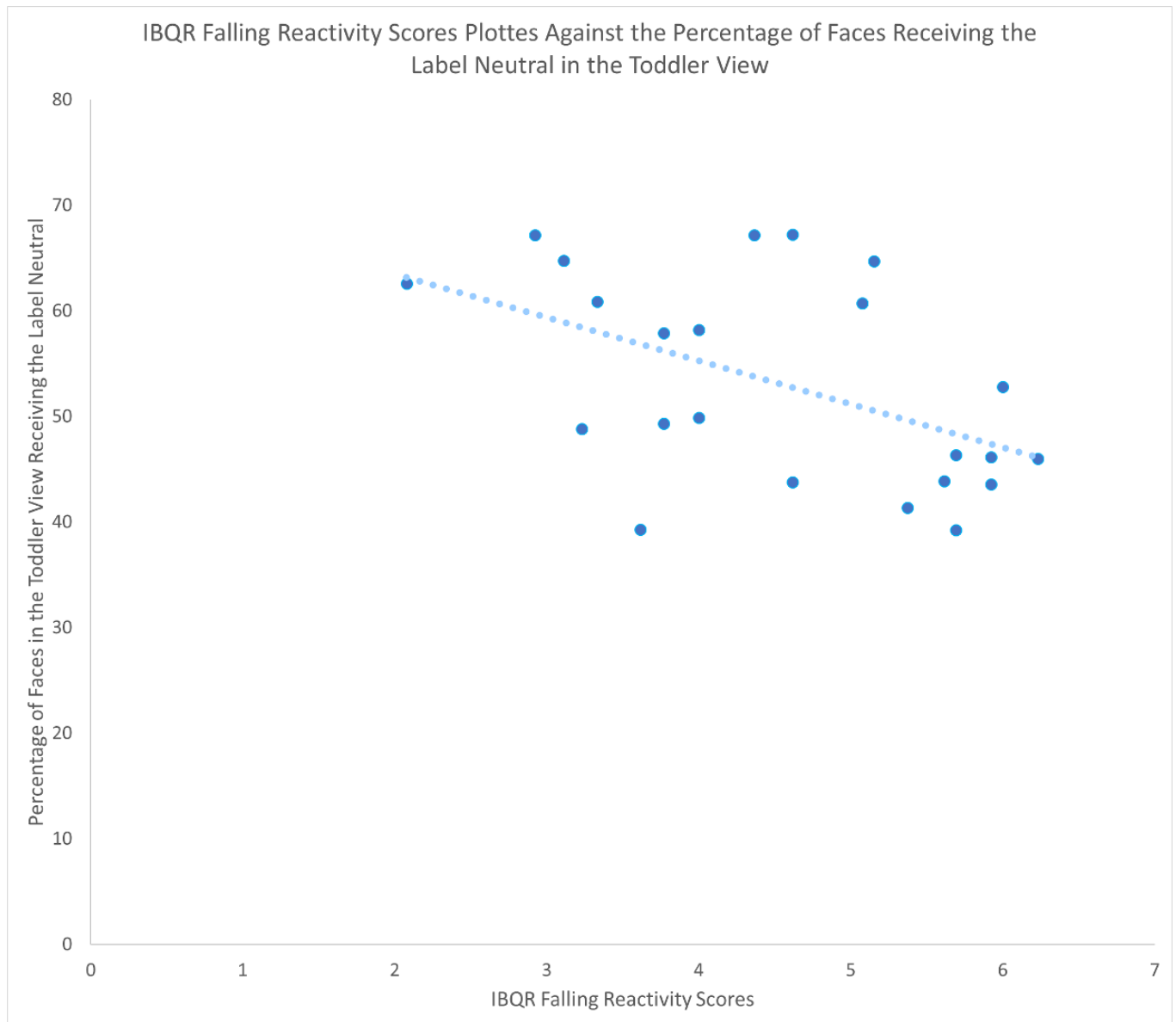
Next, the relationships between specific IBQR dimensions and the frequency of FEE in the toddler view was examined. Specifically, the dimensions loading on the factor negativity—sadness, distress to limitations, fear and falling reactivity—were examined, given the significant relationship between the factor negativity and FEE frequency above.

A multivariate multiple regression was conducted to assess the impact of the specific dimensions of the negativity factor on the frequency of the seven emotional expressions in the toddler view. The results indicated that no dimension significantly influenced the combined dependent variables: distress to limitations $V = .290$, $F(6, 13) = .885$, $p = .532$, partial $\eta^2 = .290$; fear $V = .515$, $F(6, 13) = 2.299$, $p = .098$, partial $\eta^2 = .515$; sadness $V = .309$, $F(6, 13) = .968$, $p = .483$, partial $\eta^2 = .309$; falling reactivity $V = .484$, $F(6, 13) = 2.034$, $p = .133$, partial $\eta^2 = .484$.

Univariate analyses showed that sadness was a significant predictor of disgust frequency ($B = -1.566$, $p = .050$), with higher sadness associated with less frequent disgust. Fearfulness was a significant predictor of happy frequency ($B = -2.725$, $p = .038$), with higher fearfulness associated with lower frequencies of happy faces. Falling reactivity was a significant predictor of neutral frequency ($B = -4.306$, $p = .019$), with higher falling reactivity scores associated with lower frequencies of neutral faces. No other significant effects were observed.

Figure 7

The percentage of faces receiving the label neutral as plotted against falling reactivity scores.



Note: This chart demonstrates the relationship between the frequency of neutral faces in the toddler view and falling reactivity dimension scores. Higher falling reactivity scores were associated with less frequent neutral faces, or more expressive faces overall. This relationship was significant after bootstrapping.

These results were checked with bootstrap analysis. After conducting a bootstrapping analysis with 1000 samples, the relationship between sadness and disgust was no longer significant (bootstrap $p = .348$). The 95% bootstrap confidence interval for the parameter estimate ranged from -3.409 to 1.078, suggesting substantial variability and including zero, indicating no significant effect. The relationship between fearfulness and happiness was also no longer significant (bootstrap $p = .120$). The 95% bootstrap confidence interval for the parameter estimate ranged from -5.608 to -.013, suggesting substantial variability and including zero, indicating no significant effect. However, the relationship between falling reactivity and neutral remained significant (bootstrap $p = .008$). The 95% bootstrap confidence interval for the parameter estimate ranged from -7.674 to -1.503, not including zero and suggesting a significant effect. See fig. 7 for a chart demonstrating the relationship between falling reactivity scores and the percentage of faces in the toddler view that were neutral. Higher falling reactivity scores were associated with less frequent neutral FEE, or a higher proportion of expressive, non-neutral faces.

Concurrent ECBQ Temperament and FEE Frequency

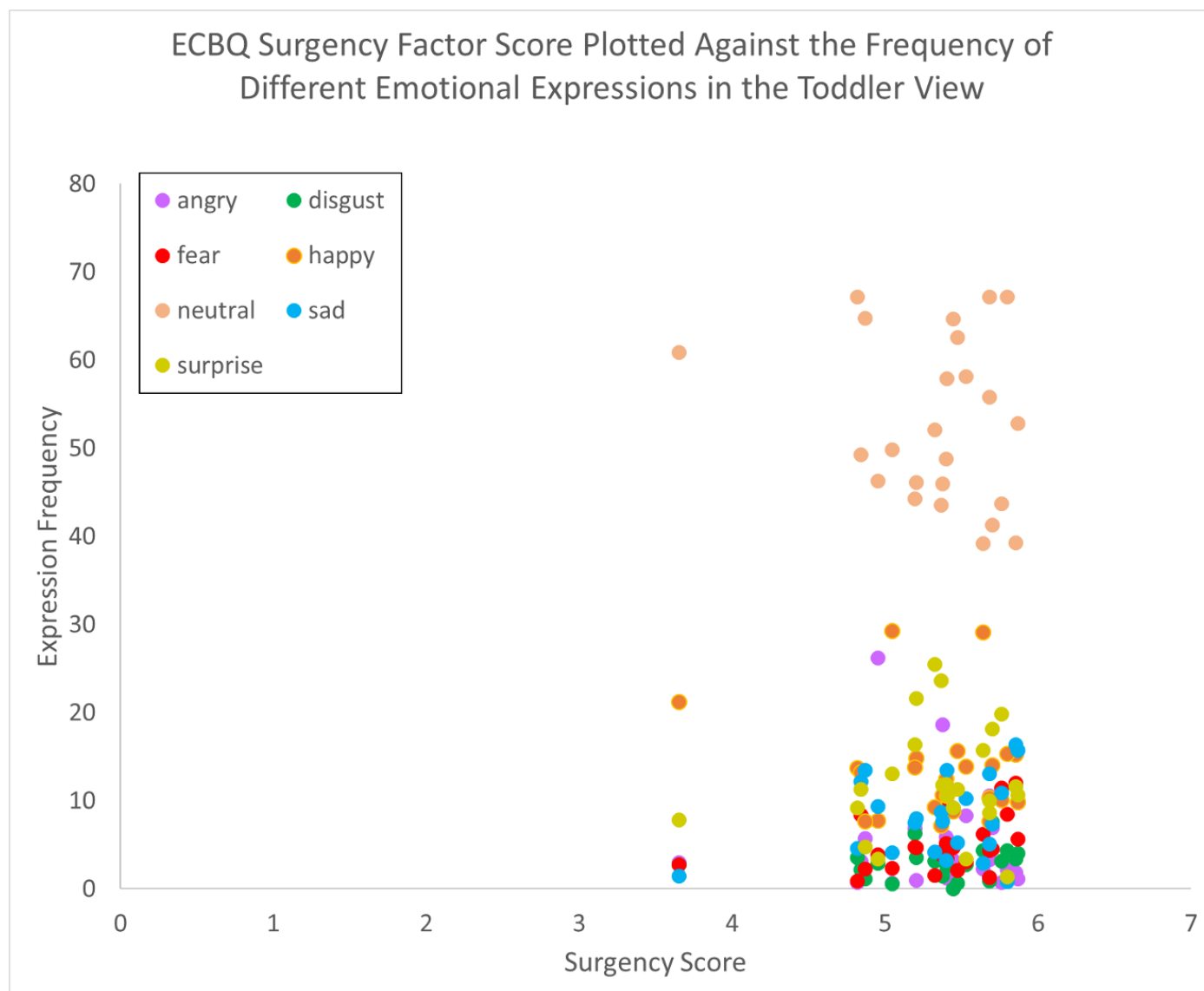
Of the 26 participants, 24 had complete ECBQ data gathered at around the time of recording available for analysis (15 female, 9 male). Associations between aspects of this concurrent personality scale and the presence of faces and bodies in the toddler view was examined.

Firstly, relationships between FEE frequency and ECBQ surgency, negativity and regulation factor scores were examined. Scatter charts were created for the frequency of each of the FEE in the toddler view and factors scores for surgency (fig. 8), negativity (fig. 9) and regulation (fig. 10).

A multivariate multiple regression was conducted to assess the impact of ECBQ personality factors (surgency, negative affectivity, and regulation) on the frequency of seven emotional expressions in the toddler view. The results indicated that the three factors were not significantly related to the combined dependent variables—surgency Pillai's Trace $V = .222$, $F(6, 15) = .715$, $p = .644$, partial $\eta^2 = .222$; negativity $V = .177$, $F(6, 15) = .538$, $p = .771$, partial $\eta^2 = .177$; regulation $V = .294$, $F(6, 15) = 1.041$, $p = .438$, partial $\eta^2 = .294$. Exploratory univariate analyses showed that regulation was a significant predictor of surprise frequency ($B = 5.685$, $p = .048$). No other significant effects were observed. These results could suggest that higher levels of regulation are associated with increased frequencies of surprise in the toddler view. Given the small sample size and the multiple regressions employed, these results were checked with bootstrap analysis. After conducting a bootstrapping analysis with 1000 samples, the relationship between regulation and surprise was no longer significant (bootstrap $p = .074$). The 95% bootstrap confidence interval for the parameter estimate ranged from $-.334$ to 13.819 , suggesting substantial variability and including zero, indicating no significant effect.

Figure 8

The percentage of faces receiving each FEE label as plotted against surgency factor scores.



Note: This chart demonstrates the relationship between the frequency of different FEE in the toddler view and surgency factor scores. None of these relationships were significant.

The percentage of faces receiving each FEE label as plotted against negativity factor scores.

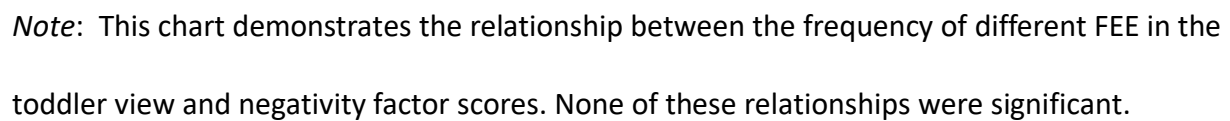
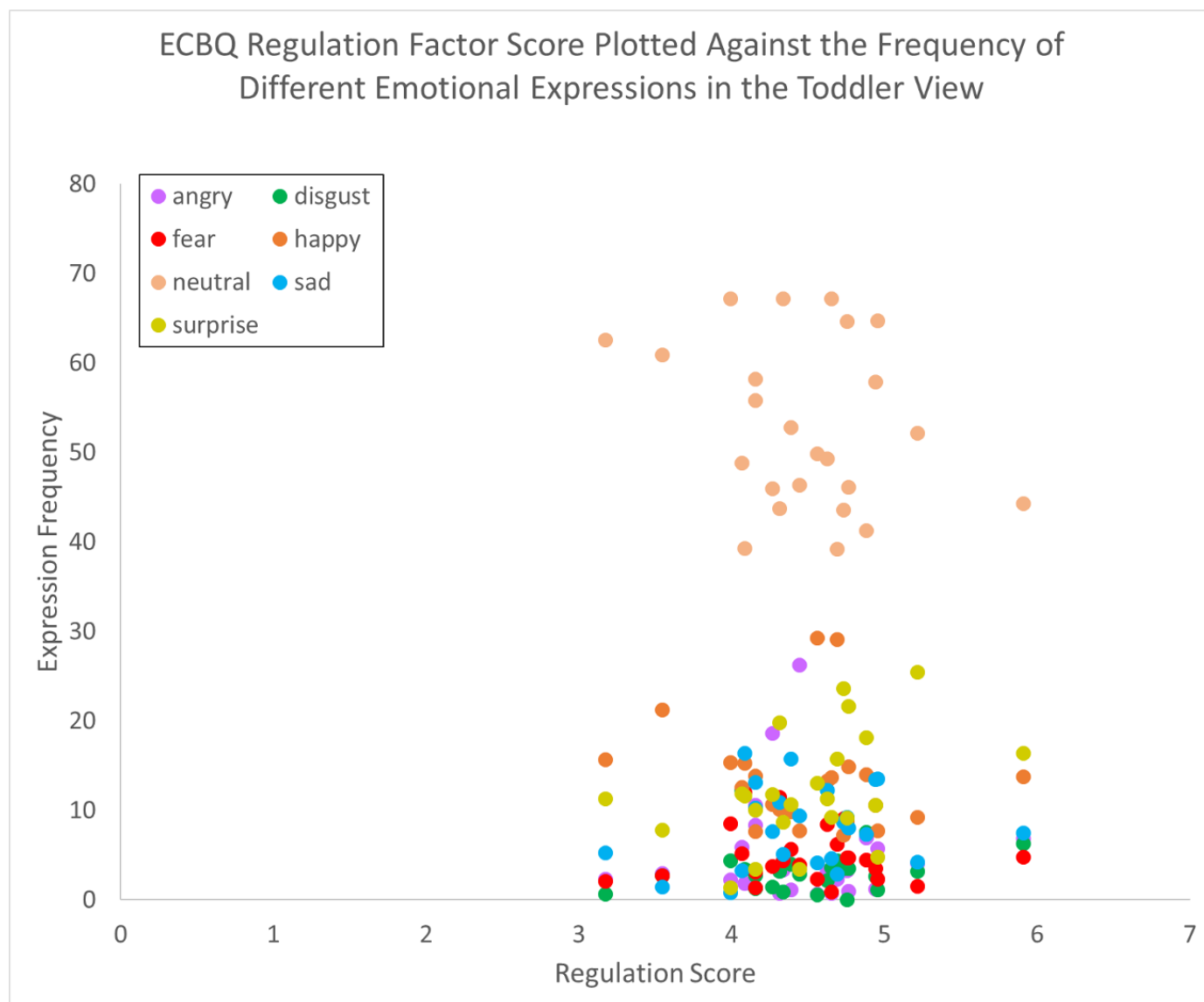


Figure 10

The percentage of faces receiving each FEE label as plotted against regulation factor scores.



Note: This chart demonstrates the relationship between the frequency of different FEE in the toddler view and regulation factor scores. None of these relationships were significant.

Next, the relationships between specific IBQR dimensions and the frequency of FEE in the toddler view was examined. Specifically, the dimensions loading on the factor regulation—Low Intensity Pleasure, Cuddliness, Attentional Focusing, Inhibitory Control, and Attentional Shifting-- were examined, since the relationship between regulation overall and the FEE input found in exploratory analyses, though not robust, could indicate a specific dimension loading on this factor related to the FEE input.

A multivariate multiple regression was conducted to assess the impact of the specific dimensions of the regulation factor on the frequency of the seven emotional expressions in the toddler view. The results indicated that no dimension significantly influenced the combined dependent variables: low intensity pleasure $V = .189$, $F(6, 13) = .504$, $p = .795$, partial $\eta^2 = .189$; cuddliness $V = .447$, $F(6, 13) = 1.751$, $p = .187$, partial $\eta^2 = .447$; attentional focusing $V = .372$, $F(6, 13) = 1.282$, $p = .331$, partial $\eta^2 = .372$; attentional shifting $V = .251$, $F(6, 13) = .725$, $p = .637$, partial $\eta^2 = .251$; inhibition $V = .280$, $F(6, 13) = .844$, $p = .558$, partial $\eta^2 = .280$.

Univariate analyses showed that cuddliness was a significant predictor of anger frequency ($B = -3.475$, $p = .022$), with higher cuddliness associated with less frequent anger. No other significant effects were observed. These results were checked with bootstrap analysis. After conducting a bootstrapping analysis using the standard value of 1000 samples, the relationship between cuddliness and anger was no longer significant (bootstrap $p = .125$). The 95% bootstrap confidence interval for the parameter estimate ranged from -6.016 to 1.470, suggesting substantial variability and including zero, indicating no significant effect.

Discussion

This study aimed to explore the relationships between internal and external toddler individual differences and the frequency of facial emotional expressions (FEE) they see, with a particular focus on sibling status, longitudinal and concurrent temperament traits.

Understanding factors shaping the variability of the toddler visual input of FEE could be an essential first step in building models of how diversity in socioemotional developmental trajectories of FEE processing emerges. There was great variation in the frequencies with which different FEE were present within the visual input of toddlers. Not only this, but this variation was also related to some degree with both external and internal factors previously indicated as being significant in socioemotional development.

While we had hypothesised that toddlers with siblings would see more negative, threat-related FEE compared to those without siblings, we did not find a difference in the percentages of angry, fearful or disgusted faces present within the views of toddlers with and without siblings. However, toddlers without siblings observed a significantly higher percentage of happy faces than those with siblings. Though not a direct confirmation of our hypothesis, it perhaps reflects a similar underlying principle. We expected toddlers with siblings to see more negative, threat-related FEE because child siblings do not have fully developed emotional regulation and have not fully internalised display rules (Campos et al., 1989; Grolnick et al., 1999; Saarni, 1979); the socially desirable happy FEE being less frequent in the views of toddlers with siblings could reflect a similar effect. So, while counter to our specific expectations, the broad concept of the

presence of siblings being related to the frequency of FEE that are moderated by display rules seems sustained.

Beyond the direct influence of the FEE of the child siblings themselves, it could be that the presence of child siblings changes the behaviour of parents and caregivers. There is some indirect support for the first idea of children with siblings perhaps receiving less positive attention, with parents of multiple children reporting higher stress levels than parents of only children (Fu et al., 2024). There is also evidence that parents of only one child are more attentive parents, at least in the context of school and learning (Wei et al., 2016).

However, to fully understand why sibling status appears to be related to the FEE input, future research will need to recruit gender balanced samples of toddlers and toddlers with a range of sibling age groups. In our sample, most children without siblings were female, which could have confounded this effect. Gender plays a role in socialisation and development, especially in the domain of socioemotional development (Brody, 2000). This may mean that gender confounds our comparison of FEE frequency in toddlers with and without siblings. While less likely given that we found no main effect of gender in our analysis of this potential confound, direct testing of a model encompassing both presence of child siblings and gender in a balanced sample is necessary to rule this out in future research. Not only this, but we did not have a sufficient sample size to compare toddlers with older and younger siblings, with our participants having a mix of both. It is likely that these sibling dynamics are unique and

separable, and so future research should seek to compare the impact of younger and older siblings on the FEE input in a gender balanced sample.

To assess the source of differences in FEE input for toddlers with and without siblings, future research should also code the sources of FEE— whether sibling faces or adult faces. This would allow for an understanding of the degree to which siblings directly shape differences in the socioemotional visual environment versus through changing parent behaviour.

We also predicted that there would be a longitudinal relationship between early temperament traits recorded at 13-months and the percentages of faces observed by toddlers with each FEE (hypothesis 2). The results partially supported this hypothesis. Negativity significantly predicted the frequency of neutral expressions, and this relationship remained significant after bootstrap analysis. In testing specific IBQR dimensions further to assess our hypothesis that some specific dimensions of temperament would be related to the FEE frequencies, we focused on dimensions loading on negativity. We found that higher falling reactivity predicted lower frequencies of neutral faces, and this relationship remained significant after bootstrapping. So, this specific dimension of longitudinal temperament was associated in some ways with the toddler FEE input. Specifically, a higher percentage of faces being expressive was associated with higher scores on falling reactivity.

We also hypothesised that the concurrently measured ECBQ factors would be associated with aspects of FEE frequency within the toddler view. The results did not support this

hypothesis, as none of the three ECBQ factors (surgency, negativity, regulation) significantly related to the combined dependent variables. In exploratory analysis, regulation initially appeared to predict surprise frequency, but this effect was not significant when bootstrapped. The specific dimensions loading on regulation were examined, and relationships between cuddliness and anger were not significant after bootstrapping. This speaks against the hypothesis that specific dimensions of concurrent temperament (specifically, we had hypothesised that shyness may be relevant) predict the degree to which toddlers see faces displaying the six basic FEE.

It was surprising that longitudinal temperament—measured using the IBQR only when participants were 13-months old— predicted aspects of the FEE input, but not concurrent temperament on the ECBQ. However, it is important to note that we did not measure visual input at the time that IBQR data was collected; it could be that IBQR temperament is not only related longitudinally to later toddler visual input, but that it is related concurrently to infant visual input. A relationship between temperament at 13-months and visual FEE input in toddlerhood may suggest that toddler temperament only has observable relationships with the socioemotional input over the long term rather than effects that are immediately observable.

Or it could perhaps be that parent variables directly formative of the later FEE environment shape the reported IBQR temperament for infants. So, it could be that variables specific to the individual parent shape the temperament that they perceive and report in their infant. There is evidence for this kind of phenomenon— for example, maternal mental health

and parental stress accounts for a significant portion of variation in perceived infant temperament (Mäntymaa et al., 2006). A variety of mental health conditions and variables have been associated with reduced or altered facial expressivity, including depression and stress (Gaebel and Wölwer, 2004; Mayo and Heilig, 2019). This could mean that parents who are experiencing differing levels of stress and mental illness report different levels of negativity for their infants and also express emotion differently. However, this would not explain why the effect is found for longitudinal but not concurrent negativity.

To better understand explanations, it is important to examine the specific IBQR relationships that were present—higher negativity significantly predicted a higher frequency of neutral expressions and higher falling reactivity predicted lower frequency of neutral faces. Falling reactivity refers to the rate at which an individual recovers from peak negative or positive arousal, and it loads negatively on negativity, so low falling reactivity is associated with high negativity (Putnam et al., 2001). So, lower regulation or higher levels of falling reactivity were associated with a greater percentage of expressive faces.

Falling reactivity is significant in acute and long-term developmental functioning in many ways. For example, higher falling reactivity has been concurrently associated with better quality mother-infant dyadic interactions (Rigato et al., 2024). Higher falling reactivity also plays a protective role in developing executive functions for infants raised in adverse socioemotional environments (Ertekin et al., 2021). Lower falling reactivity is associated with risk of externalising behavioural problems in preschoolers (Garstein et al., 2012). So, overall, higher

falling reactivity is associated with healthier child-caregiver interactions and lower risk of adverse developmental outcomes, particularly in the domain of behavioural and executive functioning.

There are several ways in which emotional negativity as a whole may directly impact on parent-child interactions and therefore FEE frequency in the toddler view. One possible explanation for this could be that early temperament shapes the infant-caregiver relationship, which shapes the later FEE input. There is some evidence for this—for example, infant temperament, especially negativity, is predictive of maternal stress within the first year of life (Oddi et al., 2013). Higher levels of maternal stress are associated with poorer mother-infant bonds (Lutkiewicz et al., 2020), which are in turn associated with poorer socioemotional developmental outcomes (Behrendt et al., 2019). So, it could be that here, infants who score higher on metrics of negativity on the IBQR tended to have more stressed caregivers and develop poorer caregiver-infant relationships. This could then lead to a more limited range of expressivity from caregivers as a result of either reduced positive affect or increased negative affect that is down-regulated—parents perceive negative FEE, particularly directed at their children, to be non-normative and downregulate them (Lin et al., 2021; Malatesta and Haviland, 1982). Indeed, there is prior lab-based evidence for parents with depressive symptoms employing more neutral and less positive FEE in interactions with infants (Aktar et al., 2017), suggesting that negative emotions on the part of the parent are often expressed as more neutral expressions rather than outright negative expressions.

These more neutral FEE responses from the parent can be stressful for the infant or toddler. An extreme example of this can be seen in the still-face paradigm (Braungart-Rieker et al., 1998), but in the long-term, lower parental responsiveness (which is associated with higher parental stress) is associated with increased child behavioural problems and poorer cognitive and prosocial outcomes (Ward and Lee, 2020). So, the FEE input could be a mediating factor between early infant negativity along with parental stress and later socioemotional disorders and deficits.

The Implications for Differences in Socioemotional Processing Development

While we cannot directly say what the significance of these individual variations in FEE frequency for long-term socioemotional development is, one possibility is that these differences in the environmental input shape later socioemotional development. As discussed, there is indirect evidence for the statistics of the FEE visual input shaping development to be seen in the ORE (Elfenbein and Ambady, 2002; Yan et al., 2016), children who have been neglected or abused (Doretto and Scivoletto, 2018; Pollak and Tolley-Shell, 2003) and visual deprivation (Gao et al., 2013). These differences in FEE processing are often highly relevant in characterising disorders and determining social functioning. Some of the group differences in socioemotional development observed within the research could be related to the effects we have observed here.

Firstly, the differences in the frequency of FEE in the views of toddlers with and without child siblings that we observed could potentially explain differences in social development

between these populations— for example, the superior theory of mind skills of children with siblings (Zhang et al., 2021). However, given that there is no research to suggest an overall difference in FEE processing according to this characteristic, it may be that other aspects of the socioemotional input provided by child siblings, such as more opportunities for play, are more formative of these differences (McAlister and Peterson, 2006).

However, differences in FEE processing are more central to many disorders with noted associations with temperament. In mood disorders such as major depressive disorder, individuals often exhibit impaired accuracy in judging FEE, frequently displaying a consistent negativity bias (Bourke et al., 2010). These kinds of cognitive biases provide a direct causal underpinning for the mood and affect symptoms experienced within the context of these disorders (Zghoul et al., 2019), meaning that understanding how differences in the ability to accurately interpret FEE cues emerge could be essential in understanding the aetiology of and possible treatments for mood disorders. Further research should examine how differences in early FEE visual input and their relationships with aspects of early temperament, specifically negativity, may fit into models of these mood disorders. There is indeed prior research to suggest that negativity is an aspect of temperament that may be a risk factor for the development of mood disorders (Clark et al., 1994). So, understanding how and why negativity is related to the FEE input will be essential in building comprehensive models of risk for these disorders.

Further, differences in FEE processing are clinically significant for disorders and populations with antisocial behaviour as a feature. For instance, sex offenders have shown significant impairments in identifying emotions such as anger, disgust, surprise, and fear (Gery et al., 2009). These differences in FEE processing are likely to be at least in part explanatory of or related to their harmful behaviour, so understanding how they might come about and whether the early environmental input of FEE is essential.

Not only this but conditions like Autism with a social cognition deficit are associated with impaired FEE processing (Loth et al., 2018). Given that Autism is associated with differences in early temperament, including negativity (Macari et al., 2017; Narvekar et al., 2022), one possible causal mechanism could be early temperamental differences shaping the visual FEE input (at least in part), and these differences in turn shaping later FEE processing proficiency. To fully assess these hypotheses, it will first be necessary to build more comprehensive models of how and why relationships between temperament and the visual input come about. It would also be necessary to directly examine samples likely to include toddlers who go on to be diagnosed with neurodevelopmental disorders.

Overall Conclusions and Analysis

These findings are the first to shed some light on the great degree of variability in the toddler visual input of facial emotional expressions and the factors that may shape this variability. Overall, the findings suggest a role for the presence of child siblings and longitudinal temperament. We find that toddlers without child siblings see proportionally more happy faces

than toddlers with siblings. We also find that higher negativity at 13-months is associated longitudinally with a higher percentage of neutral faces in the toddler view, seemingly driven by falling reactivity.

Further research with larger sample sizes is needed to confirm and extend these findings, particularly how and why sibling status and longitudinal (but not concurrent) temperament could be shaping the emotional expressions that toddlers see. It would be particularly interesting to extend this research to samples at-risk of neurodevelopmental disorders. It would also be essential to code the identities of the faces displaying FEE within the toddler view to understand the influence of siblings. As with all naturalistic developmental research, pragmatic constraints such as the high participant demands, and the time-consuming data coding process limited the number of participants and consequently the statistical power here. While the toddler-perspective ($N = 26$) data here may represent a limited number of individuals, it represents hundreds of hours of video footage that can take many hours to code using traditional methods, but limited statistical power, particularly for between-subjects comparisons on variables like sibling status. That said, some of the relationships between elements of the FEE input and temperament here were relatively robust, being significant even after conservative hypothesis testing. This forms the groundwork for future research with wider and more diverse samples to examine these relationships in greater detail.

Characterising the Naturalistic Toddler Visual Input of Social Stimuli: General Discussion

Emma J Jackson

Summary of Findings

This thesis sought to address broadly the question of what visual information about socioemotional stimuli-- particularly, faces, bodies and facial emotional expressions-- is in the view of toddlers as they navigate naturalistic home environments. This is a first, critical step in understanding what information toddlers may extract from their environments, which is essential in understanding how the environment may or may not drive development of face, FEE and body processing. This is also a necessary step in building models of how variability in this development may emerge. Prior research on what toddlers actually see and how often is very limited, with the only available research suggesting that faces are infrequent in the toddler view (making up about 7.5% to 8.8% of toddler perspective scenes) relative to bodies and relative to their frequency in the visual input of younger infants (Fausey, Jayaraman and Smith, 2016; Jayaraman et al. 2016, 2017; Sugden et al., 2013) . Research on adult visual input was also very limited, with the only available information being on frequency of faces in the adult view with data collapsed across a variety of social environments (Oruc et al., 2019). This meant that it was unclear whether infrequent faces were a temporary and unique feature of toddlerhood, or closer to a developmental milestone. Considering this, we have at several points examined the degree to which adult caregivers see faces, bodies and FEE as a point of comparison. Further, we provide an example of a methodology for coding large, rich video datasets that leverages computer vision models.

The five chapters in this thesis have characterised the naturalistic frequency of faces and bodies in the view of toddlers and their caregivers (chapter 1); how individual differences in the

frequency of faces and bodies in the toddler view may be related to the presence of child siblings and temperament (chapter 2); how typical adults actually perceive and categorise naturalistic facial movements (chapter 3); the naturalistic frequency of FEE in the view of toddlers and their caregivers (chapter 4); and how individual differences in the frequency of FEE in the toddler view may relate to the presence of child siblings and temperament (chapter 5).

We coded rich, highly naturalistic video data, utilising egocentric recordings alongside computer vision to code the visual input of caregivers and toddlers. We analysed 259 hours of egocentric footage (approximately 14 million frames) from 26 toddlers, and (for 24 participants) their corresponding caregivers. Presence of child siblings was recorded using parental report, as was temperament: parents completed the IBQR at 13-months and the ECBQ concurrently. Computer vision models were adapted and validated for this egocentric footage, in part with in-lab face and body coding and in part using the data collected for study 3. This represents a novel and effective use of computer sciences to augment coding of rich video data.

Chapter 1

Chapter 1 aimed to quantify the frequency of faces and bodies in the egocentric view of children in the second year of life and their caregivers within the home. Prior research suggests that faces specifically are infrequent in the view of toddlers, but not younger infants, with around 8% of the toddler visual input containing faces, as compared to around 25% for newborns (Fausey et al., 2016; Jayaraman et al., 2015; Jayaraman et al., 2017; Sugden et al., 2014). There is no general age-related decline of people within the visual input from birth to 24-

months, who are present in 40% of scenes (Jayaraman et al., 2017), and hands increase in frequency over this time period (Fausey et al., 2016). Prior research on the adult “face diet” would suggest that this is unique to toddlers, with 22.5% of the typical adult visual input containing faces (Oruc et al., 2019).

However, this prior research was limited in several ways. Firstly, there is the lack of a quantification of the adult caregiver view within the same environment, which means that it is impossible to know whether infrequent faces as compared to bodies are a unique feature of the toddler's visual input or the developmental endpoint. Secondly, there is the limited number of walking toddlers being examined in each sample. Then, there is coding being limited to a subsample of frames. Finally, there is a definition of face and body that excludes media, self or mirror presentations. In part, these limitations are related to the time-consuming nature of the manual coding often used. While Oruc et al. (2019) and Frank et al. (2013) used automated detection coupled with manual checking, and Sugden et al. (2014) used second-by-second manual coding, the works of Jayaraman and Fausey all utilise the same video corpus and the same coding methodology: sampling and coding a frame every 5-seconds. While an impressive corpus of videos, this reflects the time-consuming nature of manual video coding. So, we coded frame-by-frame in a wider sample of walking toddlers with a wider definition of faces and bodies.

Across all toddlers, 7.98% of frames ($SD = 4.75$) contained faces while 53.73% of frames ($SD = 6.70$) contained bodies. For caregivers, 13.81 % ($SD = 6.92$) of caregiver-perspective frames

contained faces, and 72.88% (SD = 9.06) of frames contained bodies. Faces were less frequent than bodies in both perspectives, and both faces and bodies were less frequent in the toddler than caregiver views. Not only this, the frequency of faces in the caregiver view was lower than the face frequency in the adult view found by Oruc et al. (2019). In conclusion, faces really are infrequent in the toddler view as compared to other both bodies and the face frequency in the caregiver view. However, the home environment is perhaps slightly less face-dense for all individuals than environments outside of the home might be.

Chapter 2

Chapter 2 aimed to examine potential sources, or at least correlates, of the wide variation in face and body frequency within the toddler view noted in chapter 1. We focused on the external variable of whether the toddler had a sibling and the internal variable of toddler temperament, measured both concurrent to recording and at an earlier age.

We expected that toddlers with siblings would see more faces and bodies, partially as a function of an increased number of individuals within the family home, and partially because of sibling faces being potentially more pragmatically accessible than adult faces. Prior lab-based research on the visual input of toddlers has suggested that pragmatic constraints arising from height may limit visual access to the upper visual field, which is where the faces of most adults can be found (Frank et al., 2013; Kretch, Franchak and Adolph, 2014). So, shorter siblings may offer easier visual access.

We also tested the proposal that factors internal to the toddler may shape the statistics of the visual input through their actions and motivations (Smith et al., 2018), as well as possibly through the social relationships around them. Specifically, we were interested in the internal variable of temperament, given prior research suggesting various associations between temperament and socioemotional development (e.g., Lo and Marr, 2022; McAlister and Peterson, 2006). We anticipated that longitudinal temperament as measured at 13-months and concurrent temperament would be related to the frequency of faces and bodies in the toddler view. The rationale behind this hypothesis was that firstly, temperament has been shown to shape concurrent responses to stimuli, including faces (Rennels et al., 2020). Toddlers can shape their own visual input, both through looking behaviour and through shaping the behaviour of others (Smetana, 1989), so concurrent temperament may moderate the visual input through these pathways, with toddlers with varied temperaments being more or less motivated to attend to faces and bodies and eliciting more or fewer instances of this input from others. This idea of toddlers shaping their own visual “curriculum” has been suggested already, with the proposal being that the motivations, pre-existing knowledge and developmental stage of the toddler shape their navigation of the environment and therefore input (Smith et al., 2018).

Secondly, temperament has been associated with long-term aspects of the parent-child relationship (Adreadakis et al., 2020; Liu et al., 2020), which may mean that longitudinal temperament is related to face and body frequency through its effect on the social environment. Thirdly, dimensions of early temperament are related to aspects of later social proficiency and development such as empathy and theory of mind (MacGowan et al., 2022;

Mark et al., 2002; Mink et al., 2014; Wagers and Kiel, 2019; Young et al., 1999). Given that the deficits in social processing often comorbid with deficits in face processing are also related to temperamental differences (Pijl et al., 2019), temperament seemed a likely candidate for individual differences in this development.

While there was a gender confound on presence of siblings (most no-sibling children were female), face and body frequency did not appear to differ by gender or sibling presence.

Additionally, no robust relationships were found between any measure of temperament and the frequency of either faces or bodies. While this could be a product of the relatively small sample size that we test here, it could also be that the general frequency of faces and bodies does not meaningfully vary with the presence of child siblings or temperament— it could even be that only more fine-grained elements of the socioemotional visual input, such as FEE, are related to the individual differences that we examine.

Chapter 3

With the above results in mind, we were interested in examining specific aspects of the information available within the faces in the toddler view. However, we first needed ground truth data for our FEE models. In performing a review of prior research on real-world presentations of FEE, it became clear that there is surprisingly little research on this topic of how typical adults perceive and rate real-world FEE. Most prior efforts utilise artificial, posed, static examples of facial poses. There is some limited prior research using more naturalistic stimuli. For example, Cowen and Keltner (2020), in examining 1500 static face-body photographs

sampled from the internet, found 28 reliable emotion categories– and crucially, high levels of disagreement amongst participants that varied by FEE category. Given the wide array and variety of facial movements in our datasets, we determined that the 311 excerpts that we had selected for model validation could be used to examine adult categorisation and perception of real-world facial movements while collecting ground truth data.

We examined the ratings of 149 typical adult participants each assigned to a subset of stimuli; a total of 37-38 ratings was collected for each videoclip. We tested the effect of the FEE category of a facial movement and the age group of the individual producing the facial movement on level of disagreement and average intensity/confidence ratings. Our approach was bottom-up in that we did not decide which FEE was the “correct” answer for each clip ahead of time: participant consensus determined this.

We expected to replicate the ambiguity of Cowen and Keltner (2020), as well as the effect of the FEE category. We expected the age group of the video subject to have an effect, as adults tend to moderate their facial expressions, especially when directed at children (Chong et al., 2003) while infants and young children are more dysregulated (Grolnick et al., 1999). Not only this, but non-parents (who made up most of our sample) have more difficulty categorising infant facial expressions (Maack et al., 2017).

Largely, we confirmed our hypotheses. While each one of the six “basic” emotion expressions and neutral were present within our stimulus set-- though, despite all efforts to find varied

stimuli, fear and disgust were still rare-- average agreement for a stimulus was only 69.41% (SD = 21.73). Average agreement varied by emotion, being highest for happiness and lowest for fear. When we examined emotion valence categories (positive, neutral, negative), agreement varied by emotion and age-group, with agreement being highest for negative infant emotion, but generally higher for positive emotions for other stimulus subject age groups. Confidence and intensity varied by final clip FEE valence and age group, following similar patterns to agreement. So spontaneous FEE recorded in naturalistic settings are ambiguous, and perceptions of FEE vary significantly by the valence category of observed FEE and the age group exhibiting the FEE. Negative emotions, especially fear, are more ambiguous and associated with lower confidence ratings than positive FEE like happiness for adults and children, but overall agreement is highest for negative infant emotions, which here consisted only of sadness. Another output of this data was the validation of our FEE model (for details, see the methods section), which in turn allowed for the data analysis of chapters 4 and 5.

Chapter 4

Chapter 4 aimed to examine the naturalistic visual input of adults and toddlers through quantifying the frequency of facial emotional expressions in their egocentric view using coding by the previously validated model. This endeavour is especially novel, given that very limited attempts have been made to assess the degree to which people of any age actually see FEE. While there is some prior self-report data to indicate that adults see happiness frequently, and fear especially infrequently (Calvo et al. 2014), there has been no actual quantification of this.

We hypothesised that while both positive and negative FEE would be present in the view of the toddlers, positive and neutral FEE would be more frequent. We also hypothesised that negative FEE would be more frequent in the caregiver view. These hypotheses were partially based on the preliminary prior research of Calvo et al. (2014) but were mostly based on indirect evidence. While toddlers are capable of committing transgressions that elicit admonitions (and possibly negative FEE) (Smetana, 1989), parents tend to produce positive child-directed expressions (Chong et al., 2003; Malatesta and Haviland, 1982). It does not seem likely that toddlers, who have not yet mastered emotional regulation and inhibition (Grolnick et al., 1999) would similarly moderate their negative emotional displays. Given that within the home environment toddlers are likely a major source of FEE within the caregiver view, this may mean that negative FEE are more common in the caregiver view.

Broadly, these hypotheses were confirmed, though the percentage of faces receiving each FEE label in the views of toddlers and their caregivers was markedly similar. Negative, threat-related FEE such as fear, anger and disgust were especially infrequent. Neutral and happy faces were most frequent in the views of both toddlers and their caregivers, though were slightly more frequent as a percentage of frames (rather than faces) in the caregiver view.

Chapter 5

As with faces and bodies, when analysing the data for chapter 4, we noticed marked individual differences in the degree to which different FEE were in view. In chapter 5, we sought to examine sources of correlates of this. While there was again no prior research on this topic,

there was research to suggest that temperament and the presence of siblings may be important in socioemotional development and FEE processing.

As previously mentioned, dimensions of early temperament are related to aspects of later social proficiency and development such as empathy and theory of mind (MacGowan et al., 2022; Mark et al., 2002; Mink et al., 2014; Wagers and Kiel, 2019; Young et al., 1999). Temperament is also associated with risk of ASD diagnosis (Chetcuti et al., 2020; Macari et al., 2017; Narvekar et al., 2022), which entails marked deficits in FEE processing, especially for negative and threat-related FEE. So, temperament could shape socioemotional development. It could be that this is because of or related to variation in the visual input. Toddlers are not necessarily passive receivers of information but may directly and indirectly affect the statistics of their visual input (Smith et al., 2018). Indeed, there is research to suggest a direct link between aspects of temperament such as negativity and regulation and aspects of concurrent FEE processing such as gaze behaviour and EEG response (Fu et al., 2020; Martinos et al. 2012; Ravicz et al., 2015). These differences in processing could impact on the visual input.

In terms of factors external to the child, prior research suggests that the presence of child siblings could be formative. There is some evidence to suggest that child siblings enhance or modify socioemotional development (Rodrigues and Abhyankar, 2012; Sang and Nelson, 2017; Smorti and Ponti, 2018; Zhang et al., 2021). Child siblings certainly are the source of a wide range and variety of peer interactions within the home environment (Howe and Recchia, 2005). Pragmatically speaking, toddlers may struggle to visually access the faces of standing adults

because of height differences (Frank et al., 2013; Kretch, Franchak and Adolph, 2014). Not only this, but child siblings, whose abilities to emotionally self-regulate are still emerging (Grolnick et al., 1999). So, siblings may be a source of more frequent, intense and more negative FEE.

We hypothesised that toddlers with child siblings would see more negative, threat-related FEE than toddlers without siblings, and that aspects of early temperament would be longitudinally related to the FEE that toddlers see. We also hypothesised that aspects of concurrent temperament as measured using the ECBQ would be related to the FEE that toddlers see. We found that toddlers without siblings saw proportionally more happy faces. Further, emotional negativity, specifically falling reactivity, was longitudinally related to FEE frequency in the toddler view. Toddlers scoring higher on emotional negativity and lower on falling reactivity more frequently saw neutral faces. We did not find a significant relationship between any aspect of concurrent temperament and the frequency of FEE in the toddler view.

Summary

Overall, the work presented here demonstrates that faces are infrequent in the naturalistic view of toddlers, even as compared to the frequency of bodies in their view, and the face frequency in the views of caregivers within the same environment. This would imply that toddlers rarely see faces. The frequency of faces and bodies in the toddler view does not differ with longitudinal or concurrent temperament, or whether toddlers have child siblings. In terms of FEE, we find that real-world facial movements are often much more ambiguous for typical human observers than research with prototypical FEE would suggest, especially for threat-

related FEE. Further, negative FEE are very infrequent within the views of both toddlers and their caregivers, with positive FEE like happiness and the neutral making up the vast majority of faces within both. Neutral and happy faces are in the view of caregivers slightly more frequently than the views of toddlers when not controlling for face frequency. Finally, toddlers without siblings see more happy faces, and elements of negativity, specifically falling reactivity, are associated with the frequency of neutral faces. Higher falling reactivity is associated with a higher frequency of expressive faces.

Implications of These Findings

Theoretical Implications

Collectively, these findings suggest that the visual input of toddlers contains very little direct information about faces and facial emotional expressions. Particularly, there is very little visual information available to toddlers about negative FEE like fear and disgust. Non-neutral FEE overall were infrequent in the toddler view. This would seem strange, given that socioemotional development continues at pace well into childhood and adolescence (Batty and Taylor, 2006; Vetter et al., 2013) and given that the environmental input has been proposed to be an essential component of socioemotional development (LoBue and Ogren, 2022).

Given ongoing emergence of FEE processing into adolescence (Herba and Phillips, 2004) alongside many other social cognition milestones that we would expect to depend on learning from faces and FEE within the environment— for example, explicit ability to pass false belief tasks does not emerge until about age 5-years (Astington, and Barriault, 2001)-- it seems

strange that faces and FEE are so infrequent in the view of toddlers. Faces are rich sources of information on the internal worlds of others and the wider world around the toddler (Astor et al., 2021; Hessels, 2020). Either very little visual environmental input is necessary for development of these cognitive processes, or affective and intentional information is being extracted from elsewhere (perhaps bodies) to provide developmental substrate for these abilities, or toddlers are highly efficient in their extraction of information from faces.

The relative infrequency of faces could mean that toddlers learn less from faces than young infants, or it could be that toddlers can more efficiently process infrequent faces than young infants would be able to and so still extract the same degree of information. There is some evidence for this— young children and toddlers tend to process faces inefficiently and based on features, as compared to efficient adult configural processing (Diamond and Carey, 1977). Further, there is data to directly suggest increases in face processing speed over infancy (Rose et al., 2002). So, it could be that toddlers can extract the same amount of information from these brief face presentations that much younger infants would be able to over prolonged viewing.

It could also be that toddlers, who see bodies frequently, have a greater tendency to rely on other, multimodal cues or contextual information to extract variables like identity, affective state and more. Indeed, even for adults, when there is a mismatch between the facial and bodily emotional cues, it is primarily cues from the body that determine final judgement (Meeren et al., 2005). There is evidence to indicate that infants rely on vocal emotion cues more

than FEE cues and use the former to shape categorization of the latter (Caron et al., 1988). This may then be used to shape and scaffold learning about faces and FEE from more limited visual exposure. So, it could be that when toddlers direct their own input, it is not faces that form the primary source of information about the socioemotional environment. So, if this were true, only infrequent and brief glances towards faces would be necessary for extracting the information needed for typical development and navigation of the social environment.

Toddlerhood may also be a period of increased learning about aspects of the visual environment other than faces, such as these frequent bodies. Visual deprivation early in infancy does not have the same long-term detriment for body processing as it does for face processing (Hadad, Maurer and Lewis, 2012). Not only this, but many skills that emerge very early for face processing take somewhat longer to emerge for body processing. For example, evidence for holistic processing of faces can be seen from approximately 3-months (de Haan et al., 2001; Quinn and Tanaka, 2009; Turati et al., 2010) and equivalent skills are not somewhat established until 14- to 18-months (Gillmeister et al., 2019; Slaughter et al., 2002; 2004; Turati et al., 2004). Given that (in line with prior research, Fausey et al., 2016; Jayaraman et al., 2015, 2017) we do not find generally infrequent people, but specifically infrequent faces, it could be that toddlers are motivated to attend more frequently to aspects of the environment such as these as they learn more about them.

There had been hypotheses that perhaps the later emergence of FEE processing was driven by older infants being exposed to more instances of these expressions, as they are more

capable of committing transgressions (Reid et al., 1994; Smetana, 1984, 1989). While this could well be true in a literal sense, it is certainly not true that the environments of toddlers are rich in these negative expressions. Given the strong tendency of adults to moderate negative affect when interacting with infants (Chong et al., 2003; Kim and Johnson, 2014), it seems unlikely that younger infants would see negative, threat-related emotions more frequently than the toddlers that we examined. So, it seems that infants and toddlers either learn about negative FEE from very scant input or have some in-built knowledge that emerges with little environmental input over time. There is some research to indicate that this may be the case—for example, skin conductance research has previously indicated that infants as young as 3- to 4-months subconsciously process happy and angry faces, and so it could be that early processing of FEE is more sophisticated with more innate knowledge present early in infancy than looking-behaviour tasks may indicate (Nava et al., 2016).

This would tie in with theories that maturational processes drive the development of many socioemotional skills, including elements of this negative, threat-related FEE processing development (Winnicott, 2018). Given continued development in the context of such scarce input, this could suggest that models implying reliance on environmental visual input for the development of such skills (e.g., CONSPEC/CONLERN, Morton and Johnson 1991) are not accurate. Perhaps, particularly for the especially scarce negative and threat-related FEE, innate and maturational processes are more influential in the emergence of new skills. There have indeed been prior suggestions that processing of threat-related emotions is not only more subcortically mediated than other FEE processing, but that it is driven to a larger degree by

innate processes (Adolphs, 2008; Kawai et al., 2016). So, maturational processes could complement the scarce environmental input, especially for the most infrequent FEE.

An important overall theme in this thesis has been this distinction between positive and negative FEE. As chapter 3 demonstrates, adult processing of FEE is not monolith—real-world examples of negative FEE, especially fearful FEE, are far more ambiguous for typical adults than expressions like happiness. Further, chapter 4 would suggest that this hierarchy of difficulty is reflected in the statistics of the visual input—most faces in the toddler and caregivers' views are neutral or happy, with threat-related FEE being particularly infrequent. This would reflect patterns of development, with the ability to discriminate between threat-related FEE emerging later in development (Hoehl and Striano, 2008) than happiness (Nelson, 1987) and deficits in processing in clinical populations often observed for these expressions (Farran et al., 2011; Wallace et al., 2008). Prior research has suggested that happiness is “special”, with self-report measures suggesting that adults see happiness most often (Calvo et al., 2014), adults showing enhanced perceptual sensitivity for happiness as compared to other FEE (Maher et al., 2014). If the environment is important for this developmental timeline, the later emergence and more challenging nature of negative FEE processing for even typical adults as compared to happiness could be related to its infrequent occurrence in the visual environment. However, it is important to note that this idea of more challenging negative FEE and less challenging positive FEE is nuanced and often restricted to specific tasks or expressions—as evidenced by the fact that we observed less ambiguity for ratings of infant negative FEE specifically, which consisted mostly of sadness.

The potential role for the individual differences in the presence of siblings and toddler temperament for shaping the visual input is also important. We found no relationship between these factors and the frequencies of faces and bodies in the toddler view, or between concurrent temperament as a whole and the visual input but found that the presence of child siblings and aspects of longitudinal negativity were associated with the frequencies of different FEE in the toddler view. Specifically, in chapter 5, we found that toddlers without siblings saw more frequent happiness, and that higher negativity (lower falling reactivity) at 13-months was associated with a higher percentage of the faces that the toddler saw being neutral in toddlerhood.

The role of siblings in this is not clear: it could be that the siblings themselves alter the statistics of the FEE input, or it could be that parents of multiple children differ in their behaviour. Child siblings are likely to have not fully internalised display rules that dictate happiness to be normative (Malatesta and Haviland, 1982), which may mean that they are a source of more frequent non-happy FEE in children with siblings. It could also be that parents of multiple children have less time to spend with each of these children and fewer emotional resources to expend on each of them (Öberg, 2017), resulting in fewer parent-original happy FEE in the views of toddlers without siblings. To fully assess these hypotheses, we would need to characterise the identities of the faces exhibiting different FEE within the toddler view.

The relationship between falling reactivity and the frequency of expressions could be explained in several ways. Higher negativity has been associated with externalising behavioural problems (Sanson et al., 2004) and Autism (Chetcuti et al., 2021). Further, higher falling reactivity specifically is associated with higher quality affective mutuality between mothers and infants (Rigato et al., 2024). So, these correlations between early temperament and later FEE input could reflect differences between toddlers with higher-risk temperamental profiles for social disorders early in life. They could also reflect the impact that early infant temperament has had on the caregiver-toddler relationship. The fact that the relationships here are longitudinal, and that we do not see a relationship between concurrent negativity and the statistics of the FEE visual input, could suggest more longitudinal and relationship-based mechanisms for the relationship between temperament and FEE input. However, more research examining the specific source of these differences in FEE input is necessary before these ideas can be fully assessed— for example, assessment of the types of interactions and situations in which these FEE occur, and assessments of how the toddler-caregiver relationship quality is associated with the statistics of the FEE input.

These papers are unique in that this is a richer quantification of the toddler socioemotional input than has previously been undertaken (or even possible), with a more representative sample of walking toddlers, and more aspects of the visual input extracted. They are also unique in that they quantify the adult (and specifically caregiver) visual input with respect to socio-emotional stimuli comprehensively for the first time. Largely, this has been

possible because of the validation and use of computer vision in coding this data, which is a novel application of this technology in developmental psychology.

Not only this, but here we present the first examination of adult perception and categorization of truly naturalistic, dynamic facial movements for emotional valence. Given how difficult a wide variety of truly naturalistic expressions are to obtain, our ability to extract and use a corpus of 311 examples allowed an insight into an ability that has been previously assessed only by using static images, which more often than not are also artificially posed.

Methodological and Practical Implications

This was the first research to use egocentric head camera recordings and automatic, computer vision coding in tandem. In future, these methodologies could be used in much wider-scale research, maybe in comparing visual input cross-culturally or in comparing different demographics—potentially, those with high and low risk for socioemotional disorders like ASD. The dataset of 311 facial movements that we generated ground truth data for could be essential in validating computer vision models in future. These findings may be useful in developing social AI, as well as in understanding what environmental input is necessary for typical development. With further research, they can also help us to understand what exactly may underpin diversity in socioemotional stimulus and general social processing development, through building a more thorough understanding of how differences in the input relate to later differences in social functioning.

Limitations and Future Directions for Research

While this research was innovative in the methodologies used and allowed for coding and analysis of a much greater number of participants and greater amount of data than previous, similar research, the main sample examined here was still limited in number (N=26). While this translates into multiple millions of frames coded, it does mean that our ability to interpret individual differences was potentially limited in terms of statistical power. Future research could focus on recruiting greater numbers of participants to specifically examine individual differences. Indeed, there is currently a wide project that builds on the methodologies here in process that addresses exactly these problems, alongside using specialised cameras with a wider field of view and sensors to measure physiological input (Geangu et al., 2023).

A second limitation is that while our face detection algorithm had acceptable precision and recall, performance was poorer than it would have been on typical controlled, lab-based recordings because of several factors: lighting, video quality, motion blur and more. While many of these would apply to human coding to the same, if not greater, a degree, one area of performance that was particularly difficult for algorithmic coding was identification of very partial face presentations (e.g., just a mouth) that a human would know to be a face because of contextual knowledge and greater face expertise. Future research would do well to test methodologies for tackling these partial face presentations.

Future research should further examine sources of individual differences in the visual input using a wider, more varied or cross-cultural sample. It would be especially useful to have an idea how differences in input relate longitudinally to social skills and proficiency. This would help to build a better picture of what causal role the environmental input of FEE may play and be essential in addressing the question of whether innate differences directly drive later socioemotional competency, or whether the environment is a key mediator of this relationship—i.e., whether the emergence of FEE processing and other socioemotional skills is driven more by maturation or the environment (Winnicott, 2018). This research could focus on toddlers who are and are not at risk of socioemotional processing disorders and may examine especially the role of negativity and specific dimensions of this factor in predicting the visual input and development of FEE processing abilities.

It would also be essential to establish just how toddlers learn from such infrequent examples of threat-related FEE through examining differences in how these expressions are attended to within the visual input—for example, perhaps though it is very infrequent, fear is highly salient, meaning that presentations of it in the natural visual input are more influential. This is precisely what will be attempted with the ongoing work of Geangu et al. (2023). This data will help to inform the question of whether infant and toddler attentional behaviour more covert than explicit head movements may be formative in weighting the environmental visual input. There is some prior research to suggest that this may be the case, particularly when it comes to social cognition (Terrizzi and Beier, 2016). So, the salience of negative, threat-related FEE may offset their infrequency.

It could also be that the multimodal nature of these infrequent threat-related FEE is significant in offsetting this infrequency. In this thesis, only visual information available in the environment (not auditory information, such as vocalisations) is examined, and affective information available from body posture is not quantified. There is indeed prior research to suggest that vocal expressions of negative, threat-related emotions are selectively deployed by caregivers in their interactions with infants (Dahl et al., 2014). Further, emotional expressions are processed multimodally, with information from bodies and voices being used alongside information from faces (Grossman, 2010; Meeren et al., 2005). These additional modalities could therefore provide more frequent sources of negative, threat related affective information to complement infrequent equivalent FEE. They could also act as cues to increase the saliency of threat-related FEE. For example, infrequent facial expressions of anger may be more salient, and toddlers may therefore learn more from them, if they are accompanied by loud, angry voices.

Conclusions

In conclusion, faces are not particularly frequent in the visual input of adult caregivers but are rarer in the visual input of toddlers as compared to bodies. Not only this, but facial movements (or FEE) perceived as neutral or happy make up the overwhelming majority of the facial displays for both toddlers and their caregivers. Similarly, adult ratings of the real facial movements that occur in the visual input naturalistically show much higher agreement and confidence for positive FEE like happiness than for threat-related FEE like fear. Face and FEE

processing abilities in toddlerhood seem to emerge in the context of very scarce environmental visual information, especially on negative, threat-related FEE. So, either environmental information on negative FEE is extracted very efficiently, is extracted from sources outside of the visual input, or is not essential for the development of FEE processing. It could also be that a combination of these things is true: for instance, negative FEE could be salient and tend to co-occur with multimodal input. Finally, longitudinal negativity and presence of child siblings is associated with differences in the frequencies of FEE within the toddler visual input.

The research presented in this thesis has made unique contributions to the field in that face frequency in that the visual input of toddlers has rarely been quantified before, and when it has, with smaller samples. Building on Fausey et al. (2016) and Jayaraman et al. (2015, 2017), we characterise the visual input in infancy but focus exclusively on toddlers with a wider sample at this age. We also build on the work of Oruc et al. (2019) in characterising the visual input of a new subgroup of adults—caregivers within the home. This work is also unique in that it is the first quantification of the FEE visual input for any age group. We leveraged a combination of novel applications of computer vision methodologies and adult lab ratings of truly naturalistic FEE to extract information on the FEE that toddlers and their caregivers see. We build on some prior research using self-report by Calvo et al. (2014), but the FEE visual input has never been quantified before using direct measurement in any context.

Together with these prior assessments of the face and body visual input, the presented findings indicate that faces are especially scarce as compared to bodies in the view of toddlers

specifically, and this does not extend to adult caregivers within the home. Future research should assess precisely what drives this infrequency, possibly through assessing the contexts in which toddlers access faces. It should also assess how toddlers use information from these infrequent faces. The fact that neither the presence of siblings nor temperament was related to face or body frequency in the toddler view leaves variation in this frequency opaque but suggests that simply having child siblings (who are likely to be of similar height) in the home is not sufficient to offset infrequent faces, so perhaps simple pragmatic barriers do not entirely drive infrequent faces in toddlerhood.

The presented findings also show that negative FEE are especially infrequent in the faces of toddlers and their caregivers, and that both internal and external individual differences are associated with differences in the toddler FEE input. Higher toddler negativity is associated with a higher incidence of neutral FEE, and toddlers without siblings see more frequent happiness. To fully understand the significance of this, future research should extend these findings to clinical populations and attempt to assess the individuals in the toddler view who exhibit different FEE and the contexts in which they display them.

The dissociable nature of performance on tasks employing positive and negative, threat-related FEE in adults could be related in part to their infrequency in the views of both toddlers and adults. Future research should also assess just how toddlers can learn from such infrequent negative, threat-related FEE-- is the environmental input sufficient to drive this, or are maturational processes necessary to augment such scarce visual input?

References

- Achaibou, A., Pourtois, G., Schwartz, S., & Vuilleumier, P. (2008). Simultaneous recording of EEG and facial muscle reactions during spontaneous emotional mimicry. *Neuropsychologia*, 46(4), 1104-1113.
- Achtergarde, S., Postert, C., Wessing, I., Romer, G., & Müller, J. M. (2015). Parenting and child mental health: Influences of parent personality, child temperament, and their interaction. *The Family Journal*, 23(2), 167-179.
- Adams, R. B., Ambady, N., Macrae, C. N., & Kleck, R. E. (2006). Emotional expressions forecast approach-avoidance behavior. *Motivation and emotion*, 30, 177-186.
- 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.
- Adolph, K. E., & Robinson, S. R. (2015). Motor development. *Handbook of child psychology and developmental science*, 1-45.
- Adolphs, R. (2008). Fear, faces, and the human amygdala. *Current opinion in neurobiology*, 18(2), 166-172.
- Adolphs, R., Gosselin, F., Buchanan, T. W., Tranel, D., Schyns, P., & Damasio, A. R. (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature*, 433(7021), 68-72.
- Ainsworth, M. D. S. (1969). Object relations, dependency, and attachment: A theoretical review of the infant-mother relationship. *Child development*, 969-1025.

Aktar, E., Colonnese, C., de Vente, W., Majdandžić, M., & Bögels, S. M. (2017). How do parents' depression and anxiety, and infants' negative temperament relate to parent–infant face-to-face interactions?. *Development and Psychopathology*, 29(3), 697-710.

Ali, L., & Davis, J. P. (2020). The positive relationship between face identification and facial emotion perception ability extends to super-recognisers.

Als, H. (1977). The newborn communicates. *Journal of communication*, 27(2), 66-73.

Amso, D., & Kirkham, N. (2021). A multiple - memory systems framework for examining attention and memory interactions in infancy. *Child Development Perspectives*, 15(2), 132-138.

Andreadakis, E., Laurin, J. C., Joussemet, M., & Mageau, G. A. (2020). Toddler temperament, parent stress, and autonomy support. *Journal of Child and Family Studies*, 29, 3029-3043.

Andric, S., Maric, N. P., Knezevic, G., Mihaljevic, M., Mirjanic, T., Velthorst, E., & van Os, J. (2016). Neuroticism and facial emotion recognition in healthy adults. *Early Intervention in Psychiatry*, 10(2), 160-164.

Anzures, G., Quinn, P. C., Pascalis, O., Slater, A. M., & Lee, K. (2010). Categorization, categorical perception, and asymmetry in infants' representation of face race. *Developmental science*, 13(4), 553-564.

Anzures, G., Quinn, P. C., Pascalis, O., Slater, A. M., Tanaka, J. W., & Lee, K. (2013). Developmental origins of the other-race effect. *Current directions in psychological science*, 22(3), 173-178.

- Armstrong, T., Bilsky, S. A., Zhao, M., & Olatunji, B. O. (2013). Dwelling on potential threat cues: An eye movement marker for combat - related PTSD. *Depression and anxiety*, 30(5), 497-502.
- Astington, J. W., & Barriault, T. (2001). Children's theory of mind: How young children come to understand that people have thoughts and feelings. *Infants & Young Children*, 13(3), 1-12.
- Astor, K., Thiele, M., & Gredebäck, G. (2021). Gaze following emergence relies on both perceptual cues and social awareness. *Cognitive Development*, 60, 101121.
- Aviezer, H., Ensenberg, N., & Hassin, R. R. (2017). The inherently contextualized nature of facial emotion perception. *Current opinion in psychology*, 17, 47-54.
- Aviezer, H., Trope, Y., & Todorov, A. (2012). Body cues, not facial expressions, discriminate between intense positive and negative emotions. *Science*, 338(6111), 1225-1229.
- Baird, A. D., Scheffer, I. E., & Wilson, S. J. (2011). Mirror neuron system involvement in empathy: a critical look at the evidence. *Social neuroscience*, 6(4), 327-335.
- Bambach, S., Crandall, D. J., Smith, L. B., & Yu, C. (2016). Active Viewing in Toddlers Facilitates Visual Object Learning: An Egocentric Vision Approach. In *CogSci*.
- Bambach, S., Crandall, D., Smith, L., & Yu, C. (2018). Toddler-inspired visual object learning. *Advances in neural information processing systems*, 31.
- Banks, M. S., & Salapatek, P. (1978). Acuity and contrast sensitivity in 1-, 2-, and 3-month-old human infants. *Investigative Ophthalmology & Visual Science*, 17(4), 361-365.

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.

Baron-Cohen, S., Tager-Flusberg, H., & Lombardo, M. (Eds.). (2013). *Understanding other minds: Perspectives from developmental social neuroscience*. OUP Oxford.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The “Reading the Mind in the Eyes” Test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, 42(2), 241-251.

Barrett, L. F. (2011). Was Darwin wrong about emotional expressions?. *Current Directions in Psychological Science*, 20(6), 400-406.

Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological science in the public interest*, 20(1), 1-68.

Barrett, L. F., Mesquita, B., & Gendron, M. (2011). Context in emotion perception. *Current directions in psychological science*, 20(5), 286-290.

Batty, M., & Taylor, M. J. (2003). Early processing of the six basic facial emotional expressions. *Cognitive Brain Research*, 17(3), 613-620.

Batty, M., & Taylor, M. J. (2006). The development of emotional face processing during childhood. *Developmental science*, 9(2), 207-220.

Bauser, D. S., Thoma, P., Aizenberg, V., Brüne, M., Juckel, G., & Daum, I. (2012). Face and body perception in schizophrenia: a configural processing deficit?. *Psychiatry Research*, 195(1-2), 9-17.

Bavelas, J. B., & Chovil, N. (1997). 15. Faces in dialogue. *The psychology of facial expression*, 334.

Bayet, L. (2022). How infants learn from a world of faces: implications for racial biases and mask-wearing. *Policy Insights from the Behavioral and Brain Sciences*, 9(1), 65-72.

Bayet, L., Behrendt, H. F., Cataldo, J. K., Westerlund, A., & Nelson, C. A. (2018). Recognition of facial emotions of varying intensities by three-year-olds. *Developmental psychology*, 54(12), 2240.

Bediou, B., Asri, F., Brunelin, J., Krolak-Salmon, P., D'Amato, T., Saoud, M., & Tazi, I. (2007). Emotion recognition and genetic vulnerability to schizophrenia. *The British Journal of Psychiatry*, 191(2), 126-130.

Beebe, B., Messinger, D., Bahrnick, L. E., Margolis, A., Buck, K. A., & Chen, H. (2016). A systems view of mother–infant face-to-face communication. *Developmental psychology*, 52(4), 556.

Behrendt, H. F., Scharke, W., Herpertz - Dahlmann, B., Konrad, K., & Firk, C. (2019). Like mother, like child? Maternal determinants of children's early social - emotional development. *Infant mental health journal*, 40(2), 234-247.

Bell, C. C., & McBride, D. F. (2010). Affect regulation and prevention of risky behaviors. *JAMA*, 304(5), 565-566.

Benders, T. (2013). Mommy is only happy! Dutch mothers' realisation of speech sounds in infant-directed speech expresses emotion, not didactic intent. *Infant Behavior and Development*, 36(4), 847-862.

Benitez-Quiroz, C., Srinivasan, R., & Martinez, A. M. (2016). Emotionet: An accurate, real-time algorithm for the automatic annotation of a million facial expressions in the wild. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 5562-5570).

Bennetts, R. J., Mole, J., & Bate, S. (2017). Super-recognition in development: A case study of an adolescent with extraordinary face recognition skills. *Cognitive neuropsychology*, 34(6), 357-376.

Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of cognitive neuroscience*, 8(6), 551-565.

Besel, L. D., & Yuille, J. C. (2010). Individual differences in empathy: The role of facial expression recognition. *Personality and individual differences*, 49(2), 107-112.

Bidet-Ildei, C., Kitromilides, E., Orliaguet, J. P., Pavlova, M., & Gentaz, E. (2014). Preference for point-light human biological motion in newborns: contribution of translational displacement. *Developmental psychology*, 50(1), 113.

Biedinger, N. (2011). The influence of education and home environment on the cognitive outcomes of preschool children in Germany. *Child Development Research*, 2011.

- Biotti, F., & Cook, R. (2016). Impaired perception of facial emotion in developmental prosopagnosia. *Cortex*, 81, 126-136.
- Biotti, F., Gray, K. L., & Cook, R. (2017). Impaired body perception in developmental prosopagnosia. *Cortex*, 93, 41-49.
- Bjornsdottir, R. T., Tskhay, K. O., Ishii, K., & Rule, N. O. (2017). Cultural differences in perceiving and processing emotions: a holistic approach to person perception. *Culture and Brain*, 5, 105-124.
- Blair, C., Granger, D., Willoughby, M., Kivlighan, K., & Family Life Project Investigators. (2006). Maternal sensitivity is related to hypothalamic - pituitary - adrenal axis stress reactivity and regulation in response to emotion challenge in 6 - month - old infants. *Annals of the New York Academy of Sciences*, 1094(1), 263-267.
- Blehar, M. C., Lieberman, A. F., & Ainsworth, M. D. S. (1977). Early face-to-face interaction and its relation to later infant-mother attachment. *Child development*, 182-194.
- Bobbitt, K. C., & Gershoff, E. T. (2016). Chaotic experiences and low-income children's social-emotional development. *Children and youth services review*, 70, 19-29.
- Bochkovskiy, A., Wang, C.Y. and Liao, H.Y.M., 2020. Yolov4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*.
- Booth, A. E., Pinto, J., & Bertenthal, B. I. (2002). Perception of the symmetrical patterning of human gait by infants. *Developmental Psychology*, 38(4), 554.

Borjon, J. I., Schroer, S. E., Bambach, S., Slone, L. K., Abney, D. H., Crandall, D. J., & Smith, L. B. (2018). A view of their own: Capturing the egocentric view of infants and toddlers with head-mounted cameras. *JoVE (Journal of Visualized Experiments)*, (140), e58445.

Bould, E., Morris, N., & Wink, B. (2008). Recognising subtle emotional expressions: The role of facial movements. *Cognition and emotion*, 22(8), 1569-1587.

Bourke, C., Douglas, K., & Porter, R. (2010). Processing of facial emotion expression in major depression: a review. *Australian & New Zealand Journal of Psychiatry*, 44(8), 681-696.

Brand, R. J., Baldwin, D. A., & Ashburn, L. A. (2002). Evidence for ‘motionese’ : modifications in mothers’ infant - directed action. *Developmental science*, 5(1), 72-83.

Braungart-Rieker, J., Garwood, M. M., Powers, B. P., & Notaro, P. C. (1998). Infant affect and affect regulation during the still-face paradigm with mothers and fathers: the role of infant characteristics and parental sensitivity. *Developmental Psychology*, 34(6), 1428.

Bremond-Gignac, D., Daruich, A., Robert, M. P., & Valleix, S. (2020). Recent developments in the management of congenital cataract. *Annals of translational medicine*, 8(22).

Brenna, V., Proietti, V., Montiroso, R., & Turati, C. (2013). Positive, but not negative, facial expressions facilitate 3-month-olds’ recognition of an individual face. *International Journal of Behavioral Development*, 37(2), 137-142.

Brennan, M., & Kirkland, J. (1982). Classification of infant cries using descriptive scales. *Infant Behavior and Development*, 5(2-4), 341-346.

Brody, L. R. (2000). The socialization of gender differences in emotional expression: Display rules, infant temperament, and differentiation. *Gender and emotion: Social psychological perspectives*, 2(11), 122-137.

Brody, L. R. (2013). On understanding gender differences in the expression of emotion: Gender roles, socialization, and language. In *Human feelings* (pp. 87-121). Routledge.

Brownell, C. A., Kopp, C. B., Brownell, C., & Kopp, C. (2007). Transitions in toddler socioemotional development. *Socioemotional development in the toddler years: Transitions and transformations*, 1, 40.

Bryan, A. E., & Dix, T. (2009). Mothers' emotions and behavioral support during interactions with toddlers: The role of child temperament. *Social Development*, 18(3), 647-670.

Bryant, G., & Barrett, H. C. (2008). Vocal emotion recognition across disparate cultures. *Journal of Cognition and Culture*, 8(1-2), 135-148.

Buck, R. W., Savin, V. J., Miller, R. E., & Caul, W. F. (1972). Communication of affect through facial expressions in humans. *Journal of personality and social psychology*, 23(3), 362.

Bulgarelli, D., & Molina, P. (2016). Social cognition in preschoolers: Effects of early experience and individual differences. *Frontiers in Psychology*, 7, 1762.

Burton, A.M. (2013). Why has research in face recognition progressed so slowly? The importance of variability. *The Quarterly Journal of Experimental Psychology*, 66(8), 1467-1485.

Bushneil, I. W. R., Sai, F., & Mullin, J. T. (1989). Neonatal recognition of the mother's face. *British Journal of Developmental Psychology*, 7(1), 3-15.

Butti, N., Finisguerra, A., & Urgesi, C. (2022). Holistic processing of body stimuli: Evidence of body composite illusion in adults and children. *Developmental Psychology*.

Cabeza, R., Bruce, V., Kato, T., & Oda, M. (1999). The prototype effect in face recognition: Extension and limits. *Memory & cognition*, 27, 139-151.

Calkins, S. D., & Mackler, J. S. (2011). Temperament, emotion regulation, and social development. *Social development: Relationships in infancy, childhood, and adolescence*, 2011, 44-70.

Calvo, M. G., & Nummenmaa, L. (2008). Detection of emotional faces: salient physical features guide effective visual search. *Journal of Experimental Psychology: General*, 137(3), 471.

Calvo, M. G., Gutiérrez-García, A., Fernández-Martín, A., & Nummenmaa, L. (2014). Recognition of facial expressions of emotion is related to their frequency in everyday life. *Journal of Nonverbal Behavior*, 38, 549-567.

Campos, J. J., Campos, R. G., & Barrett, K. C. (1989). Emergent themes in the study of emotional development and emotion regulation. *Developmental psychology*, 25(3), 394.

Camras, L. A., Oster, H., Bakeman, R., Meng, Z., Ujiie, T., & Campos, J. J. (2007). Do infants show distinct negative facial expressions for fear and anger? Emotional expression in 11 - month - old European American, Chinese, and Japanese infants. *Infancy*, 11(2), 131-155.

Canedo, D., & Neves, A. J. (2019). Facial expression recognition using computer vision: A systematic review. *Applied Sciences*, 9(21), 4678.

- Canli, T., & Lesch, K. P. (2007). Long story short: the serotonin transporter in emotion regulation and social cognition. *Nature neuroscience*, 10(9), 1103-1109.
- Cannon, E. N., & Woodward, A. L. (2012). Infants generate goal - based action predictions. *Developmental science*, 15(2), 292-298.
- Carbajal-Valenzuela, C. C., Santiago-Rodríguez, E., Quirarte, G. L., & Harmony, T. (2017). Development of emotional face processing in premature and full-term infants: a quantitative EEG study. *Clinical EEG and neuroscience*, 48(2), 88-95.
- Carey, S., Diamond, R., & Woods, B. (1980). Development of face recognition: A maturational component?. *Developmental psychology*, 16(4), 257.
- Caron, A. J., Caron, R. F., & MacLean, D. J. (1988). Infant discrimination of naturalistic emotional expressions: The role of face and voice. *Child development*, 604-616.
- Carranza Carnicero, J. A., Pérez - López, J., Del Carmen González Salinas, M., & Martínez - Fuentes, M. T. (2000). A longitudinal study of temperament in infancy: Stability and convergence of measures. *European Journal of Personality*, 14(1), 21-37.
- Carrasco, M. Á., Holgado-Tello, F. P., Delgado, B., & González-Peña, P. (2016). Reactive temperament traits and behavioural problems in children: The mediating role of effortful control across sex and age. *European Journal of Developmental Psychology*, 13(2), 197-212.
- Carver, L. J., Dawson, G., Panagiotides, H., Meltzoff, A. N., McPartland, J., Gray, J., & Munson, J. (2003). Age - related differences in neural correlates of face recognition during the toddler and

preschool years. *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology*, 42(2), 148-159.

Cassia, V. M., Kuefner, D., Picozzi, M., & Vescovo, E. (2009). Early experience predicts later plasticity for face processing: Evidence for the reactivation of dormant effects. *Psychological science*, 20(7), 853-859.

Cassia, V. M., Proietti, V., & Pisacane, A. (2013). Early and later experience with one younger sibling affects face processing abilities of 6-year-old children. *International Journal of Behavioral Development*, 37(2), 160-165.

Castelli, F. (2005). Understanding emotions from standardized facial expressions in autism and normal development. *Autism*, 9(4), 428-449.

Cavallini, A., Fazzi, E., Viviani, V., Astori, M. G., Zaverio, S., Bianchi, P. E., & Lanzi, G. (2002). Visual acuity in the first two years of life in healthy term newborns: an experience with the teller acuity cards. *Functional neurology*, 17(2), 87-92.

Chaplin, T. M., Cole, P. M., & Zahn-Waxler, C. (2005). Parental socialization of emotion expression: gender differences and relations to child adjustment. *Emotion*, 5(1), 80.

Cheal, J. L., & Rutherford, M. D. (2011). Categorical perception of emotional facial expressions in preschoolers. *Journal of Experimental Child Psychology*, 110(3), 434-443.

Chetcuti, L., Uljarević, M., Ellis-Davies, K., Hardan, A. Y., Whitehouse, A. J., Hedley, D., ... & Prior, M. R. (2021). Temperament in individuals with autism spectrum disorder: A systematic review. *Clinical Psychology Review*, 85, 101984.

- Chetcuti, L., Uljarević, M., Varcin, K. J., Boutrus, M., Wan, M. W., Green, J., ... & AICES Team. (2020). Subgroups of Temperament Associated with Social–Emotional Difficulties in Infants with Early Signs of Autism. *Autism Research*, 13(12), 2094-2101.
- Chiarella, S. S., & Poulin - Dubois, D. (2013). Cry babies and pollyannas: Infants can detect unjustified emotional reactions. *Infancy*, 18, E81-E96.
- Chiroro, P., & Valentine, T. (1995). An investigation of the contact hypothesis of the own-race bias in face recognition. *The Quarterly Journal of Experimental Psychology Section A*, 48(4), 879-894.
- Chong, S. C. F., Werker, J. F., Russell, J. A., & Carroll, J. M. (2003). Three facial expressions mothers direct to their infants. *Infant and Child Development: An International Journal of Research and Practice*, 12(3), 211-232.
- Chouchourelou, A., Matsuka, T., Harber, K., & Shiffrar, M. (2006). The visual analysis of emotional actions. *Social Neuroscience*, 1(1), 63-74.
- Chung, M. S., & Thomson, D. M. (1995). Development of face recognition. *British Journal of Psychology*, 86(1), 55-87.
- Clark, L. A., Watson, D., & Mineka, S. (1994). Temperament, personality, and the mood and anxiety disorders. *Journal of abnormal psychology*, 103(1), 103.
- Clifford, S. M., Hudry, K., Elsabbagh, M., Charman, T., & Johnson, M. H. (2013). Temperament in the first 2 years of life in infants at high-risk for autism spectrum disorders. *Journal of autism and developmental disorders*, 43(3), 673-686.

Clifford, S. M., Hudry, K., Elsabbagh, M., Charman, T., Johnson, M. H., & BASIS Team. (2013). Temperament in the first 2 years of life in infants at high-risk for autism spectrum disorders. *Journal of autism and developmental disorders*, 43, 673-686.

Cloninger, C. R., Cloninger, K. M., Zwir, I., & Keltikangas-Järvinen, L. (2019). The complex genetics and biology of human temperament: a review of traditional concepts in relation to new molecular findings. *Translational psychiatry*, 9(1), 290.

Cohen, L. B. (1976). Habituation of infant visual attention. *Habituation: Perspectives from child development, animal behavior, and neurophysiology*, 207-238.

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.

Collin, L., Bindra, J., Raju, M., Gillberg, C., & Minnis, H. (2013). Facial emotion recognition in child psychiatry: a systematic review. *Research in developmental disabilities*, 34(5), 1505-1520.

Colson, E. R., & Dworkin, P. H. (1997). Toddler development. *Pediatrics in review*, 18(8), 255-259.

Coman, A. (2021). Perceptual Sensitivity and Response Bias Towards Infant and Adult Facial Expressions in Mothers Versus Non-Mothers (Master's thesis, Reichman University (Israel)).

Conte, S., Proietti, V., Quadrelli, E., Turati, C., & Cassia, V. M. (2019). Sibling experience prevents neural tuning to adult faces in 10-month-old infants. *Neuropsychologia*, 129, 72-82.

Cook, R., Brewer, R., Shah, P., & Bird, G. (2013). Alexithymia, not autism, predicts poor recognition of emotional facial expressions. *Psychological science*, 24(5), 723-732.

Courchesne, E. (1977). Event-related brain potentials: Comparison between children and adults. *Science*, 197(4303), 589-592.

Cowen, A. S., & Keltner, D. (2020). What the face displays: Mapping 28 emotions conveyed by naturalistic expression. *American Psychologist*, 75(3), 349.

Crespo-Llado, M. M., Vanderwert, R. E., & Geangu, E. (2018). Individual differences in infants' neural responses to their peers' cry and laughter. *Biological psychology*, 135, 117-127.

Crockenberg, S. C., Leerkes, E. M., & J6, P. S. B. (2008). Predicting aggressive behavior in the third year from infant reactivity and regulation as moderated by maternal behavior. *Development and Psychopathology*, 20(1), 37-54.

Dahl, A., Sherlock, B. R., Campos, J. J., & Theunissen, F. E. (2014). Mothers' tone of voice depends on the nature of infants' transgressions. *Emotion*, 14(4), 651.

Dalrymple, K. A., Corrow, S., Yonas, A., & Duchaine, B. (2012). Developmental prosopagnosia in childhood. *Cognitive neuropsychology*, 29(5-6), 393-418.

Daltrozzo, J., Wioland, N., & Kotchoubey, B. (2012). The N400 and late positive complex (LPC) effects reflect controlled rather than automatic mechanisms of sentence processing. *Brain sciences*, 2(3), 267-297.

Damasio, A. R. (1985). Prosopagnosia. *Trends in Neurosciences*, 8, 132-135.

Darwin, C. (1872). 1965. The expression of the emotions in man and animals, 2.

Darwin, C. (1872). The expression of the emotions in man and animals, 1872.

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

Davenport, E., Yap, M. B., Simmons, J. G., Sheeber, L. B., & Allen, N. B. (2011). Maternal and adolescent temperament as predictors of maternal affective behavior during mother–adolescent interactions. *Journal of adolescence*, 34(5), 829-839.

Davidson, R. J., & Fox, N. A. (1989). Frontal brain asymmetry predicts infants' response to maternal separation. *Journal of abnormal psychology*, 98(2), 127.

Davis, D. H., & Thoman, E. B. (1988). The early social environment of premature and fullterm infants. *Early human development*, 17(2-3), 221-232.

Davis, M. H., Luce, C., & Kraus, S. J. (1994). The heritability of characteristics associated with dispositional empathy. *Journal of personality*, 62(3), 369-391.

Dawel, A., O’Kearney, R., McKone, E., & Palermo, R. (2012). Not just fear and sadness: Meta-analytic evidence of pervasive emotion recognition deficits for facial and vocal expressions in psychopathy. *Neuroscience & Biobehavioral Reviews*, 36(10), 2288-2304.

Dawson, G., Webb, S. J., Carver, L., Panagiotides, H., & McPartland, J. (2004). Young children with autism show atypical brain responses to fearful versus neutral facial expressions of emotion. *Developmental science*, 7(3), 340-359.

de Gelder, B., & Rouw, R. (2000). Configural face processes in acquired and developmental prosopagnosia: evidence for two separate face systems?. *Neuroreport*, 11(14), 3145-3150.

De Gelder, B., Frissen, I., Barton, J., & Hadjikhani, N. (2003). A modulatory role for facial expressions in prosopagnosia. *Proceedings of the National Academy of Sciences*, 100(22), 13105-13110.

de Gelder, B., Teunisse, J. P., & Benson, P. J. (1997). Categorical perception of facial expressions: Categories and their internal structure. *Cognition & Emotion*, 11(1), 1-23.

De Haan, M. (2013). *Infant EEG and event-related potentials*. Psychology Press.

De Haan, M., Belsky, J., Reid, V., Volein, A., & Johnson, M. H. (2004). Maternal personality and infants' neural and visual responsivity to facial expressions of emotion. *Journal of Child Psychology and Psychiatry*, 45(7), 1209-1218.

De Haan, M., Johnson, M. H., & Halit, H. (2003). Development of face-sensitive event-related potentials during infancy: a review. *International Journal of Psychophysiology*, 51(1), 45-58.

De Haan, M., Johnson, M. H., Maurer, D., & Perrett, D. I. (2001). Recognition of individual faces and average face prototypes by 1-and 3-month-old infants. *Cognitive development*, 16(2), 659-678.

De Heering, A., De Liedekerke, C., Deboni, M., & Rossion, B. (2010). The role of experience during childhood in shaping the other - race effect. *Developmental science*, 13(1), 181-187.

De Sonnevile, L. M. J., Verschoor, C. A., Njokiktjien, C., Op het Veld, V., Toorenaar, N., & Vranken, M. (2002). Facial identity and facial emotions: speed, accuracy, and processing strategies in children and adults. *Journal of Clinical and experimental neuropsychology*, 24(2), 200-213.

Deater - Deckard, K., Mullineaux, P. Y., Beekman, C., Petrill, S. A., Schatschneider, C., & Thompson, L. A. (2009). Conduct problems, IQ, and household chaos: A longitudinal multi - informant study. *Journal of Child Psychology and Psychiatry*, 50(10), 1301-1308.

Decety, J., & Meyer, M. (2008). From emotion resonance to empathic understanding: A social developmental neuroscience account. *Development and psychopathology*, 20(4), 1053-1080.

Deen, B., Richardson, H., Dilks, D. D., Takahashi, A., Keil, B., Wald, L. L., ... & Saxe, R. (2017). Organization of high-level visual cortex in human infants. *Nature communications*, 8(1), 1-10.

Del Vecchio, T., & Rhoades, K. A. (2010). Bidirectional influences in mother–toddler dyads: an examination of the relative influence of mothers' and children's behaviours. *Infant and child development*, 19(5), 516-529.

Deng, J., Guo, J., Ververas, E., Kotsia, I., & Zafeiriou, S. (2020). Retinaface: Single-shot multi-level face localisation in the wild. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 5203-5212).

Deng, J., Guo, J., Xue, N., & Zafeiriou, S. (2019). Arcface: Additive angular margin loss for deep face recognition. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 4690-4699).

Dennis, T. A., Malone, M. M., & Chen, C. C. (2009). Emotional face processing and emotion regulation in children: an ERP study. *Developmental neuropsychology*, 34(1), 85-102.

Derntl, B., Finkelmeyer, A., Toygar, T. K., Hülsmann, A., Schneider, F., Falkenberg, D. I., & Habel, U. (2009). Generalized deficit in all core components of empathy in schizophrenia.

Schizophrenia research, 108(1-3), 197-206.

Dhall, A., Goecke, R., Lucey, S., & Gedeon, T. (2011). Acted facial expressions in the wild database. Australian National University, Canberra, Australia, Technical Report TR-CS-11, 2(1).

Di Giorgio, E., Leo, I., Pascalis, O., & Simion, F. (2012). Is the face-perception system human-specific at birth?. *Developmental psychology*, 48(4), 1083.

Di Lorenzo, R., van den Boomen, C., Kemner, C., & Junge, C. (2020). Charting development of ERP components on face-categorization: Results from a large longitudinal sample of infants. *Developmental cognitive neuroscience*, 45, 100840.

Diamond, R., & Carey, S. (1977). Developmental changes in the representation of faces. *Journal of experimental child psychology*, 23(1), 1-22.

Diego, M. A., Field, T., Hart, S., Hernandez - Reif, M., Jones, N., Cullen, C., ... & Kuhn, C. (2002). Facial expressions and EEG in infants of intrusive and withdrawn mothers with depressive symptoms. *Depression and Anxiety*, 15(1), 10-17.

Dimberg, U. (1982). Facial reactions to facial expressions. *Psychophysiology*, 19(6), 643-647.

Ding, X., Zhang, X., Ma, N., Han, J., Ding, G. and Sun, J., 2021. Repvgg: Making vgg-style convnets great again. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 13733-13742).

Dittrich, W. H., Troscianko, T., Lea, S. E., & Morgan, D. (1996). Perception of emotion from dynamic point-light displays represented in dance. *Perception*, 25(6), 727-738.

Dobel, C., Bölte, J., Aicher, M., & Schweinberger, S. R. (2007). Prosopagnosia without apparent cause: Overview and diagnosis of six cases. *Cortex*, 43(6), 718-733.

Doi, H., Koga, T., & Shinohara, K. (2009). 18-Month-olds can perceive Mooney faces. *Neuroscience Research*, 64(3), 317-322.

Domnich, A., & Anbarjafari, G. (2021). Responsible AI: Gender bias assessment in emotion recognition. arXiv preprint arXiv:2103.11436.

Dondi, M., Simion, F., & Caltran, G. (1999). Can newborns discriminate between their own cry and the cry of another newborn infant?. *Developmental psychology*, 35(2), 418.

Doretto, V., & Scivoletto, S. (2018). Effects of early neglect experience on recognition and processing of facial expressions: A systematic review. *Brain Sciences*, 8(1), 10.

Dotti Sani, G. M., & Treas, J. (2016). Educational gradients in parents' child - care time across countries, 1965 – 2012. *Journal of Marriage and Family*, 78(4), 1083-1096.

Doty, T. J., Japee, S., Ingvar, M., & Ungerleider, L. G. (2013). Fearful face detection sensitivity in healthy adults correlates with anxiety-related traits. *Emotion*, 13(2), 183.

Downey, D. B., Condron, D. J., & Yucel, D. (2015). Number of siblings and social skills revisited among American fifth graders. *Journal of Family Issues*, 36(2), 273-296.

Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science*, 293(5539), 2470-2473.

Drimalla, H., Landwehr, N., Hess, U., & Dziobek, I. (2019). From face to face: the contribution of facial mimicry to cognitive and emotional empathy. *Cognition and Emotion*.

Duchaine, B. C., Parker, H., & Nakayama, K. (2003). Normal recognition of emotion in a prosopagnosic. *Perception*, 32(7), 827-838.

Dunbar, R. I. (2012). Bridging the bonding gap: the transition from primates to humans. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1597), 1837-1846.

Dunbar, R. I., & Shultz, S. (2007). Understanding primate brain evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1480), 649-658.

Dupré, D., Krumhuber, E. G., Küster, D., & McKeown, G. J. (2020). A performance comparison of eight commercially available automatic classifiers for facial affect recognition. *Plos one*, 15(4), e0231968.

Duque, A., & Vázquez, C. (2015). Double attention bias for positive and negative emotional faces in clinical depression: Evidence from an eye-tracking study. *Journal of Behavior Therapy and Experimental Psychiatry*, 46, 107-114.

Durán, J. I., & Fernández-Dols, J. M. (2021). Do emotions result in their predicted facial expressions? A meta-analysis of studies on the co-occurrence of expression and emotion. *Emotion*, 21(7), 1550.

Durán, J. I., Reisenzein, R., & Fernández-Dols, J. M. (2017). Coherence between emotions and facial expressions. *The science of facial expression*, 107-129.

Eack, S. M., Mazefsky, C. A., & Minshew, N. J. (2015). Misinterpretation of facial expressions of emotion in verbal adults with autism spectrum disorder. *Autism*, 19(3), 308-315.

Easterbrook, M. A., Kisilevsky, B. S., Muir, D. W., & Laplante, D. P. (1999). Newborns discriminate schematic faces from scrambled faces. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 53(3), 231.

Ekas, N. V., Lickenbrock, D. M., & Braungart - Rieker, J. M. (2013). Developmental trajectories of emotion regulation across infancy: Do age and the social partner influence temporal patterns. *Infancy*, 18(5), 729-754.

Ekman, P. (1999). Basic emotions. *Handbook of cognition and emotion*, 98(45-60), 16.

Ekman, P. (2003). Darwin, deception, and facial expression. *Ann. N. Y. Acad. Sci.* 1000, 205–221. doi: 10.1196/annals.1280.010

Ekman, P. (2009). Lie catching and microexpressions. *The philosophy of deception*, 1(2), 5.

Ekman, P. (Ed.). (2006). *Darwin and facial expression: A century of research in review*. Ishk.

Ekman, P. (2014). Expression and the nature of emotion. *Approaches to emotion*, 319-343.

Ekman, P., & Cordaro, D. (2011). What is meant by calling emotions basic. *Emotion review*, 3(4), 364-370.

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

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of personality and social psychology*, 17(2), 124.

Ekman, P., & Friesen, W. V. (1978). Facial action coding systems. Consulting Psychologists Press.

Ekman, P., & Keltner, D. (1970). Universal facial expressions of emotion. *California mental health research digest*, 8(4), 151-158.

Ekman, P., & Keltner, D. (1997). Universal facial expressions of emotion. Segerstrale U, P. Molnar P, eds. *Nonverbal communication: Where nature meets culture*, 27-46.

Ekman, P., & Oster, H. (1979). Facial expressions of emotion. *Annual review of psychology*, 30(1), 527-554.

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

Ekman, P., Sorenson, E. R., & Friesen, W. V. (1969). Pan-cultural elements in facial displays of emotion. *Science*, 164(3875), 86-88.

Elfenbein, H. A. (2013). Nonverbal dialects and accents in facial expressions of emotion. *Emotion Review*, 5(1), 90-96.

Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: a meta-analysis. *Psychological bulletin*, 128(2), 203.

Elfenbein, H. A., & Ambady, N. (2003). When familiarity breeds accuracy: cultural exposure and facial emotion recognition. *Journal of personality and social psychology*, 85(2), 276.

Elicker, J., Ruprecht, K. M., & Anderson, T. (2014). Observing infants' and toddlers' relationships and interactions in group care. *Lived spaces of infant-toddler education and care: Exploring diverse perspectives on theory, research and practice*, 131-145.

Enlow, M. B., Kitts, R. L., Blood, E., Bizarro, A., Hofmeister, M., & Wright, R. J. (2011). Maternal posttraumatic stress symptoms and infant emotional reactivity and emotion regulation. *Infant Behavior and Development*, 34(4), 487-503.

Enlow, M. B., White, M. T., Hails, K., Cabrera, I., & Wright, R. J. (2016). The Infant Behavior Questionnaire-Revised: Factor structure in a culturally and sociodemographically diverse sample in the United States. *Infant Behavior and Development*, 43, 24-35.

Ertekin, Z., Gunnar, M. R., & Berument, S. K. (2021). Temperament moderates the effects of early deprivation on infant attention. *Infancy*, 26(3), 455-468.

Essler, S., & Paulus, M. (2021). When do caregivers begin to view their child as a moral agent? Comparing moral and non-moral reactions to young children's moral transgressions. *Journal of Moral Education*, 50(3), 330-342.

Etkin, A., Klemenhagen, K. C., Dudman, J. T., Rogan, M. T., Hen, R., Kandel, E. R., & Hirsch, J. (2004). Individual differences in trait anxiety predict the response of the basolateral amygdala to unconsciously processed fearful faces. *Neuron*, 44(6), 1043-1055.

Fairchild, G., Hagan, C. C., Passamonti, L., Walsh, N. D., Goodyer, I. M., & Calder, A. J. (2014). Atypical neural responses during face processing in female adolescents with conduct disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, 53(6), 677-687.

Fairchild, G., Van Goozen, S. H., Calder, A. J., Stollery, S. J., & Goodyer, I. M. (2009). Deficits in facial expression recognition in male adolescents with early - onset or adolescence - onset conduct disorder. *Journal of Child Psychology and Psychiatry*, 50(5), 627-636.

Farah, M. J. (1996). Is face recognition 'special'? Evidence from neuropsychology. *Behavioural brain research*, 76(1-2), 181-189.

Farah, M. J., Rabinowitz, C., Quinn, G. E., & Liu, G. T. (2000). Early commitment of neural substrates for face recognition. *Cognitive neuropsychology*, 17(1-3), 117-123.

Farran, E. K., Branson, A., & King, B. J. (2011). Visual search for basic emotional expressions in autism; impaired processing of anger, fear and sadness, but a typical happy face advantage. *Research in Autism Spectrum Disorders*, 5(1), 455-462.

Farrell, A. K., Simpson, J. A., Carlson, E. A., Englund, M. M., & Sung, S. (2017). The impact of stress at different life stages on physical health and the buffering effects of maternal sensitivity. *Health Psychology*, 36(1), 35.

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.

Fausey, C. M., Jayaraman, S., & Smith, L. B. (2016). From faces to hands: Changing visual input in the first two years. *Cognition*, 152, 101-107.

Feinman, S. (1982). Social referencing in infancy. *Merrill-Palmer Quarterly* (1982-), 445-470.

FELDMAN, R. S., McGEE, G. G., Mann, L., & Strain, P. (1993). Nonverbal affective decoding ability in children with autism and in typical preschoolers. *Journal of Early Intervention*, 17(4), 341-350.

Feldman, R., & Eidelman, A. I. (2009). Biological and environmental initial conditions shape the trajectories of cognitive and social - emotional development across the first years of life. *Developmental science*, 12(1), 194-200.

Field, T. M. (1981). Infant gaze aversion and heart rate during face-to-face interactions. *Infant behavior and development*, 4, 307-315.

Field, T. M., & Walden, T. A. (1982). Production and perception of facial expressions in infancy and early childhood. *Advances in child development and behavior*, 16, 169-211.

Field, T. M., Woodson, R., Greenberg, R., & Cohen, D. (1983). Facial expression by neonates. *Annual Progress in Child Psychiatry and Child Development*, 16, 119-125.

Fiset, D., Blais, C., Royer, J., Richoz, A. R., Dugas, G., & Caldara, R. (2017). Mapping the impairment in decoding static facial expressions of emotion in prosopagnosia. *Social cognitive and affective neuroscience*, 12(8), 1334-1341.

- Foster, M. A., Lambert, R., Abbott-Shim, M., McCarty, F., & Franze, S. (2005). A model of home learning environment and social risk factors in relation to children's emergent literacy and social outcomes. *Early childhood research quarterly*, 20(1), 13-36.
- Fox, N. A., & Henderson, H. A. (1999). Does infancy matter? Predicting social behavior from infant temperament. *Infant Behavior and Development*, 22(4), 445-455.
- Fox, R., & McDaniel, C. (1982). The perception of biological motion by human infants. *Science*, 218(4571), 486-487.
- Framorando, D., Bapst, M., Vuille, N., & Pegna, A. J. (2017). Naso-temporal asymmetries: suppression of emotional faces in the temporal visual hemifield. *Frontiers in neuroscience*, 11, 14.
- Franchak, J. M. (2019). Changing opportunities for learning in everyday life: Infant body position over the first year. *Infancy*, 24(2), 187-209.
- Franchak, J. M., Kretch, K. S., & Adolph, K. E. (2018). See and be seen: Infant–caregiver social looking during locomotor free play. *Developmental science*, 21(4), e12626.
- Frank, M., Simmons, K., Yurovsky, D., & Pusiol, G. (2013). Developmental and postural changes in children's visual access to faces. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 35, No. 35).
- Fridlund, A. J., & Russell, J. A. (2006). The functions of facial expressions: what's in a face?.
- Frith, C. (2009). Role of facial expressions in social interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1535), 3453-3458.

Frith, C., & Frith, U. (2005). Theory of mind. *Current biology*, 15(17), R644-R645.

Fu, X., Morales, S., LoBue, V., Buss, K. A., & Pérez - Edgar, K. (2020). Temperament moderates developmental changes in vigilance to emotional faces in infants: Evidence from an eye - tracking study. *Developmental Psychobiology*, 62(3), 339-352.

Fu, Z. W., Wang, K. L., Du, N. Y., Li, Y. J., Duan, J. J., & Zhao, S. X. (2024). Comparative study on emotional behavior and parental job stress of only-child and non-only-child preschool children. *World Journal of Clinical Cases*, 12(21), 4642-4651.

Fugate, J., Gouzoules, H., & Barrett, L. F. (2010). Reading chimpanzee faces: evidence for the role of verbal labels in categorical perception of emotion. *Emotion*, 10(4), 544.

Furl, N., Garrido, L., Dolan, R. J., Driver, J., & Duchaine, B. (2011). Fusiform gyrus face selectivity relates to individual differences in facial recognition ability. *Journal of cognitive neuroscience*, 23(7), 1723-1740.

Gaebel, W., & Wölwer, W. (2004). Facial expressivity in the course of schizophrenia and depression. *European archives of psychiatry and clinical neuroscience*, 254, 335-342.

Galati, D., Sini, B., Schmidt, S., & Tinti, C. (2003). Spontaneous facial expressions in congenitally blind and sighted children aged 8–11. *Journal of Visual Impairment & Blindness*, 97(7), 418-428.

Gao, C., Conte, S., Richards, J. E., Xie, W., & Hanayik, T. (2019). The neural sources of N170: Understanding timing of activation in face-selective areas. *Psychophysiology*, 56(6), e13336.

Gao, X., Maurer, D., & Nishimura, M. (2013). Altered representation of facial expressions after early visual deprivation. *Frontiers in Psychology*, 4, 878.

Gartstein, M. A., & Rothbart, M. K. (2003). Studying infant temperament via the revised infant behavior questionnaire. *Infant Behavior and Development*, 26(1), 64-86.

Gartstein, M. A., Putnam, S. P., & Rothbart, M. K. (2012). Etiology of preschool behavior problems: Contributions of temperament attributes in early childhood. *Infant Mental Health Journal*, 33(2), 197-211.

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature neuroscience*, 3(2), 191-197.

Geangu, E., & Vuong, Q. C. (2020). Look up to the body: An eye-tracking investigation of 7-months-old infants' visual exploration of emotional body expressions. *Infant Behavior and Development*, 60, 101473.

Geangu, E., & Vuong, Q. C. (2023). Seven - months - old infants show increased arousal to static emotion body expressions: Evidence from pupil dilation. *Infancy*.

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.

Geangu, E., Senna, I., Croci, E., & Turati, C. (2015). The effect of biomechanical properties of motion on infants' perception of goal-directed grasping actions. *Journal of experimental child psychology*, 129, 55-67.

Geangu, E., Smith, W. A., Mason, H. T., Martinez-Cedillo, A. P., Hunter, D., Knight, M. I., ... & Muller, B. R. (2023). EgoActive: integrated wireless wearable sensors for capturing infant egocentric auditory–visual statistics and autonomic nervous system function ‘in the wild’. *Sensors*, 23(18), 7930.

Gee, D. G., Humphreys, K. L., Flannery, J., Goff, B., Telzer, E. H., Shapiro, M., ... & Tottenham, N. (2013). A developmental shift from positive to negative connectivity in human amygdala–prefrontal circuitry. *Journal of Neuroscience*, 33(10), 4584-4593.

Geldart, S., Mondloch, C. J., Maurer, D., De Schonen, S., & Brent, H. P. (2002). The effect of early visual deprivation on the development of face processing. *Developmental Science*, 5(4), 490-501.

Gery, I., Miljkovitch, R., Berthoz, S., & Soussignan, R. (2009). Empathy and recognition of facial expressions of emotion in sex offenders, non-sex offenders and normal controls. *Psychiatry research*, 165(3), 252-262.

Gillmeister, H., Stets, M., Grigorova, M., & Rigato, S. (2019). How do bodies become special? Electrophysiological evidence for the emergence of body-related cortical processing in the first 14 months of life. *Developmental psychology*, 55(10), 2025.

Golan, O., Baron-Cohen, S., Hill, J. J., & Rutherford, M. D. (2007). The ‘Reading the Mind in the Voice’ test-revised: a study of complex emotion recognition in adults with and without autism spectrum conditions. *Journal of autism and developmental disorders*, 37(6), 1096-1106.

Golarai, G., Liberman, A., & Grill-Spector, K. (2015). Experience shapes the development of neural substrates of face processing in human ventral temporal cortex. *Cerebral Cortex*, 27(2), bhv314.

Goldfield, E. C. (1989). Transition from rocking to crawling: Postural constraints on infant movement. *Developmental Psychology*, 25(6), 913.

Goldsmith, H. H., & Campos, J. J. (1982). Toward a theory of infant temperament. The development of attachment and affiliative systems, 161-193.

Gorno-Tempini, M. L., Pradelli, S., Serafini, M., Pagnoni, G., Baraldi, P., Porro, C., ... & Nichelli, P. (2001). Explicit and incidental facial expression processing: an fMRI study. *Neuroimage*, 14(2), 465-473.

Grady, C. L., Mondloch, C. J., Lewis, T. L., & Maurer, D. (2014). Early visual deprivation from congenital cataracts disrupts activity and functional connectivity in the face network. *Neuropsychologia*, 57, 122-139.

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.

Grenville, E., & Dwyer, D. M. (2022). Face masks have emotion-dependent dissociable effects on accuracy and confidence in identifying facial expressions of emotion. *Cognitive research: principles and implications*, 7(1), 15.

Grolnick, W. S., Bridges, L. J., & Connell, J. P. (1996). Emotion regulation in two - year - olds: Strategies and emotional expression in four contexts. *Child development*, 67(3), 928-941.

Grolnick, W. S., McMenamy, J. M., Kurowski, C. O., Balter, L., & Tamis-LeMonda, C. S. (1999). Emotional and self-regulation in infancy and toddlerhood. *Child psychology*, 1.

Grossmann, T. (2010). The development of emotion perception in face and voice during infancy. *Restorative neurology and neuroscience*, 28(2), 219-236.

Grossmann, T., Striano, T., & Friederici, A. D. (2007). Developmental changes in infants' processing of happy and angry facial expressions: A neurobehavioral study. *Brain and cognition*, 64(1), 30-41.

Guo, K. (2012). Holistic gaze strategy to categorize facial expression of varying intensities.

Hadad, B. S., Maurer, D., & Lewis, T. L. (2012). Sparing of sensitivity to biological motion but not of global motion after early visual deprivation. *Developmental Science*, 15(4), 474-481.

Hadders-Algra, M. (2022). Human face and gaze perception is highly context specific and involves bottom-up and top-down neural processing. *Neuroscience & Biobehavioral Reviews*, 132, 304-323.

Hale, S. (1990). A global developmental trend in cognitive processing speed. *Child development*, 61(3), 653-663.

Haley, D. W., & Stansbury, K. (2003). Infant stress and parent responsiveness: Regulation of physiology and behavior during still - face and reunion. *Child development*, 74(5), 1534-1546.

Halit, H., De Haan, M., & Johnson, M. H. (2003). Cortical specialisation for face processing: face-sensitive event-related potential components in 3-and 12-month-old infants. *Neuroimage*, 19(3), 1180-1193.

Hallock, M. B., Worobey, J., & Self, P. A. (1989). Behavioural development in chimpanzee (*Pan troglodytes*) and human newborns across the first month of life. *International Journal of Behavioral Development*, 12(4), 527-540.

Hancock, K. J., & Rhodes, G. (2008). Contact, configural coding and the other - race effect in face recognition. *British Journal of Psychology*, 99(1), 45-56.

Harris, A., Vyas, D. B., & Reed, C. L. (2016). Holistic processing for bodies and body parts: New evidence from stereoscopic depth manipulations. *Psychonomic Bulletin & Review*, 23, 1513-1519.

Hartling, C., Fan, Y., Weigand, A., Trilla, I., Gärtner, M., Bajbouj, M., ... & Grimm, S. (2019). Interaction of HPA axis genetics and early life stress shapes emotion recognition in healthy adults. *Psychoneuroendocrinology*, 99, 28-37.

Hartshorne, T. S., & Schmitt, M. C. (2016). Social-emotional development in children and youth who are deafblind. *American annals of the deaf*, 161(4), 444-453.

Hashiya, K., Meng, X., Uto, Y., & Tajiri, K. (2019). Overt congruent facial reaction to dynamic emotional expressions in 9–10-month-old infants. *Infant Behavior and Development*, 54, 48-56.

Hayden, A., Bhatt, R. S., Joseph, J. E., & Tanaka, J. W. (2007). The other - race effect in infancy: Evidence using a morphing technique. *Infancy*, 12(1), 95-104.

Heck, A., Chroust, A., White, H., Jubran, R., & Bhatt, R. S. (2018). Development of body emotion perception in infancy: From discrimination to recognition. *Infant Behavior and Development*, 50, 42-51.

Henderson, A. J., Holzleitner, I. J., Talamas, S. N., & Perrett, D. I. (2016). Perception of health from facial cues. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1693), 20150380.

Henderson, H. A., Fox, N. A., & Rubin, K. H. (2001). Temperamental contributions to social behavior: The moderating roles of frontal EEG asymmetry and gender. *Journal of the American Academy of Child & Adolescent Psychiatry*, 40(1), 68-74.

Herba, C., & Phillips, M. (2004). Annotation: Development of facial expression recognition from childhood to adolescence: Behavioural and neurological perspectives. *Journal of Child Psychology and Psychiatry*, 45(7), 1185-1198.

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.

Heron, M., & Slaughter, V. (2010). Infants' responses to real humans and representations of humans. *International Journal of Behavioral Development*, 34(1), 34-45.

Hess, U., Blairy, S., & Kleck, R. E. (1997). The intensity of emotional facial expressions and decoding accuracy. *Journal of Nonverbal Behavior*, 21, 241-257.

Hessels, R. S. (2020). How does gaze to faces support face-to-face interaction? A review and perspective. *Psychonomic bulletin & review*, 27(5), 856-881.

Hirschler - Guttenberg, Y., Golan, O., Ostfeld - Etzion, S., & Feldman, R. (2015). Mothering, fathering, and the regulation of negative and positive emotions in high - functioning preschoolers with autism spectrum disorder. *Journal of Child Psychology and Psychiatry*, 56(5), 530-539.

Hock, A., White, H., Jubran, R., & Bhatt, R. S. (2016). The whole picture: Holistic body posture recognition in infancy. *Psychonomic Bulletin & Review*, 23, 426-431.

Hock, A., White, H., Jubran, R., & Bhatt, R. S. (2016). The whole picture: Holistic body posture recognition in infancy. *Psychonomic Bulletin & Review*, 23(2), 426-431.

Hoehl, S., & Striano, T. (2008). Neural processing of eye gaze and threat - related emotional facial expressions in infancy. *Child development*, 79(6), 1752-1760.

Hoehl, S., & Striano, T. (2010). The development of emotional face and eye gaze processing. *Developmental Science*, 13(6), 813-825.

Hoehl, S., & Wahl, S. (2012). Recording infant ERP data for cognitive research. *Developmental neuropsychology*, 37(3), 187-209.

Hoehl, S., Wiese, L., & Striano, T. (2008). Young infants' neural processing of objects is affected by eye gaze direction and emotional expression. *PLoS One*, 3(6), e2389.

Hoemann, K., Xu, F., & Barrett, L. F. (2019). Emotion words, emotion concepts, and emotional development in children: A constructionist hypothesis. *Developmental psychology*, 55(9), 1830.

Hoffmann, H., Traue, H. C., Bachmayr, F., & Kessler, H. (2010). Perceived realism of dynamic facial expressions of emotion: Optimal durations for the presentation of emotional onsets and offsets. *Cognition and Emotion*, 24(8), 1369-1376.

Hoffner, C., & Badzinski, D. M. (1989). Children's integration of facial and situational cues to emotion. *Child development*, 411-422.

Homae, F., Watanabe, H., Otobe, T., Nakano, T., Go, T., Konishi, Y., & Taga, G. (2010). Development of global cortical networks in early infancy. *Journal of neuroscience*, 30(14), 4877-4882.

Horlin, C., Falkmer, M., Fitzgerald, P., Leung, D., Ordqvist, A., & Falkmer, T. (2013). The influence of static versus naturalistic stimuli on face processing in children with and without Asperger syndrome or high-functioning autism. *Research in Autism Spectrum Disorders*, 7(12), 1617-1624.

Horning, S. M., Cornwell, R. E., & Davis, H. P. (2012). The recognition of facial expressions: An investigation of the influence of age and cognition. *Aging, Neuropsychology, and Cognition*, 19(6), 657-676.

Horvath, S., Kueser, J. B., Kelly, J., & Borovsky, A. (2022). Difference or delay? Syntax, semantics, and verb vocabulary development in typically developing and late-talking toddlers. *Language Learning and Development*, 18(3), 352-376.

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.

Howe, N., & Recchia, H. (2005). Playmates and teachers: reciprocal and complementary interactions between siblings. *Journal of Family Psychology*, 19(4), 497.

Humphreys, K., Avidan, G., & Behrmann, M. (2007). A detailed investigation of facial expression processing in congenital prosopagnosia as compared to acquired prosopagnosia. *Experimental Brain Research*, 176(2), 356-373.

Hunnius, S. (2022). Early cognitive development: Five lessons from infant learning. In *Oxford research encyclopedia of psychology*.

Huxter, K., Atkin, A. E., & Singhal, A. (2023). Perception of emotion in the facial expressions and body language of athletes. *Social Behavior and Personality: an international journal*, 51(4), 1-12.

Hyde, D. C., Jones, B. L., Flom, R., & Porter, C. L. (2011). Neural signatures of face – voice synchrony in 5 - month - old human infants. *Developmental Psychobiology*, 53(4), 359-370.

Itier, R. J., & Taylor, M. J. (2004). N170 or N1? Spatiotemporal differences between object and face processing using ERPs. *Cerebral cortex*, 14(2), 132-142.

Izard, C. E. (2007). Basic emotions, natural kinds, emotion schemas, and a new paradigm. *Perspectives on psychological science*, 2(3), 260-280.

Jack, R. E., & Schyns, P. G. (2017). Toward a social psychophysics of face communication. *Annual review of psychology*, 68, 269-297.

Jack, R. E., Garrod, O. G., Yu, H., Caldara, R., & Schyns, P. G. (2012). Facial expressions of emotion are not culturally universal. *Proceedings of the National Academy of Sciences*, 109(19), 7241-7244.

Jackson, I., & Sirois, S. (2009). Infant cognition: going full factorial with pupil dilation. *Developmental science*, 12(4), 670-679.

Jaksic, C., & Schlegel, K. (2020). Accuracy in judging others' personalities: the role of emotion recognition, emotion understanding, and trait emotional intelligence. *Journal of Intelligence*, 8(3), 34.

Jayaraman, S., Fausey, C. M., & Smith, L. B. (2015). The faces in infant-perspective scenes change over the first year of life. *PloS one*, 10(5), e0123780.

Jayaraman, S., Fausey, C. M., & Smith, L. B. (2017). Why are faces denser in the visual experiences of younger than older infants?. *Developmental psychology*, 53(1), 38.

Jenkins, R., & Burton, A. M. (2011). Stable face representations. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1571), 1671-1683.

Jenkins, R., Dowsett, A. J., & Burton, A. M. (2018). How many faces do people know?. *Proceedings of the Royal Society B*, 285(1888), 20181319.

Jessen, S., & Kotz, S. A. (2011). The temporal dynamics of processing emotions from vocal, facial, and bodily expressions. *Neuroimage*, 58(2), 665-674.

Jia, S., Wang, S., Hu, C., Webster, P. J., & Li, X. (2021). Detection of genuine and posed facial expressions of emotion: databases and methods. *Frontiers in psychology*, 11, 580287.

Jiang, Z., Recio, G., Li, W., Zhu, P., He, J., & Sommer, W. (2023). The other-race effect in facial expression processing: Behavioral and ERP evidence from a balanced cross-cultural study in women. *International Journal of Psychophysiology*, 183, 53-60.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & psychophysics*, 14, 201-211.

Johnson, M. H. (2005). Subcortical face processing. *Nature Reviews Neuroscience*, 6(10), 766-774.

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, 40(1-2), 1-19.

Johnson, M. H., Senju, A., & Tomalski, P. (2015). The two-process theory of face processing: modifications based on two decades of data from infants and adults. *Neuroscience & Biobehavioral Reviews*, 50, 169-179.

Jones, S. S. (2007). Imitation in infancy: The development of mimicry. *Psychological science*, 18(7), 593-599.

Joosen, K. J., Mesman, J., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2012). Maternal sensitivity to infants in various settings predicts harsh discipline in toddlerhood. *Attachment & human development*, 14(2), 101-117.

Joseph, J. E., Gathers, A. D., & Bhatt, R. S. (2011). Progressive and regressive developmental changes in neural substrates for face processing: testing specific predictions of the Interactive Specialization account. *Developmental science*, 14(2), 227-241.

- Jozwik, K. M., O’Keeffe, J., Storrs, K. R., Guo, W., Golan, T., & Kriegeskorte, N. (2022). Face dissimilarity judgments are predicted by representational distance in morphable and image-computable models. *Proceedings of the National Academy of Sciences*, 119(27), e2115047119.
- Kadosh, K. C., & Johnson, M. H. (2007). Developing a cortex specialized for face perception. *Trends in cognitive sciences*, 11(9), 367-369.
- Kaiser, S., & Wehrle, T. (2001). Facial expressions as indicators of appraisal processes. *Appraisal processes in emotion: Theory, methods, research*, 285-300.
- Kalogeiton, V., & Zisserman, A. (2020). Constrained video face clustering using 1nn relations.
- Kammermeier, M., Duran Perez, L., König, L., & Paulus, M. (2020). Attachment security and attention to facial emotional expressions in preschoolers: An eye - tracking study. *British Journal of Developmental Psychology*, 38(2), 167-185.
- Kanwisher, N., & Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 361(1476), 2109-2128.
- Karahan, S., Yildirim, M. K., Kirtac, K., Rende, F. S., Butun, G., & Ekenel, H. K. (2016, September). How image degradations affect deep CNN-based face recognition?. In 2016 international conference of the biometrics special interest group (BIOSIG) (pp. 1-5). IEEE.
- Karasik, L. B., Tamis - LeMonda, C. S., & Adolph, K. E. (2011). Transition from crawling to walking and infants’ actions with objects and people. *Child development*, 82(4), 1199-1209.

Katainen, S., Räikkönen, K., & Keltikangas - Järvinen, L. (1997). Childhood temperament and mother's child - rearing attitudes: stability and interaction in a three - year follow - up study.

European Journal of Personality, 11(4), 249-265.

Kataja, E. L., Karlsson, L., Leppänen, J. M., Pelto, J., Häikiö, T., Nolvi, S., ... & Karlsson, H.

(2020). Maternal depressive symptoms during the pre - and postnatal periods and infant attention to emotional faces. Child development, 91(2), e475-e480.

Katarzyna, C., Fred, V., & Ami, K. (2010). Limited attentional bias for faces in toddlers with autism spectrum disorders. Archives of general psychiatry, 67(2), 178-185.

Kaufman, G., & Grönlund, A. (2021). Displaying parenthood,(un) doing gender: parental leave, daycare and working-time adjustments in Sweden and the UK. Families, relationships and societies, 10(2), 213-229.

Kawai, N., Kubo, K., Masataka, N., & Hayakawa, S. (2016). Conserved evolutionary history for quick detection of threatening faces. Animal cognition, 19, 655-660.

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.

Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Gibson, A., Smith, M., ... & Pascalis, O. (2005).

Three - month - olds, but not newborns, prefer own - race faces. Developmental science, 8(6), F31-F36.

- Kesler, M. L., Andersen, A. H., Smith, C. D., Avison, M. J., Davis, C. E., Kryscio, R. J., & Blonder, L. X. (2001). Neural substrates of facial emotion processing using fMRI. *Cognitive Brain Research*, 11(2), 213-226.
- Kessels, R. P., Montagne, B., Hendriks, A. W., Perrett, D. I., & De Haan, E. H. (2014). Assessment of perception of morphed facial expressions using the Emotion Recognition Task: Normative data from healthy participants aged 8–75. *Journal of neuropsychology*, 8(1), 75-93.
- Kiel, E. J., Price, N. N., & Premo, J. E. (2020). Maternal comforting behavior, toddlers' dysregulated fear, and toddlers' emotion regulatory behaviors. *Emotion*, 20(5), 793.
- Kim, G., & Kwak, K. (2011). Uncertainty matters: Impact of stimulus ambiguity on infant social referencing. *Infant and Child Development*, 20(5), 449-463.
- Kim, H. I., & Johnson, S. P. (2014). Detecting 'infant - directedness' in face and voice. *Developmental Science*, 17(4), 621-627.
- Kim, M., Kim, J., & Kim, S. Y. (2022). Attentional Bias to Emotional Stimuli and Effects of Anxiety on the Bias in Neurotypical Adults and Adolescents. *Science of Emotion and Sensibility*, 25(4), 107-118.
- Kim, W., Lee, S., Kim, S., Jo, S., Yoo, C., Hwang, I., ... & Song, J. (2020). Dyadic Mirror: Everyday Second-person Live-view for Empathetic Reflection upon Parent-child Interaction. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 4(3), 1-29.

Kiuru, N., Hirvonen, R., & Ahonen, T. (2019). Assessing temperament among Finnish early adolescents and their parents: psychometric properties of the short forms of the temperament questionnaires. *Merrill-Palmer Quarterly*, 65(3), 294-328.

Klinnert, M. D., Emde, R. N., Butterfield, P., & Campos, J. J. (1986). Social referencing: The infant's use of emotional signals from a friendly adult with mother present. *Developmental Psychology*, 22(4), 427.

Konishi, Y., Takaya, R., Kimura, K., Konishi, K., Fujii, Y., Saito, M., & Sudo, M. (1994). Development of posture in prone and supine positions during the prenatal period in low risk preterm infants. *Archives of Disease in Childhood-Fetal and Neonatal Edition*, 70(3), F188-F191.

Kornreich, C., Philippot, P., Foisy, M. L., Blairy, S., Raynaud, E., Dan, B., ... & Verbanck, P. (2002). Impaired emotional facial expression recognition is associated with interpersonal problems in alcoholism. *Alcohol and alcoholism*, 37(4), 394-400.

Kosakowski, H. L., Cohen, M. A., Herrera, L., Nichoson, I., Kanwisher, N., & Saxe, R. (2024). Cortical Face-Selective Responses Emerge Early in Human Infancy. *eneuro*, 11(7).

Kostyrka - Allchorne, K., Wass, S. V., & Sonuga - Barke, E. J. (2020). Research Review: Do parent ratings of infant negative emotionality and self - regulation predict psychopathology in childhood and adolescence? A systematic review and meta - analysis of prospective longitudinal studies. *Journal of Child Psychology and Psychiatry*, 61(4), 401-416.

- Kotsoni, E., de Haan, M., & Johnson, M. H. (2001). Categorical perception of facial expressions by 7-month-old infants. *Perception*, 30(9), 1115-1125.
- Kret, M. E. (2015). Emotional expressions beyond facial muscle actions. A call for studying autonomic signals and their impact on social perception. *Frontiers in psychology*, 6, 711.
- Kret, M. E., Prochazkova, E., Sterck, E. H., & Clay, Z. (2020). Emotional expressions in human and non-human great apes. *Neuroscience & Biobehavioral Reviews*, 115, 378-395.
- Kret, M. E., Prochazkova, E., Sterck, E. H., & Clay, Z. (2020). Emotional expressions in human and non-human great apes. *Neuroscience & Biobehavioral Reviews*.
- 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.
- Kretch, K. S., Franchak, J. M., & Adolph, K. E. (2014). Crawling and walking infants see the world differently. *Child development*, 85(4), 1503-1518.
- Kreutzer, M. A., & Charlesworth, W. R. (1973). Infants' Reactions to Different Expressions of Emotions.
- Krogh-Jespersen, S., & Woodward, A. L. (2018). Reaching the goal: Active experience facilitates 8-month-old infants' prospective analysis of goal-based actions. *Journal of Experimental Child Psychology*, 171, 31-45.
- Krumhuber, E. G., Küster, D., Namba, S., Shah, D., & Calvo, M. G. (2021). Emotion recognition from posed and spontaneous dynamic expressions: Human observers versus machine analysis. *Emotion*, 21(2), 447.

Kubicek, C., Jovanovic, B., & Schwarzer, G. (2017). How manual object exploration is associated with 7 - to 8 - month - old infants' visual prediction abilities in spatial object processing. *Infancy*, 22(6), 857-873.

Kuchuk, A., Vibbert, M., & Bornstein, M. H. (1986). The perception of smiling and its experiential correlates in three-month-old infants. *Child development*, 1054-1061.

Kuefner, D., De Heering, A., Jacques, C., Palmero-Soler, E., & Rossion, B. (2010). Early visually evoked electrophysiological responses over the human brain (P1, N170) show stable patterns of face-sensitivity from 4 years to adulthood. *Frontiers in human neuroscience*, 67.

Küntzler, T., Höfling, T. T. A., & Alpers, G. W. (2021). Automatic facial expression recognition in standardized and non-standardized emotional expressions. *Frontiers in psychology*, 12, 1086.

Labouvie-Vief, G., Grünh, D., & Mouras, H. (2009). Dynamic emotion-cognition interactions in adult development: Arousal, stress, and the processing of affect.

Lambert, S. R. (2004). Treatment of congenital cataract. *British journal of ophthalmology*, 88(7), 854-855.

Landa, R. J., Holman, K. C., & Garrett-Mayer, E. (2007). Social and communication development in toddlers with early and later diagnosis of autism spectrum disorders. *Archives of general psychiatry*, 64(7), 853-864.

Lander, K., Bruce, V., & Bindemann, M. (2018). Use-inspired basic research on individual differences in face identification: Implications for criminal investigation and security. *Cognitive Research: Principles and Implications*, 3(1), 1-13.

- Lange, J., Heerdink, M. W., & Van Kleef, G. A. (2022). Reading emotions, reading people: Emotion perception and inferences drawn from perceived emotions. *Current opinion in psychology*, 43, 85-90.
- Laukka, P. (2005). Categorical perception of vocal emotion expressions. *Emotion*, 5(3), 277.
- Lawrence, J., Haszard, J. J., Taylor, B., Galland, B., Gray, A., Sayers, R., ... & Taylor, R. (2021). A longitudinal study of parental discipline up to 5 years. *Journal of Family Studies*, 27(4), 589-606.
- Leahy-Warren, P., & McCarthy, G. (2007). Postnatal depression: prevalence, mothers' perspectives, and treatments. *Archives of psychiatric nursing*, 21(2), 91-100.
- Lee, D. H., & Anderson, A. K. (2017). Form and function of facial expressive origins. *The science of facial expression*, 173-194.
- Lee, S. B., Koo, S. J., Song, Y. Y., Lee, M. K., Jeong, Y. J., Kwon, C., ... & An, S. K. (2014). Theory of mind as a mediator of reasoning and facial emotion recognition: findings from 200 healthy people. *Psychiatry Investigation*, 11(2), 105.
- Lee, T. H., Choi, J. S., & Cho, Y. S. (2012). Context modulation of facial emotion perception differed by individual difference. *PLOS one*, 7(3), e32987.
- Leinbach, M. D., & Fagot, B. I. (1993). Categorical habituation to male and female faces: Gender schematic processing in infancy. *Infant behavior and development*, 16(3), 317-332.
- Leppänen, J. M., & Nelson, C. A. (2006). The development and neural bases of facial emotion recognition. *Advances in child development and behavior*, 34, 207-246.

- Leppänen, J. M., & Nelson, C. A. (2012). Early development of fear processing. *Current Directions in Psychological Science*, 21(3), 200-204.
- Leppänen, J. M., Moulson, M. C., Vogel - Farley, V. K., & Nelson, C. A. (2007). An ERP study of emotional face processing in the adult and infant brain. *Child development*, 78(1), 232-245.
- Levitzky, S., & Cooper, R. (2000). Infant colic syndrome—maternal fantasies of aggression and infanticide. *Clinical pediatrics*, 39(7), 395-400.
- Lewis, C., Freeman, N. H., Hagestadt, C., & Douglas, H. (1994). Narrative access and production in preschoolers' false belief reasoning. *Cognitive Development*, 9(4), 397-424.
- Lewis, M. (2012). *Social cognition and the acquisition of self*. Springer Science & Business Media.
- Lewis, M., & Brooks - Gunn, J. (1979). Toward a theory of social cognition: The development of self. *New directions for child and adolescent development*, 1979(4), 1-20.
- Lewis, M., & Takahashi, K. (Eds.). (2005). *Beyond the dyad: Conceptualization of social networks*. Karger Medical and Scientific Publishers.
- Li, C., Li, L., Jiang, H., Weng, K., Geng, Y., Li, L., Ke, Z., Li, Q., Cheng, M., Nie, W. and Li, Y., 2022. YOLOv6: a single-stage object detection framework for industrial applications. *arXiv preprint arXiv:2209.02976*.
- Lin, Gao-Xian, Logan Hansotte, Dorota Szczygiel, Loes Meeussen, Isabelle Roskam, and Moïra Mikolajczak. "Parenting with a smile: Display rules, regulatory effort, and parental burnout." *Journal of Social and Personal Relationships* 38, no. 9 (2021): 2701-2721.

Linnankoski, I., Laakso, M., Aulanko, R., & Leinonen, L. (1994). Recognition of emotions in macaque vocalizations by children and adults. *Language & Communication*, 14(2), 183-192.

Liu, P., Kryski, K. R., Smith, H. J., Joannisse, M. F., & Hayden, E. P. (2020). Transactional relations between early child temperament, structured parenting, and child outcomes: A three-wave longitudinal study. *Development and psychopathology*, 32(3), 923-933.

Liu, S., Qi, L., Qin, H., Shi, J. and Jia, J., 2018. Path aggregation network for instance segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 8759-8768).

Liu, S., Xiao, N. G., Quinn, P. C., Zhu, D., Ge, L., Pascalis, O., & Lee, K. (2015). Asian infants show preference for own-race but not other-race female faces: The role of infant caregiving arrangements. *Frontiers in Psychology*, 6, 593.

Lo, R. F., & Mar, R. A. (2022). Having siblings is associated with better mentalizing abilities in adults. *Cognitive Development*, 63, 101193.

LoBue, V., & Ogren, M. (2022). How the emotional environment shapes the emotional life of the child. *Policy insights from the behavioral and brain sciences*, 9(1), 137-144.

Lombardo, M. V., Barnes, J. L., Wheelwright, S. J., & Baron-Cohen, S. (2007). Self-referential cognition and empathy in autism. *PloS one*, 2(9), e883.

Long, B., Kachergis, G., Agrawal, K., & Frank, M. C. (2020). Detecting social information in a dense database of infants' natural visual experience. In *CogSci*.

- Lord, C., Brugha, T. S., Charman, T., Cusack, J., Dumas, G., Frazier, T., ... & Veenstra-VanderWeele, J. (2020). Autism spectrum disorder. *Nature reviews Disease primers*, 6(1), 1-23.
- Loth, E., Garrido, L., Ahmad, J., Watson, E., Duff, A., & Duchaine, B. (2018). Facial expression recognition as a candidate marker for autism spectrum disorder: how frequent and severe are deficits?. *Molecular autism*, 9, 1-11.
- Lozier, L. M., Vanmeter, J. W., & Marsh, A. A. (2014). Impairments in facial affect recognition associated with autism spectrum disorders: a meta-analysis. *Development and psychopathology*, 26(4), 933-945.
- Ludemann, P. M., & Nelson, C. A. (1988). Categorical representation of facial expressions by 7-month-old infants. *Developmental Psychology*, 24(4), 492.
- Lutkiewicz, K., Bieleninik, Ł., Cieślak, M., & Bidzan, M. (2020). Maternal–infant bonding and its relationships with maternal depressive symptoms, stress and anxiety in the early postpartum period in a Polish sample. *International journal of environmental research and public health*, 17(15), 5427.
- Maack, J. K., Bohne, A., Nordahl, D., Livsdatter, L., Lindahl, Å. A., Øvervoll, M., ... & Pfuhl, G. (2017). The Tromsø Infant Faces Database (TIF): development, validation and application to assess parenting experience on clarity and intensity ratings. *Frontiers in psychology*, 409.
- Macari, S. L., Koller, J., Campbell, D. J., & Chawarska, K. (2017). Temperamental markers in toddlers with autism spectrum disorder. *Journal of Child Psychology and Psychiatry*, 58(7), 819-828.

MacGowan, T. L., Colonnese, C., Nikolić, M., & Schmidt, L. A. (2022). Expressions of shyness and theory of mind in children: A psychophysiological study. *Cognitive Development*, 61, 101138.

Maclean, P. C., Erickson, S. J., & Lowe, J. R. (2009). Comparing emotional reactivity and regulation in infants born ELGA and VLGA. *Infant Behavior and Development*, 32(3), 336-339.

MacLin, M. K., Downs, C., MacLin, O. H., & Caspers, H. M. (2009). The effect of defendant facial expression on mock juror decision-making: The power of remorse. *North American Journal of Psychology*, 11(2), 323-332.

Maher, S., Ekstrom, T., & Chen, Y. (2014). Greater perceptual sensitivity to happy facial expression. *Perception*, 43(12), 1353-1364.

Malatesta, C. Z., & Haviland, J. M. (1982). Learning display rules: The socialization of emotion expression in infancy. *Child development*, 991-1003.

Malsert, J., Palama, A., & Gentaz, E. (2020). Emotional facial perception development in 7, 9 and 11 year-old children: The emergence of a silent eye-tracked emotional other-race effect. *Plos one*, 15(5), e0233008.

Mäntymaa, M., Puura, K., Luoma, I., Salmelin, R. K., & Tamminen, T. (2006). Mother's early perception of her infant's difficult temperament, parenting stress and early mother–infant interaction. *Nordic journal of psychiatry*, 60(5), 379-386.

Mark, I. L. V. D., IJzendoorn, M. H. V., & Bakermans - Kranenburg, M. J. (2002). Development of empathy in girls during the second year of life: Associations with parenting, attachment, and temperament. *Social development*, 11(4), 451-468.

Markestad, T., Kaaresen, P. I., Rønnestad, A., Reigstad, H., Lossius, K., Medbø, S., ... & Irgens, L. M. (2005). Early death, morbidity, and need of treatment among extremely premature infants. *Pediatrics*, 115(5), 1289-1298.

Marsden, J., Glazebrook, C., Tully, R., & Völlm, B. (2019). Do adult males with antisocial personality disorder (with and without co-morbid psychopathy) have deficits in emotion processing and empathy? A systematic review. *Aggression and violent behavior*, 48, 197-217.

Marsh, A. A., Elfenbein, H. A., & Ambady, N. (2003). Nonverbal “accents” cultural differences in facial expressions of emotion. *Psychological Science*, 14(4), 373-376.

Martin - Key, N. A., Graf, E. W., Adams, W. J., & Fairchild, G. (2018). Facial emotion recognition and eye movement behaviour in conduct disorder. *Journal of Child Psychology and Psychiatry*, 59(3), 247-257.

Martinos, M., Matheson, A., & de Haan, M. (2012). Links between infant temperament and neurophysiological measures of attention to happy and fearful faces. *Journal of Child Psychology and Psychiatry*, 53(11), 1118-1127.

Martins, A. T., Faísca, L., Vieira, H., & Gonçalves, G. (2019). Emotional recognition and empathy both in deaf and blind adults. *The Journal of Deaf Studies and Deaf Education*, 24(2), 119-127.

Marwick, K., & Hall, J. (2008). Social cognition in schizophrenia: a review of face processing. *British Medical Bulletin*, 88(1), 43-58.

Masi, I., Trần, A. T., Hassner, T., Leksut, J. T., & Medioni, G. (2016). Do we really need to collect millions of faces for effective face recognition?. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part V* 14 (pp. 579-596). Springer International Publishing.

Mason, D. A., & Frick, P. J. (1994). The heritability of antisocial behavior: A meta-analysis of twin and adoption studies. *Journal of Psychopathology and Behavioral Assessment*, 16(4), 301-323.

Mastropieri, D., & Turkewitz, G. (1999). Prenatal experience and neonatal responsiveness to vocal expressions of emotion. *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology*, 35(3), 204-214.

Matsumoto, D. (1990). Cultural similarities and differences in display rules. *Motivation and emotion*, 14, 195-214.

Matsumoto, D. (1990). Cultural similarities and differences in display rules. *Motivation and emotion*, 14(3), 195-214.

Matsumoto, D. (1992). American-Japanese cultural differences in the recognition of universal facial expressions. *Journal of cross-cultural psychology*, 23(1), 72-84.

Matsumoto, D., Keltner, D., Shiota, M. N., O'Sullivan, M., & Frank, M. (2008). Facial expressions of emotion.

- Maurer, D., Lewis, T. L., & MONDLOCH, C. (2008). 25 Plasticity of the Visual System. *Handbook of developmental cognitive neuroscience*, 415.
- Maurer, D., Lewis, T. L., Brent, H. P., & Levin, A. V. (1999). Rapid improvement in the acuity of infants after visual input. *Science*, 286(5437), 108-110.
- Mauro, R., & Kubovy, M. (1992). Caricature and face recognition. *Memory & Cognition*, 20(4), 433-440.
- Mayo, L. M., & Heilig, M. (2019). In the face of stress: Interpreting individual differences in stress-induced facial expressions. *Neurobiology of stress*, 10, 100166.
- McAlister, A., & Peterson, C. C. (2006). Mental playmates: Siblings, executive functioning and theory of mind. *British Journal of Developmental Psychology*, 24(4), 733-751.
- McClure, E. B. (2000). A meta-analytic review of sex differences in facial expression processing and their development in infants, children, and adolescents. *Psychological bulletin*, 126(3), 424.
- McConachie, H. R. (1976). Developmental prosopagnosia. A single case report. *Cortex*, 12(1), 76-82.
- McConnell, B. A., & Bryson, S. E. (2005). Visual attention and temperament: Developmental data from the first 6 months of life. *Infant Behavior and Development*, 28(4), 537-544.
- McKone, E., & Robbins, R. (2011). Are faces special. *Oxford handbook of face perception*, 149-176.
- Meadows, S. (2013). Parenting behaviour and children's cognitive development. *Psychology Press*.

Meeren, H. K., van Heijnsbergen, C. C., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proceedings of the National Academy of Sciences*, 102(45), 16518-16523.

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law*, 7(1), 3.

Meltzoff, A. N. (1988). Imitation of televised models by infants. *Child development*, 59(5), 1221.

Meltzoff, A. N., & Marshall, P. J. (2020). Importance of body representations in social-cognitive development: new insights from infant brain science. *Progress in brain research*, 254, 25-48.

Meltzoff, A. N., & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science*, 198(4312), 75-78.

Mendoza, R., Cabral-Calderin, Y., Domínguez, M., Garcia, A., Borrego, M., Caballero, A., ... &

Mendoza, R., Cabral-Calderin, Y., Domínguez, M., Garcia, A., Borrego, M., Caballero, A., ... & Reyes, M. M. (2011). Impairment of emotional expression recognition in schizophrenia: A Cuban familial association study. *Psychiatry research*, 185(1-2), 44-48.

Messinger, D. S. (2002). Positive and negative: Infant facial expressions and emotions. *Current Directions in Psychological Science*, 11(1), 1-6.

Messinger, D. S., & Fogel, A. (1998). Give and take: The development of conventional infant gestures. *Merrill-Palmer Quarterly* (1982-), 566-590.

- Miller, E. J., Krumhuber, E. G., & Dawel, A. (2022). Observers perceive the Duchenne marker as signaling only intensity for sad expressions, not genuine emotion. *Emotion*, 22(5), 907.
- Mink, D., Henning, A., & Aschersleben, G. (2014). Infant shy temperament predicts preschoolers theory of mind. *Infant Behavior and Development*, 37(1), 66-75.
- Missana, M., Atkinson, A. P., & Grossmann, T. (2015). Tuning the developing brain to emotional body expressions. *Developmental Science*, 18(2), 243-253.
- Mondloch, C. J., Lewis, T. L., Budreau, D. R., Maurer, D., Dannemiller, J. L., Stephens, B. R., & Kleiner-Gathercoal, K. A. (1999). Face perception during early infancy. *Psychological Science*, 10(5), 419-422.
- Moore, C. (2008). The development of gaze following. *Child Development Perspectives*, 2(2), 66-70.
- Moore, C., & Povinelli, D. J. (2007). Differences in how 12 - and 24 - month - olds interpret the gaze of adults. *Infancy*, 11(3), 215-231.
- Morita, T., Slaughter, V., Katayama, N., Kitazaki, M., Kakigi, R., & Itakura, S. (2012). Infant and adult perceptions of possible and impossible body movements: An eye-tracking study. *Journal of Experimental Child Psychology*, 113(3), 401-414.
- Morningstar, M., Nelson, E. E., & Dirks, M. A. (2018). Maturation of vocal emotion recognition: Insights from the developmental and neuroimaging literature. *Neuroscience & Biobehavioral Reviews*, 90, 221-230.

- Morton, J., & Johnson, M. H. (1991). CONSPEC and CONLERN: a two-process theory of infant face recognition. *Psychological review*, 98(2), 164.
- Motley, M. T., & Camden, C. T. (1988). Facial expression of emotion: A comparison of posed expressions versus spontaneous expressions in an interpersonal communication setting. *Western Journal of Communication (includes Communication Reports)*, 52(1), 1-22.
- Moulson, M. C., Shutts, K., Fox, N. A., Zeanah, C. H., Spelke, E. S., & Nelson, C. A. (2015). Effects of early institutionalization on the development of emotion processing: a case for relative sparing?. *Developmental science*, 18(2), 298-313.
- Mumme, D. L., & Fernald, A. (2003). The infant as onlooker: Learning from emotional reactions observed in a television scenario. *Child development*, 74(1), 221-237.
- Murray, A. D. (1979). Infant crying as an elicitor of parental behavior: an examination of two models. *Psychological bulletin*, 86(1), 191.
- Narvekar, N., Carter Leno, V., Pasco, G., Johnson, M. H., Jones, E. J., & Charman, T. (2022). A prospective study of associations between early fearfulness and perceptual sensitivity and later restricted and repetitive behaviours in infants with typical and elevated likelihood of autism. *Autism*, 26(8), 1947-1958.
- Nava, E., Romano, D., Grassi, M., & Turati, C. (2016). Skin conductance reveals the early development of the unconscious processing of emotions. *Cortex*, 84, 124-131.
- Nelson, C. A. (1987). The recognition of facial expressions in the first two years of life: Mechanisms of development. *Child development*, 889-909.

Nelson, C. A., & De Haan, M. (1997). A neurobehavioral approach to the recognition of facial expressions in infancy. *The psychology of facial expression*, 176-204.

Nelson, C. A., & Horowitz, F. D. (1983). The perception of facial expressions and stimulus motion by two-and five-month-old infants using holographic stimuli. *Child Development*, 868-877.

Nelson, C. A., Morse, P. A., & Leavitt, L. A. (1979). Recognition of facial expressions by seven-month-old infants. *Child development*, 1239-1242.

Neppl, T. K., Donnellan, M. B., Scaramella, L. V., Widaman, K. F., Spilman, S. K., Ontai, L. L., & Conger, R. D. (2010). Differential stability of temperament and personality from toddlerhood to middle childhood. *Journal of research in personality*, 44(3), 386-396.

Neuhaus, E., Jones, E. J., Barnes, K., Sterling, L., Estes, A., Munson, J., ... & Webb, S. J. (2016). The relationship between early neural responses to emotional faces at age 3 and later autism and anxiety symptoms in adolescents with autism. *Journal of autism and developmental disorders*, 46(7), 2450-2463.

Nguyen, V. T., & Cunnington, R. (2014). The superior temporal sulcus and the N170 during face processing: single trial analysis of concurrent EEG–fMRI. *Neuroimage*, 86, 492-502.

Nielsen, M., Dissanayake, C., & Kashima, Y. (2003). A longitudinal investigation of self–other discrimination and the emergence of mirror self-recognition. *Infant Behavior and Development*, 26(2), 213-226.

Nitta, H., & Hashiya, K. (2021). Self-face perception in 12-month-old infants: A study using the morphing technique. *Infant Behavior and Development*, 62, 101479.

Noyes, E., Davis, J. P., Petrov, N., Gray, K. L., & Ritchie, K. L. (2021). The effect of face masks and sunglasses on identity and expression recognition with super-recognizers and typical observers. *Royal Society open science*, 8(3), 201169.

Noyes, E., Phillips, P. J., & O'Toole, A. J. (2017). What is a super-recogniser?. In *Face processing: Systems, disorders and cultural differences* (pp. 173-201). Nova Science Publishers Inc.

Oakes, L. M. (2023). Understanding developmental cascades and experience: Diversity matters. *Infancy*, 28(3), 492-506.

Öberg, S. (2017). Too many is not enough: Studying how children are affected by their number of siblings and resource dilution in families. *The History of the Family*, 22(2-3), 157-174.

Oddi, K. B., Murdock, K. W., Vadnais, S., Bridgett, D. J., & Gartstein, M. A. (2013). Maternal and infant temperament characteristics as contributors to parenting stress in the first year postpartum. *Infant and Child Development*, 22(6), 553-579.

OGDEN, J. A. (1985). Autotopagnosia: Occurrence in a patient without nominal aphasia and with an intact ability to point to parts of animals and objects. *Brain*, 108(4), 1009-1022.

Öhman, A., Lundqvist, D., & Esteves, F. (2001). The face in the crowd revisited: a threat advantage with schematic stimuli. *Journal of personality and social psychology*, 80(3), 381.

Onal Ertugrul, I., Ahn, Y. A., Bilalpur, M., Messinger, D. S., Speltz, M. L., & Cohn, J. F. (2023). Infant AFAR: Automated facial action recognition in infants. *Behavior research methods*, 55(3), 1024-1035.

Onitsuka, T., Oribe, N., & Kanba, S. (2013). Neurophysiological findings in patients with bipolar disorder. In *Supplements to Clinical neurophysiology* (Vol. 62, pp. 197-206). Elsevier.

Oostenbroek, J., Suddendorf, T., Nielsen, M., Redshaw, J., Kennedy-Costantini, S., Davis, J., ... & Slaughter, V. (2016). Comprehensive longitudinal study challenges the existence of neonatal imitation in humans. *Current Biology*, 26(10), 1334-1338.

Oruc, I., Shafai, F., Murthy, S., Lages, P., & Ton, T. (2019). The adult face-diet: A naturalistic observation study. *Vision research*, 157, 222-229.

Otten, M., Seth, A. K., & Pinto, Y. (2017). A social Bayesian brain: How social knowledge can shape visual perception. *Brain and cognition*, 112, 69-77.

Palermo, R., & Coltheart, M. (2004). Photographs of facial expression: Accuracy, response times, and ratings of intensity. *Behavior Research Methods, Instruments, & Computers*, 36, 634-638.

Palermo, R., O'Connor, K. B., Davis, J. M., Irons, J., & McKone, E. (2013). New tests to measure individual differences in matching and labelling facial expressions of emotion, and their association with ability to recognise vocal emotions and facial identity. *PloS one*, 8(6), e68126.

Palumbo, L., & Jellema, T. (2013). Beyond face value: does involuntary emotional anticipation shape the perception of dynamic facial expressions?. *PloS one*, 8(2), e56003.

Parr, L. A. (2011). The evolution of face processing in primates. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1571), 1764-1777.

Pascalis, O., & Kelly, D. J. (2009). The origins of face processing in humans: Phylogeny and ontogeny. *Perspectives on psychological science*, 4(2), 200-209.

Pascalis, O., de Haan, M., & Nelson, C. A. (2002). Is face processing species-specific during the first year of life?. *Science*, 296(5571), 1321-1323.

Pascalis, O., de Schonen, S., Morton, J., Deruelle, C., & Fabre-Grenet, M. (1995). Mother's face recognition by neonates: A replication and an extension. *Infant Behavior and Development*, 18(1), 79-85.

Pascalis, O., Scott, L. S., Kelly, D. J., Shannon, R. W., Nicholson, E., Coleman, M., & Nelson, C. A. (2005). Plasticity of face processing in infancy. *Proceedings of the national academy of sciences*, 102(14), 5297-5300.

Paulmann, S., Bleichner, M., & Kotz, S. A. (2013). Valence, arousal, and task effects in emotional prosody processing. *Frontiers in Psychology*, 4, 345.

Paulus, M., & Fikkert, P. (2014). Conflicting social cues: Fourteen-and 24-month-old infants' reliance on gaze and pointing cues in word learning. *Journal of Cognition and Development*, 15(1), 43-59.

Pelphrey, K. A., Sasson, N. J., Reznick, J. S., Paul, G., Goldman, B. D., & Piven, J. (2002). Visual scanning of faces in autism. *Journal of autism and developmental disorders*, 32(4), 249-261.

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.

Peltola, M. J., Leppänen, J. M., Palokangas, T., & Hietanen, J. K. (2008). Fearful faces modulate looking duration and attention disengagement in 7 - month - old infants. *Developmental science*, 11(1), 60-68.

Pérez-Edgar, K., Morales, S., LoBue, V., Taber-Thomas, B. C., Allen, E. K., Brown, K. M., & Buss, K. A. (2017). The impact of negative affect on attention patterns to threat across the first 2 years of life. *Developmental psychology*, 53(12), 2219.

Peterson, C. C. (2016). Empathy and theory of mind in deaf and hearing children. *Journal of deaf studies and deaf education*, 21(2), 141-147.

Petridis, S., Martinez, B., & Pantic, M. (2013). The MAHNOB laughter database. *Image and Vision Computing*, 31(2), 186-202.

Pfister, T., Li, X., Zhao, G., & Pietikäinen, M. (2011, November). Differentiating spontaneous from posed facial expressions within a generic facial expression recognition framework. In 2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops) (pp. 868-875). IEEE.

Philip, R. C. M., Whalley, H. C., Stanfield, A. C., Sprengelmeyer, R., Santos, I. M., Young, A. W., ... & Hall, J. (2010). Deficits in facial, body movement and vocal emotional processing in autism spectrum disorders. *Psychological medicine*, 40(11), 1919-1929.

Philip, R. C., Whalley, H. C., Stanfield, A. C., Sprengelmeyer, R., Santos, I. M., Young, A. W., ... & Hall, J. (2010). Deficits in facial, body movement and vocal emotional processing in autism spectrum disorders. *Psychological medicine*, 40(11), 1919-1929.

Phillips, A. T., Wellman, H. M., & Spelke, E. S. (2002). Infants' ability to connect gaze and emotional expression to intentional action. *Cognition*, 85(1), 53-78.

Picton, T. W., Hillyard, S. A., Krausz, H. I., & Galambos, R. (1974). Human auditory evoked potentials. I: Evaluation of components. *Electroencephalography and clinical neurophysiology*, 36, 179-190.

Piek, J. P. (2006). *Infant motor development* (Vol. 10). Human kinetics.

Pijl, M. K. J., Bussu, G., Charman, T., Johnson, M. H., Jones, E. J., Pasco, G., ... & Buitelaar, J. K. (2019). Temperament as an early risk marker for Autism Spectrum Disorders? A longitudinal study of high-risk and low-risk infants. *Journal of Autism and Developmental Disorders*, 49(5), 1825-1836.

Pinto, A. I., Pessanha, M., & Aguiar, C. (2013). Effects of home environment and center-based child care quality on children's language, communication, and literacy outcomes. *Early Childhood Research Quarterly*, 28(1), 94-101.

Plank, I. S., Christiansen, L. N., Kunas, S. L., Dziobek, I., & Bermpohl, F. (2022). Mothers need more information to recognise associated emotions in child facial expressions. *Cognition and Emotion*, 36(7), 1299-1312.

Plate, R. C., Bloomberg, Z., Bolt, D. M., Bechner, A. M., Roeber, B. J., & Pollak, S. D. (2019). Abused children experience high anger exposure. *Frontiers in psychology*, 10, 440.

Pollak, S. D., & Tolley-Schell, S. A. (2003). Selective attention to facial emotion in physically abused children. *Journal of abnormal psychology*, 112(3), 323.

Pollux, P. M., Craddock, M., & Guo, K. (2019). Gaze patterns in viewing static and dynamic body expressions. *Acta psychologica*, 198, 102862.

Porter, R. E., & Samovar, L. A. (1996). Cultural influences on emotional expression: implications for intercultural communication. In *Handbook of communication and emotion* (pp. 451-472). Academic Press.

Potapova, N. V., Gartstein, M. A., & Bridgett, D. J. (2014). Paternal influences on infant temperament: Effects of father internalizing problems, parenting-related stress, and temperament. *Infant Behavior and Development*, 37(1), 105-110.

Poulin-Dubois, D., Hastings, P. D., Chiarella, S. S., Geangu, E., Hauf, P., Ruel, A., & Johnson, A. (2018). The eyes know it: Toddlers' visual scanning of sad faces is predicted by their theory of mind skills. *PloS one*, 13(12), e0208524.

Powers, A., Fani, N., Murphy, L., Briscione, M., Bradley, B., Tone, E. B., ... & Jovanovic, T. (2019). Attention bias toward threatening faces in women with PTSD: eye tracking correlates by symptom cluster. *European journal of psychotraumatology*, 10(1), 1568133.

Prince, E. B., Martin, K. B., Messinger, D. S., & Allen, M. (2015). Facial action coding system. *Environmental Psychology & Nonverbal Behavior*.

Puce, A., & Perrett, D. (2003). Electrophysiology and brain imaging of biological motion. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 358(1431), 435-445.

Putnam, S. P., Ellis, L. K., & Rothbart, M. K. (2001). The structure of temperament from infancy through adolescence. *Advances in research on temperament*, 165, 182.

Putnam, S. P., Gartstein, M. A., & Rothbart, M. K. (2006). Measurement of fine-grained aspects of toddler temperament: The Early Childhood Behavior Questionnaire. *Infant behavior and development*, 29(3), 386-401.

Putnam, S. P., Rothbart, M. K., & Gartstein, M. A. (2008). Homotypic and heterotypic continuity of fine - grained temperament during infancy, toddlerhood, and early childhood. *Infant and Child Development: An International Journal of Research and Practice*, 17(4), 387-405.

Quinn, P. C., & Tanaka, J. W. (2009). Infants' processing of featural and configural information in the upper and lower halves of the face. *Infancy*, 14(4), 474-487.

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.

Rajhans, P., Jessen, S., Missana, M., & Grossmann, T. (2016). Putting the face in context: Body expressions impact facial emotion processing in human infants. *Developmental cognitive neuroscience*, 19, 115-121.

- Rajhans, P., Missana, M., Krol, K. M., & Grossmann, T. (2015). The association of temperament and maternal empathy with individual differences in infants' neural responses to emotional body expressions. *Development and Psychopathology*, 27(4pt1), 1205-1216.
- Ramon, M., Bobak, A. K., & White, D. (2019). Super - recognizers: From the lab to the world and back again. *British journal of psychology*, 110(3), 461-479.
- Ramsey, J. L., Langlois, J. H., & Marti, N. C. (2005). Infant categorization of faces: Ladies first. *Developmental Review*, 25(2), 212-246.
- Rat-Fischer, L., O'regan, J. K., & Fagard, J. (2012). The emergence of tool use during the second year of life. *Journal of experimental child psychology*, 113(3), 440-446.
- Ravicz, M. M., Perdue, K. L., Westerlund, A., Vanderwert, R. E., & Nelson, C. A. (2015). Infants' neural responses to facial emotion in the prefrontal cortex are correlated with temperament: a functional near-infrared spectroscopy study. *Frontiers in psychology*, 6, 922.
- Reed, C. L., Beall, P. M., Stone, V. E., Kopeliov, L., Pulham, D. J., & Hepburn, S. L. (2007). Brief report: Perception of body posture—what individuals with autism spectrum disorder might be missing. *Journal of Autism and Developmental Disorders*, 37, 1576-1584.
- Reed, C. L., Stone, V. E., Bozova, S., & Tanaka, J. (2003). The body-inversion effect. *Psychological science*, 14(4), 302-308.
- Reid, M. J., O'Leary, S. G., & Wolff, L. S. (1994). Effects of maternal distraction and reprimands on toddlers' transgressions and negative affect. *Journal of Abnormal Child Psychology*, 22, 237-245.

Renneberg, B., Heyn, K., Gebhard, R., & Bachmann, S. (2005). Facial expression of emotions in borderline personality disorder and depression. *Journal of behavior therapy and experimental psychiatry*, 36(3), 183-196.

Rennels, J. L., Kayl, A. J., & Kulhanek, K. M. (2020). Individual Differences in Infants' Temperament Affect Face Processing. *Brain Sciences*, 10(8), 474.

Reyes, M. M. (2011). Impairment of emotional expression recognition in schizophrenia: A Cuban familial association study. *Psychiatry research*, 185(1-2), 44-48.

Reynolds, G. D. (2015). Infant visual attention and object recognition. *Behavioural brain research*, 285, 34-43.

Rhodes, G., Jeffery, L., Watson, T. L., Clifford, C. W., & Nakayama, K. (2003). Fitting the mind to the world: Face adaptation and attractiveness aftereffects. *Psychological science*, 14(6), 558-566.

Rice, A., Phillips, P. J., Natu, V., An, X., & O'Toole, A. J. (2013). Unaware person recognition from the body when face identification fails. *Psychological science*, 24(11), 2235-2243.

Richards, J. E. (2003). Attention affects the recognition of briefly presented visual stimuli in infants: An ERP study. *Developmental Science*, 6(3), 312-328.

Riddoch, M. J., Johnston, R. A., Bracewell, R. M., Boutsen, L., & Humphreys, G. W. (2008). Are faces special? A case of pure prosopagnosia. *Cognitive Neuropsychology*, 25(1), 3-26.

Rigato, S., Menon, E., Johnson, M. H., & Farroni, T. (2011). The interaction between gaze direction and facial expressions in newborns. *European Journal of Developmental Psychology*, 8(5), 624-636.

Rigato, S., Vrticka, P., Stets, M., & Holmboe, K. (2024). Mother-infant interaction characteristics associate with infant falling reactivity and child peer problems at pre-school age. *Plos one*, 19(6), e0302661.

Rigby, S. N., Stoesz, B. M., & Jakobson, L. S. (2018). Empathy and face processing in adults with and without autism spectrum disorder. *Autism Research*, 11(6), 942-955.

Rivolta, D., Lawson, R. P., & Palermo, R. (2017). More than just a problem with faces: Altered body perception in a group of congenital prosopagnosics. *Quarterly Journal of Experimental Psychology*, 70(2), 276-286.

Robbins, A. M., Koch, D. B., Osberger, M. J., Zimmerman-Phillips, S., & Kishon-Rabin, L. (2004). Effect of age at cochlear implantation on auditory skill development in infants and toddlers. *Archives of Otolaryngology–Head & Neck Surgery*, 130(5), 570-574.

Robertson, D. J., Black, J., Chamberlain, B., Megreya, A. M., & Davis, J. P. (2020). Super - recognisers show an advantage for other race face identification. *Applied Cognitive Psychology*, 34(1), 205-216.

Robertson, D. J., Noyes, E., Dowsett, A. J., Jenkins, R., & Burton, A. M. (2016). Face recognition by metropolitan police super-recognisers. *PloS one*, 11(2), e0150036.

Robson, P. (1984). Prewalking locomotor movements and their use in predicting standing and walking. *Child: care, health and development*, 10(5), 317-330.

Rodrigues, N., & Abhyankar, S. C. (2012). Emotional Intelligence and Anger in only Children and Children with Siblings. *Journal of Psychosocial Research*, 7(2).

Rodriguez, M. L., Ayduk, O., Aber, J. L., Mischel, W., Sethi, A., & Shoda, Y. (2005). A contextual approach to the development of self - regulatory competencies: The role of maternal unresponsivity and toddlers' negative affect in stressful situations. *Social Development*, 14(1), 136-157.

Rose, S. A., Jankowski, J. J., & Feldman, J. F. (2002). Speed of processing and face recognition at 7 and 12 months. *Infancy*, 3(4), 435-455.

Rossion, B. (2013). The composite face illusion: A whole window into our understanding of holistic face perception. *Visual Cognition*, 21(2), 139-253.

Rothbart, M. K. (1986). Longitudinal observation of infant temperament. *Developmental Psychology*, 22(3), 356.

Rothbart, M. K. (1988). Temperament and the development of inhibited approach. *Child development*, 1241-1250.

Rothbart, M. K. (1989). Temperament in childhood: A framework. *Temperament in childhood*, 5973.

Rothbart, M. K., & Bates, J. E. (1998). *Temperament*.

- Rothbart, M. K., & Bates, J. E. (2007). Temperament. *Handbook of child psychology*, 3.
- Rothbart, M. K., Ellis, L. K., & Posner, M. I. (2004). Temperament and self-regulation. *Handbook of self-regulation: Research, theory, and applications*, 2, 441-460.
- Ruffman, T., Lorimer, B., & Scarf, D. (2017). Do infants really experience emotional contagion?. *Child Development Perspectives*, 11(4), 270-274.
- Russell, J. A. (1989). Measures of emotion. In *The measurement of emotions* (pp. 83-111). Academic Press.
- Russell, J. A. (2015). Moving on from the basic emotion theory of facial expressions. *Emotion Researcher*.
- Russell, R., Duchaine, B., & Nakayama, K. (2009). Super-recognizers: People with extraordinary face recognition ability. *Psychonomic bulletin & review*, 16(2), 252-257.
- Russell, R., Porcheron, A., Sweda, J. R., Jones, A. L., Mauger, E., & Morizot, F. (2016). Facial contrast is a cue for perceiving health from the face. *Journal of Experimental Psychology: Human Perception and Performance*, 42(9), 1354.
- Saarni, C. (1979). Children's understanding of display rules for expressive behavior. *Developmental psychology*, 15(4), 424.
- Saarni, C. (1984). An observational study of children's attempts to monitor their expressive behavior. *Child development*, 1504-1513.
- Saayman, G., Ames, E. W., & Moffett, A. (1964). Response to novelty as an indicator of visual discrimination in the human infant. *Journal of Experimental Child Psychology*, 1(2), 189-198.

Safar, K., Kusec, A., & Moulson, M. C. (2017). Face experience and the attentional bias for fearful expressions in 6-and 9-month-old infants. *Frontiers in psychology*, 8, 1575.

Safdar, S., Friedlmeier, W., Matsumoto, D., Yoo, S. H., Kwantes, C. T., Kakai, H., & Shigemasa, E. (2009). Variations of emotional display rules within and across cultures: A comparison between Canada, USA, and Japan. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 41(1), 1.

Sanefuji, W., Wada, K., Yamamoto, T., Mohri, I., & Taniike, M. (2014). Development of preference for conspecific faces in human infants. *Developmental Psychology*, 50(4), 979.

Sang, S. A., & Nelson, J. A. (2017). The effect of siblings on children's social skills and perspective taking. *Infant and Child Development*, 26(6), e2023.

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.

Sangrigoli, S., Pallier, C., Argenti, A. M., Ventureyra, V. A., & de Schonen, S. (2005). Reversibility of the other-race effect in face recognition during childhood. *Psychological Science*, 16(6), 440-444.

Sanson, A., Hemphill, S. A., & Smart, D. (2004). Connections between temperament and social development: A review. *Social development*, 13(1), 142-170.

Santesso, D. L., Schmidt, L. A., & Trainor, L. J. (2007). Frontal brain electrical activity (EEG) and heart rate in response to affective infant-directed (ID) speech in 9-month-old infants. *Brain and cognition*, 65(1), 14-21.

- Sasson, N. J. (2006). The development of face processing in autism. *Journal of autism and developmental disorders*, 36, 381-394.
- Sasson, N. J. (2006). The development of face processing in autism. *Journal of autism and developmental disorders*, 36(3), 381-394.
- Sasson, N. J., Lam, K. S., Parlier, M., Daniels, J. L., & Piven, J. (2013). Autism and the broad autism phenotype: familial patterns and intergenerational transmission. *Journal of Neurodevelopmental Disorders*, 5(1), 11.
- Sato, W., & Yoshikawa, S. (2007). Spontaneous facial mimicry in response to dynamic facial expressions. *Cognition*, 104(1), 1-18.
- Sato, W., Kochiyama, T., Uono, S., Yoshikawa, S., & Toichi, M. (2017). Direction of amygdala–neocortex interaction during dynamic facial expression processing. *Cerebral Cortex*, 27(3), 1878-1890.
- Saudino, K. J. (2005). Behavioral genetics and child temperament. *Journal of Developmental & Behavioral Pediatrics*, 26(3), 214-223.
- Saxe, R. (2006). Uniquely human social cognition. *Current opinion in neurobiology*, 16(2), 235-239.
- Saxe, R., & Houlihan, S. D. (2017). Formalizing emotion concepts within a Bayesian model of theory of mind. *Current opinion in Psychology*, 17, 15-21.
- Schaffer, M., Clark, S., & Jeglic, E. L. (2009). The role of empathy and parenting style in the development of antisocial behaviors. *Crime & Delinquency*, 55(4), 586-599.

Scheerer, N. E., Behich, J., Liu, H., & Jones, J. A. (2013). ERP correlates of the magnitude of pitch errors detected in the human voice. *Neuroscience*, 240, 176-185.

Scherer, K. R. (1984). On the nature and function of emotion: A component process approach. *Approaches to emotion*, 2293(317), 31.

Scherer, K. R., Johnstone, T., & Klasmeyer, G. (2003). Vocal expression of emotion. *Handbook of affective sciences*, 433-456.

Scherer, K. R., & Scherer, U. (2011). Assessing the ability to recognize facial and vocal expressions of emotion: Construction and validation of the Emotion Recognition Index. *Journal of Nonverbal Behavior*, 35, 305-326.

Schlegel, K., Grandjean, D., & Scherer, K. R. (2012). Emotion recognition: Unidimensional ability or a set of modality-and emotion-specific skills?. *Personality and Individual Differences*, 53(1), 16-21.

Schmitow, C., Stenberg, G., Billard, A., & Hofsten, C. V. (2013). Using a head-mounted camera to infer attention direction. *International Journal of Behavioral Development*, 37(5), 468-474.

Schreckenbach, T., Ochsendorf, F., Sterz, J., Rüsseler, M., Bechstein, W. O., Bender, B., & Bechtoldt, M. N. (2018). Emotion recognition and extraversion of medical students interact to predict their empathic communication perceived by simulated patients. *BMC medical education*, 18, 1-10.

Schwabe, L., Höffken, O., Tegenthoff, M., & Wolf, O. T. (2013). Opposite effects of noradrenergic arousal on amygdala processing of fearful faces in men and women. *Neuroimage*, 73, 1-7.

Schwarzer, G., Zauner, N., & Jovanovic, B. (2007). Evidence of a shift from featural to configural face processing in infancy. *Developmental Science*, 10(4), 452-463.

Scott, L. S., & Fava, E. (2013). The own-species face bias: A review of developmental and comparative data. *Visual Cognition*, 21(9-10), 1364-1391.

Scott, R. M. (2017). Surprise! 20-month-old infants understand the emotional consequences of false beliefs. *Cognition*, 159, 33-47.

Seavey, C. A., Katz, P. A., & Zalk, S. R. (1975). Baby x. Sex roles, 1(2), 103-109.

Segal, S. C., Reyes, B. N., Gobin, K. C., & Moulson, M. C. (2019). Children's recognition of emotion expressed by own-race versus other-race faces. *Journal of experimental child psychology*, 182, 10

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: The Journal of the International Society for Developmental Psychobiology*, 25(6), 411-425.

Shamay-Tsoory, S. G., Aharon-Peretz, J., & Perry, D. (2009). Two systems for empathy: a double dissociation between emotional and cognitive empathy in inferior frontal gyrus versus ventromedial prefrontal lesions. *Brain*, 132(3), 617-627.

Shapiro, P. N., & Penrod, S. (1986). Meta-analysis of facial identification studies. *Psychological bulletin*, 100(2), 139.

Shariff, A. F., & Tracy, J. L. (2009). Knowing who's boss: Implicit perceptions of status from the nonverbal expression of pride. *Emotion*, 9(5), 631.

Shariff, A. F., & Tracy, J. L. (2011). What are emotion expressions for?. *Current Directions in Psychological Science*, 20(6), 395-399.

Shen, G., Meltzoff, A. N., Weiss, S. M., & Marshall, P. J. (2020). Body representation in infants: Categorical boundaries of body parts as assessed by somatosensory mismatch negativity. *Developmental Cognitive Neuroscience*, 44, 100795.

Shih, R. A., Belmonte, P. L., & Zandi, P. P. (2004). A review of the evidence from family, twin and adoption studies for a genetic contribution to adult psychiatric disorders. *International review of psychiatry*, 16(4), 260-283.

Shimo, S. (2020, November). Risks of Bias in AI-Based Emotional Analysis Technology from Diversity Perspectives. In 2020 IEEE International Symposium on Technology and Society (ISTAS) (pp. 66-68). IEEE.

Short, L. A., Hatry, A. J., & Mondloch, C. J. (2011). The development of norm-based coding and race-specific face prototypes: An examination of 5-and 8-year-olds' face space. *Journal of Experimental Child Psychology*, 108(2), 338-357.

Shuman, V., & Scherer, K. R. (2014). Concepts and structures of emotions. *International handbook of emotions in education*, 13-35.

Simion, F., & Giorgio, E. D. (2015). Face perception and processing in early infancy: inborn predispositions and developmental changes. *Frontiers in psychology*, 6, 969.

Simion, F., & Giorgio, E. D. (2015). Face perception and processing in early infancy: inborn pretemperaments and developmental changes. *Frontiers in psychology*, 6, 969.

Simion, F., Valenza, E., Macchi Cassia, V., Turati, C., & Umiltà, C. (2002). Newborns' preference for up–down asymmetrical configurations. *Developmental Science*, 5(4), 427-434.

Simion, F., Valenza, E., Umiltà, C., & Barba, B. D. (1998). Preferential orienting to faces in newborns: A temporal–nasal asymmetry. *Journal of Experimental Psychology: Human perception and performance*, 24(5), 1399.

Simner, M. L. (1971). Newborn's response to the cry of another infant. *Developmental psychology*, 5(1), 136.

Simons, G., Ellgring, H., & Smith Pasqualini, M. (2003). Disturbance of spontaneous and posed facial expressions in Parkinson's disease. *Cognition & Emotion*, 17(5), 759-778.

Sininger, Y. S., & Abdala, C. (1996). Hearing threshold as measured by auditory brain stem response in human neonates. *Ear and hearing*, 17(5), 395-401.

Sirigu, A., Grafman, J., Bressler, K., & Sunderland, T. (1991). Multiple representations contribute to body knowledge processing: Evidence from a case of autotopagnosia. *Brain*, 114(1), 629-642.

- Skuse, D. H., Lori, A., Cubells, J. F., Lee, I., Conneely, K. N., Puura, K., ... & Young, L. J. (2014). Common polymorphism in the oxytocin receptor gene (OXTR) is associated with human social recognition skills. *Proceedings of the National Academy of Sciences*, 111(5), 1987-1992.
- Slater, A. (2002). Visual perception in the newborn infant: issues and debates. *Intellectica*, 34, 57-76.
- Slater, A., Earle, D. C., Morison, V., & Rose, D. (1985). Pattern preferences at birth and their interaction with habituation-induced novelty preferences. *Journal of Experimental Child Psychology*, 39(1), 37-54.
- Slaughter, V., Heron, M., & Sim, S. (2002). Development of preferences for the human body shape in infancy. *Cognition*, 85(3), B71-B81.
- Slaughter, V., Heron, M., Jenkins, L., Tilse, E., Müller, U., & Liebermann, D. (2004). Origins and early development of human body knowledge. *Monographs of the Society for Research in Child Development*, i-113.
- Slaughter, V., Stone, V. E., & Reed, C. (2004). Perception of faces and bodies: similar or different?. *Current directions in psychological science*, 13(6), 219-223.
- Sliwinska, M. W., & Pitcher, D. (2018). TMS demonstrates that both right and left superior temporal sulci are important for facial expression recognition. *Neuroimage*, 183, 394-400.
- Sloutsky, V. M. (1997). Institutional care and developmental outcomes of 6-and 7-year-old children: A contextualist perspective. *International Journal of Behavioral Development*, 20(1), 131-151.

- Smetana, J. G. (1984). Toddlers' social interactions regarding moral and conventional transgressions. *Child Development*, 1767-1776.
- Smetana, J. G. (1989). Toddlers' social interactions in the context of moral and conventional transgressions in the home. *Developmental Psychology*, 25(4), 499.
- Smith, A. (2006). Cognitive empathy and emotional empathy in human behavior and evolution. *The Psychological Record*, 56(1), 3-21.
- Smith, L. B., Jayaraman, S., Clerkin, E., & Yu, C. (2018). The developing infant creates a curriculum for statistical learning. *Trends in cognitive sciences*, 22(4), 325-336.
- Smith, N. A., & Trainor, L. J. (2008). Infant - directed speech is modulated by infant feedback. *Infancy*, 13(4), 410-420.
- Smorti, M., & Ponti, L. (2018). How does sibling relationship affect children's prosocial behaviors and best friend relationship quality?. *Journal of Family Issues*, 39(8), 2413-2436.
- Soderstrom, M., Reimchen, M., Sauter, D., & Morgan, J. L. (2017). Do infants discriminate non-linguistic vocal expressions of positive emotions?. *Cognition and Emotion*, 31(2), 298-311.
- Sorce, J. F., Emde, R. N., Campos, J. J., & Klinnert, M. D. (1985). Maternal emotional signaling: its effect on the visual cliff behavior of 1-year-olds. *Developmental psychology*, 21(1), 195.
- Spittle, A. J., Treyvaud, K., Doyle, L. W., Roberts, G., Lee, K. J., Inder, T. E., ... & Anderson, P. J. (2009). Early emergence of behavior and social-emotional problems in very preterm infants. *Journal of the American Academy of Child & Adolescent Psychiatry*, 48(9), 909-918.

Stapel, J. C., van Wijk, I., Bekkering, H., & Hunnius, S. (2017). Eighteen - month - old infants show distinct electrophysiological responses to their own faces. *Developmental Science*, 20(5), e12437.

Stefanics, G., Háden, G. P., Sziller, I., Balázs, L., Beke, A., & Winkler, I. (2009). Newborn infants process pitch intervals. *Clinical Neurophysiology*, 120(2), 304-308.

Stephan, B. C. M., Breen, N., & Caine, D. (2006). The recognition of emotional expression in prosopagnosia: Decoding whole and part faces. *Journal of the International Neuropsychological Society*, 12(6), 884-895.

Stevenage, S. V., Clarke, G., & McNeill, A. (2012). The “other-accent” effect in voice recognition. *Journal of Cognitive Psychology*, 24(6), 647-653.

Strand, P. S., Cerna, S., & Downs, A. (2008). Shyness and emotion - processing skills in preschoolers: A 6 - month longitudinal study. *Infant and Child Development: An International Journal of Research and Practice*, 17(2), 109-120.

Striano, T., Brennan, P. A., & Vanman, E. J. (2002). Maternal depressive symptoms and 6 - month - old infants' sensitivity to facial expressions. *Infancy*, 3(1), 115-126.

Striano, T., Kopp, F., Grossmann, T., & Reid, V. M. (2006). Eye contact influences neural processing of emotional expressions in 4-month-old infants. *Social cognitive and affective neuroscience*, 1(2), 87-94.

- Sugden, N. A., & Moulson, M. C. (2019). These are the people in your neighbourhood: Consistency and persistence in infants' exposure to caregivers', relatives', and strangers' faces across contexts. *Vision Research*, 157, 230-241.
- Sugden, N. A., Mohamed - Ali, M. I., & Moulson, M. C. (2014). I spy with my little eye: Typical, daily exposure to faces documented from a first - person infant perspective. *Developmental psychobiology*, 56(2), 249-261.
- Sugita, Y. (2008). Face perception in monkeys reared with no exposure to faces. *Proceedings of the National Academy of Sciences*, 105(1), 394-398.
- Sullivan, P. F., Neale, M. C., & Kendler, K. S. (2000). Genetic epidemiology of major depression: review and meta-analysis. *American journal of psychiatry*, 157(10), 1552-1562.
- Sundaram, M., & Mani, A. (2016). Face recognition: demystification of multifarious aspect in evaluation metrics. *Face Recognition-Semisupervised Classification, Subspace Projection and Evaluation Methods*, 75-92.
- Surakka, V., & Hietanen, J. K. (1998). Facial and emotional reactions to Duchenne and non-Duchenne smiles. *International journal of psychophysiology*, 29(1), 23-33.
- Susilo, T., & Duchaine, B. (2013). Advances in developmental prosopagnosia research. *Current opinion in neurobiology*, 23(3), 423-429.
- Suurland, J., van der Heijden, K. B., Huijbregts, S. C., Smaling, H. J., de Sonnevile, L. M., Van Goozen, S. H., & Swaab, H. (2016). Parental perceptions of aggressive behavior in preschoolers:

Inhibitory control moderates the association with negative emotionality. *Child development*, 87(1), 256-269.

Suzuki, A., Hoshino, T., & Shigemasu, K. (2010). Happiness is unique: A latent structure of emotion recognition traits revealed by statistical model comparison. *Personality and Individual Differences*, 48(2), 196-201.

Székely, E., Lucassen, N., Tiemeier, H., Bakermans-Kranenburg, M. J., Van Ijzendoorn, M. H., Kok, R., ... & Herba, C. M. (2014). Maternal depressive symptoms and sensitivity are related to young children's facial expression recognition: The Generation R Study. *Development and Psychopathology*, 26(2), 333-345.

Tamis-LeMonda, C., & Lockman, J. J. (2023). *Developmental cascades*. Elsevier.

Tangen, S., Olsen, A., & Sandseter, E. B. H. (2022). A GoPro look on how children aged 17–25 months assess and manage risk during free exploration in a varied natural environment. *Education Sciences*, 12(5), 361.

Tarabulsy, G. M., Provost, M. A., Bordeleau, S., Trudel-Fitzgerald, C., Moran, G., Pederson, D. R., ... & Pierce, T. (2009). Validation of a short version of the maternal behavior Q-set applied to a brief video record of mother–infant interaction. *Infant Behavior and Development*, 32(1), 132-136.

Tarabulsy, G. M., Provost, M. A., Deslandes, J., St-Laurent, D., Moss, E., Lemelin, J. P., ... & Dassylva, J. F. (2003). Individual differences in infant still-face response at 6 months. *Infant behavior and development*, 26(3), 421-438.

Tardif, J., Morin Duchesne, X., Cohan, S., Royer, J., Blais, C., Fiset, D., ... & Gosselin, F.

(2019). Use of face information varies systematically from developmental prosopagnosics to super-recognizers. *Psychological science*, 30(2), 300-308.

Taylor - Colls, S., & Pasco Fearon, R. M. (2015). The effects of parental behavior on infants' neural processing of emotion expressions. *Child development*, 86(3), 877-888.

Taylor, M. J., Batty, M., & Itier, R. J. (2004). The faces of development: a review of early face processing over childhood. *Journal of cognitive neuroscience*, 16(8), 1426-1442.

Terrizzi, B. F., & Beier, J. S. (2016). Automatic cueing of covert spatial attention by a novel agent in preschoolers and adults. *Cognitive Development*, 40, 111-119.

Thelen, E., & Spencer, J. P. (1998). Postural control during reaching in young infants: a dynamic systems approach. *Neuroscience & Biobehavioral Reviews*, 22(4), 507-514.

Thomas, L. A., De Bellis, M. D., Graham, R., & LaBar, K. S. (2007). Development of emotional facial recognition in late childhood and adolescence. *Developmental science*, 10(5), 547-558.

Thompson, L. A., & Trevathan, W. R. (2009). Cortisol reactivity, maternal sensitivity, and infant preference for mother's familiar face and rhyme in 6 - month - old infants. *Journal of reproductive and infant psychology*, 27(2), 143-167.

Tiedens, L. Z. (2001). Anger and advancement versus sadness and subjugation: the effect of negative emotion expressions on social status conferral. *Journal of personality and social psychology*, 80(1), 86.

- Toda, S., & Fogel, A. (1993). Infant response to the still-face situation at 3 and 6 months. *Developmental psychology*, 29(3), 532.
- Travis, L. L., & Sigman, M. (1998). Social deficits and interpersonal relationships in autism. *Mental retardation and developmental disabilities Research Reviews*, 4(2), 65-72.
- Trevathan, W. R., & Rosenberg, K. R. (Eds.). (2016). *Costly and cute: Helpless infants and human evolution*. University of New Mexico Press.
- Trevisan, D. A., & Birmingham, E. (2016). Are emotion recognition abilities related to everyday social functioning in ASD? A meta-analysis. *Research in autism spectrum disorders*, 32, 24-42.
- Turati, C., Di Giorgio, E., Bardi, L., & Simion, F. (2010). Holistic face processing in newborns, 3 - month - old infants, and adults: Evidence from the composite face effect. *Child development*, 81(6), 1894-1905.
- Turati, C., Montiroso, R., Brenna, V., Ferrara, V., & Borgatti, R. (2011). A smile enhances 3 - month - olds' recognition of an individual face. *Infancy*, 16(3), 306-317.
- Turati, C., Sangrigoli, S., Rueli, J., & de Schonen, S. (2004). Evidence of the face inversion effect in 4 - month - old infants. *Infancy*, 6(2), 275-297.
- Uher, R. (2014). Gene–environment interactions in severe mental illness. *Frontiers in psychiatry*, 5, 48.
- Uhlhaas, P. J., Roux, F., Rodriguez, E., Rotarska-Jagiela, A., & Singer, W. (2010). Neural synchrony and the development of cortical networks. *Trends in cognitive sciences*, 14(2), 72-80.

Ursache, A., Blair, C., Stifter, C., & Voegtline, K. (2013). Emotional reactivity and regulation in infancy interact to predict executive functioning in early childhood. *Developmental psychology*, 49(1), 127.

Vaish, A., & Striano, T. (2004). Is visual reference necessary? Contributions of facial versus vocal cues in 12 - month - olds' social referencing behavior. *Developmental Science*, 7(3), 261-269.

Valentine, T., & Bruce, V. (1986). The effects of distinctiveness in recognising and classifying faces. *Perception*, 15(5), 525-535.

Valenza, E., & Bulf, H. (2007). The role of kinetic information in newborns' perception of illusory contours. *Developmental science*, 10(4), 492-501.

van Boxtel, J. J., Peng, Y., Su, J., & Lu, H. (2017). Individual differences in high-level biological motion tasks correlate with autistic traits. *Vision research*, 141, 136-144.

van den Heuvel, M. I., Donkers, F. C., Winkler, I., Otte, R. A., & Van den Bergh, B. R. (2015). Maternal mindfulness and anxiety during pregnancy affect infants' neural responses to sounds. *Social cognitive and affective neuroscience*, 10(3), 453-460.

Van Der Zant, T., & Nelson, N. L. (2021). Motion increases recognition of naturalistic postures but not facial expressions. *Journal of nonverbal behavior*, 45(4), 587-600.

Van IJzendoorn, M. H., Palacios, J., Sonuga - Barke, E. J., Gunnar, M. R., Vorria, P., McCall, R. B., ... & Juffer, F. (2011). I. Children in institutional care: Delayed development and resilience. *Monographs of the Society for Research in Child Development*, 76(4), 8-30.

Vetter, N. C., Altgassen, M., Phillips, L., Mahy, C. E., & Kliegel, M. (2013). Development of affective theory of mind across adolescence: Disentangling the role of executive functions.

Developmental Neuropsychology, 38(2), 114-125.

Vogel, M., Monesson, A., & Scott, L. S. (2012). Building biases in infancy: The influence of race on face and voice emotion matching. *Developmental science*, 15(3), 359-372.

Vogindroukas, I., Chelas, E. N., & Petridis, N. E. (2014). Reading the Mind in the Eyes Test (children's version): a comparison study between children with typical development, children with high-functioning autism and typically developed adults. *Folia Phoniatica et Logopaedica*, 66(1-2), 18-24.

Volling, B. L., & Feagans, L. V. (1995). Infant day care and children's social competence. *Infant Behavior and Development*, 18(2), 177-188.

Vouloumanos, A., Hauser, M. D., Werker, J. F., & Martin, A. (2010). The tuning of human neonates' preference for speech. *Child development*, 81(2), 517-527.

Vuong, Q. C., & Geangu, E. (2023). The development of emotion processing of body expressions from infancy to early childhood: A meta-analysis. *Frontiers in Cognition*, 2, 1155031.

Wagers, K. B., & Kiel, E. J. (2019). The influence of parenting and temperament on empathy development in toddlers. *Journal of Family Psychology*, 33(4), 391.

Walker-Andrews, A. S., & Grolnick, W. (1983). Discrimination of vocal expressions by young infants. *Infant Behavior & Development*.

Walker, G. (2011). Postcommunist deinstitutionalization of children with disabilities in Romania: Human rights, adoption, and the ecology of disabilities in Romania. *Journal of Disability Policy Studies*, 22(3), 150-159.

Walker, P. M., & Hewstone, M. (2006). A perceptual discrimination investigation of the own - race effect and intergroup experience. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, 20(4), 461-475.

Wallace, S., Coleman, M., & Bailey, A. (2008). An investigation of basic facial expression recognition in autism spectrum disorders. *Cognition and Emotion*, 22(7), 1353-1380.

Walton, G. E., & Bower, T. G. R. (1993). Newborns form “prototypes” in less than 1 minute. *Psychological Science*, 4(3), 203-205.

Wang, J., Yang, Y., Tang, Y., Wu, M., Jiang, S., & Zou, H. (2021). Longitudinal links among parent-child attachment, emotion parenting, and problem behaviors of preadolescents. *Children and Youth Services Review*, 121, 105797.

Wang, M., & Deng, W. (2021). Deep face recognition: A survey. *Neurocomputing*, 429, 215-244.

Ward, K. P., & Lee, S. J. (2020). Mothers’ and fathers’ parenting stress, responsiveness, and child wellbeing among low-income families. *Children and Youth Services Review*, 116, 105218.

Wei, W., Wu, Y., Lv, B., Zhou, H., Han, X., Liu, Z., & Luo, L. (2016). The relationship between parental involvement and elementary students’ academic achievement in China: One-only children vs. children with siblings. *Journal of Comparative Family Studies*, 47(4), 483-500.

- Weisleder, A., & Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological science*, 24(11), 2143-2152.
- Weldon, K. B., Taubert, J., Smith, C. L., & Parr, L. A. (2013). How the Thatcher illusion reveals evolutionary differences in the face processing of primates. *Animal cognition*, 16(5), 691-700.
- Wenzler, S., Levine, S., van Dick, R., Oertel-Knöchel, V., & Aviezer, H. (2016). Beyond pleasure and pain: Facial expression ambiguity in adults and children during intense situations. *Emotion*, 16(6), 807.
- Werkmann, N. L., Luczejko, A. A., Hagelweide, K., Stark, R., Weigelt, S., Christiansen, H., ... & COMPARE-family research group. (2024). Facial emotion recognition in children of parents with a mental illness. *Frontiers in Psychiatry*, 15, 1366005.
- White, S. F., Briggs-Gowan, M. J., Voss, J. L., Petitcherc, A., McCarthy, K., R. Blair, R. J., & Wakschlag, L. S. (2016). Can the fear recognition deficits associated with callous-unemotional traits be identified in early childhood?. *Journal of clinical and experimental neuropsychology*, 38(6), 672-684.
- Whittaker, J. E. V., Harden, B. J., See, H. M., Meisch, A. D., & T'Pring, R. W. (2011). Family risks and protective factors: Pathways to Early Head Start toddlers' social-emotional functioning. *Early childhood research quarterly*, 26(1), 74-86.
- Wickline, V. B., Bailey, W., & Nowicki, S. (2009). Cultural in-group advantage: Emotion recognition in African American and European American faces and voices. *The Journal of genetic psychology*, 170(1), 5-30.

Willems, S., Vrancken, L., Germeys, F., & Verfaillie, K. (2014). Holistic processing of human body postures: evidence from the composite effect. *Frontiers in Psychology*, 5, 618.

Wilmer, J. B., Germine, L., Chabris, C. F., Chatterjee, G., Williams, M., Loken, E., ... & Duchaine, B. (2010). Human face recognition ability is specific and highly heritable. *Proceedings of the National Academy of sciences*, 107(11), 5238-5241.

Wingenbach, T. S., Ashwin, C., & Brosnan, M. (2017). Diminished sensitivity and specificity at recognising facial emotional expressions of varying intensity underlie emotion-specific recognition deficits in autism spectrum disorders. *Research in Autism Spectrum Disorders*, 34, 52-61.

Winnicott, D. W. (2018). *The maturational processes and the facilitating environment: Studies in the theory of emotional development*. Routledge.

Wong, N., Beidel, D. C., Sarver, D. E., & Sims, V. (2012). Facial emotion recognition in children with high functioning autism and children with social phobia. *Child Psychiatry & Human Development*, 43, 775-794.

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, 69(1), 1-34.

Worobey, J. (1986). Convergence among assessments of temperament in the first month. *Child Development*, 47-55.

Worobey, J., & Blajda, V. M. (1989). Temperament ratings at 2 weeks, 2 months, and 1 year: Differential stability of activity and emotionality. *Developmental Psychology*, 25(2), 257.

Wu, M., Kujawa, A., Lu, L. H., Fitzgerald, D. A., Klumpp, H., Fitzgerald, K. D., ... & Phan, K. L. (2016). Age - related changes in amygdala - frontal connectivity during emotional face processing from childhood into young adulthood. *Human brain mapping*, 37(5), 1684-1695.

Wu, Y., Muentener, P., & Schulz, L. E. (2017). One-to four-year-olds connect diverse positive emotional vocalizations to their probable causes. *Proceedings of the National Academy of Sciences*, 114(45), 11896-11901.

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.

Xie, W., Leppänen, J. M., Kane-Grade, F. E., & Nelson, C. A. (2021). Converging neural and behavioral evidence for a rapid, generalized response to threat-related facial expressions in 3-year-old children. *NeuroImage*, 229, 117732.

Xu, T., Chen, Y., & Smith, L. (2011, August). It's the child's body: The role of toddler and parent in selecting toddler's visual experience. In *2011 IEEE International Conference on Development and Learning (ICDL)* (Vol. 2, pp. 1-6). IEEE.

Yan, X., Andrews, T. J., Jenkins, R., & Young, A. W. (2016). Cross-cultural differences and similarities underlying other-race effects for facial identity and expression.

Yi, C. Y., Murry, M. W., & Gentzler, A. L. (2016). Perception of emotional expressions in adults. *Journal of Individual Differences*.

Young, A. W., & Burton, A. M. (2018). Are we face experts?. *Trends in cognitive sciences*, 22(2), 100-110.

Young, A. W., Hellowell, D. J., Van de Wal, C., & Johnson, M. (1996). Facial expression processing after amygdalotomy. *Neuropsychologia*, 34(1), 31-39.

Young, S. K., Fox, N. A., & Zahn-Waxler, C. (1999). The relations between temperament and empathy in 2-year-olds. *Developmental psychology*, 35(5), 1189.

Yovel, G., Wilmer, J. B., & Duchaine, B. (2014). What can individual differences reveal about face processing?. *Frontiers in human neuroscience*, 8, 562.

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.

Zeng, J., Shan, S., & Chen, X. (2018). Facial expression recognition with inconsistently annotated datasets. In *Proceedings of the European conference on computer vision (ECCV)* (pp. 222-237).

Zghoul, T., Cowen, P. J., & Harmer, C. J. (2019). A perspective: from the serotonin hypothesis to cognitive neuropsychological approaches. In *The Serotonin System* (pp. 95-104). Academic Press.

Zhang, D., Wang, L., Luo, Y., & Luo, Y. (2012). Individual differences in detecting rapidly presented fearful faces. *PloS one*, 7(11), e49517.

- Zhang, Z., Yu, H., Long, M., & Li, H. (2021). Worse Theory of Mind in Only-Children Compared to Children With Siblings and Its Intervention. *Frontiers in psychology*, 12, 5073.
- Zhao, C., Chronaki, G., Schiessl, I., Wan, M. W., & Abel, K. M. (2019). Is infant neural sensitivity to vocal emotion associated with mother-infant relational experience?. *PloS one*, 14(2), e0212205.
- Zhao, C., Schiessl, I., Wan, M. W., Chronaki, G., & Abel, K. M. (2020). Development of the neural processing of vocal emotion during the first year of life. *Child Neuropsychology*, 1-18.
- Zhao, G., Zhang, Y., & Ge, Y. (2018). Frontal EEG asymmetry and middle line power difference in discrete emotions. *Frontiers in behavioral neuroscience*, 12, 225.
- Zhou, X., Chen, C. M., Mondloch, C. J., Chien, S. H. L., & Moulson, M. (2019). Perceptual Experience and Within-Person Variability Affect the Magnitude of the Other-Race Effect. *Journal of Vision*, 19(10), 153c-153c.
- Zhou, X., Elshiekh, A., & Moulson, M. C. (2019). Lifetime perceptual experience shapes face memory for own-and other-race faces. *Visual Cognition*, 27(9-10), 687-700.
- Zhou, X., Siddiqui, H., & Rutherford, M. D. (2024). Face perception and social cognitive development in early autism: A prospective longitudinal study from 3 months to 7 years of age. *Child Development*.
- Zieber, N., Kangas, A., Hock, A., & Bhatt, R. S. (2014). Infants' perception of emotion from body movements. *Child development*, 85(2), 675-684.

Zych, A. D., & Gogolla, N. (2021). Expressions of emotions across species. *Current opinion in neurobiology*, 68, 57-66.