Adults' Responses to Infant Vocalisations: A Neurobehavioural Investigation

Katherine S. Young

Jesus College

Trinity Term, 2013

Thesis submitted in partial fulfilment for the degree of
Doctor of Philosophy at the University of Oxford

Department of Psychiatry

University of Oxford

Adults' Responses to Infant Vocalisations: A Neurobehavioural Investigation

Thesis submitted in partial fulfilment for the degree of
Doctor of Philosophy at the University of Oxford

Katherine S. Young, Jesus College, Trinity Term 2013

Abstract

Infant vocalisations are uniquely salient sounds in the environment. They universally attract attention and compel the listener to respond with speed and care. They provide a wealth of information to parents about their infant's needs and affective state. There is a scientific consensus that early parenting has a profound impact on child development. In particular, the sensitivity with which parents respond to their infant's communicative cues has been shown to affect cognitive and socio-emotional outcomes. The mechanisms underlying such sensitivity are not well understood.

In this thesis, adults' sensitivity to infant cues will be considered in terms of two components, the 'promptness' and 'appropriateness' of responses, as originally conceptualised by Bell and Ainsworth (1972). Promptness of responses is considered in terms of adults' ability to move with speed and effort after listening to infant vocalisations. Appropriateness, on the other hand, is considered in terms of adults' ability to differentiate between functionally significant parameters in infant vocalisations. The effect of modifiable environmental factors on the promptness and appropriateness of responses is also investigated. Finally, a focused investigation of the brain basis of responses to infant vocalisations is presented.

Overall, findings demonstrated that infant vocalisations undergo privileged, specialised processing in the adult brain. After hearing an infant cry, adults with and without depression were found to move with greater coordination and effort. Adults were also found to be attuned to subtle parameters in infant cries. This sensitivity was shown to be affected by two participant-level factors, depression and previous musical training. Furthermore, this sensitivity could be enhanced through intervention, as evidenced by findings from short-term, perceptual discrimination training. The notion of privileged processing of infant vocalisations is further supported by evidence of early discrimination of infant sounds in a survival-related subcortical brain structure.

Future directions for this work include directly relating current experimental measures of adults' responses to infant cues with parental sensitivity to infant communication during dynamic interactions. Translating current findings into applied settings would require an investigation of the effects of factors such as musical and perceptual training on sensitivity to infant cues in at-risk populations, such as mothers and fathers with depression. Lastly, an increased understanding of the brain basis of adults' sensitivity to infant cues will provide insight into our greatest challenge: parenting our young.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisors, Professors Morten Kringelbach and Alan Stein for giving me the opportunity, the freedom and the support to pursue the work in this thesis. The combination of their enthusiasm and wide-ranging expertise has vastly broadened my research horizons over the past few years and for this I am particularly grateful.

I am also deeply indebted to Dr Christine Parsons for her supervision, support and friendship throughout this work. I am very grateful that she has shared with me her unrelenting passion for research, her meticulous attention to detail, her constant encouragement, all the good times (and the bad) and many a green soup. This thesis certainly would not have been the same without her.

I wish to thank the Medical Research Council and the Department of Psychiatry for granting me an MRC Supported DPhil Studentship to carry out this work. I would also like to thank Jesus College for their support in the form of a Graduate Scholarship.

I would like to thank the staff at Oxford Functional Neurosurgery for their assistance in testing DBS patients. Many thanks also to Dr Elvira Brattico and Morten Joensson for assistance with the development of control stimuli and preliminary analysis of LFP data. In addition, I would like to thank all the participants and patients involved in this research, without whom none of this would have been possible.

I am very grateful for the contributions of Geraldine Parsons, Eloise Stark and Tim van Hartevelt and their keen attention to detail during the proofreading stages. I would also like to thank all the members (past and present) of the Hedonia Research Group and the Section of Child and Adolescent Psychiatry who have helped to make work an interesting and fun place to be.

I can never thank enough my friends in Oxford and further afield who have supported me along the way, not least for their incredible patience when I wanted them to listen to ‘just a few more’ baby cries. Special thanks go to Simone, Isa, Jess, Ceri and Eugene.

Last but not least, I do not know how to express my gratitude to my parents, brother and sisters. Their unwavering belief in me has been a source of tremendous support as have their kind words and care packages. I could not have done this without them.

Contents

TABLES

FIGURES

List of Abbreviations

A1	Primary Auditory Cortex
AA	Adobe Audition
ACC	Anterior Cingulate Cortex
AFC	Alternative Forced Choice
ANOVA	Analysis of Variance
CT	Computerised Tomography
DBS	Deep Brain Stimulation
DSM-IV	Diagnostic and Statistical Manual of Mental Disorders (IV)
EEG	Electroencephalography
EPDS	Edinburgh Postnatal Depression Scale
ERP	Event-related Potential
F_0	Fundamental Frequency
FFG	Fusiform Gyrus
FFT	Fast Fourier Transform
FLIRT	FMRIB's Linear Image Registration Tool
fMRI	Functional Magnetic Resonance Imaging
FSL	FMRIB Software Library
GAD	Generalised Anxiety Disorder
GAD-Q-IV	Generalised Anxiety Disorder Questionnaire (IV)
GEMEP-CS	Geneva Multimodal Emotional Portrayal – Core Set
ICC	Intra-Class Correlation
ISI	Inter-Stimulus Interval
K-S	Kolmogorov-Smirnoff
KMO	Kaiser-Meyer-Olkin
LFP	Local Field Potential
MAV	Montreal Affective Voices
MEG	Magnetoencephalography
MNI	Montreal Neurological Institute
MRI	Magnetic Resonance Imaging
MTG	Middle Temporal Gyrus
OFC	Orbitofrontal Cortex
PAG	Periaqueductal Gray
PET	Positron Emission Tomography
PFC	Prefrontal Cortex
RMS	Root Mean Square
SCID	Structured Clinical Interview for DSM-IV
STG	Superior Temporal Gyrus
STS	Superior Temporal Sulcus
VAS	Visual Analogue Scale
VTA	Ventral Tegmental Area

Publications Associated with this work

Parsons, C. E.*, Young, K. S.*, Parsons, E., Stein, A., & Kringelbach, M. L. (2012). Listening to infant distress vocalizations enhances effortful motor performance. *Acta Paediatrica, International Journal of Paediatrics, 101*(4), e189.

Young, K. S., Parsons, C. E., Stein, A., & Kringelbach, M. L. (in prep). Overall motor performance is reduced in depression but differential emotional reactivity is not affected.

Young, K. S., Parsons, C. E., Stein, A., & Kringelbach, M. L. (2012). Interpreting infant vocal distress: the ameliorative effect of musical training in depression *Emotion, 12*(6), 1200.

Parsons, C. E.*, Young, K. S.*, Joensson, M., Brattico, E., Hyam, J. A., Stein, A., Green, A. L., Aziz, T. Z. Kringelbach, M. L. (2013). Ready for action: a role for the human midbrain in responding to infant vocalizations. *Social Cognitive and Affective Neuroscience*. doi: 10.1093/scan/nst076

Parsons, C. E., Stark, E. A., Young, K. S., Stein, A., & Kringelbach, M. L. (in press). Understanding the human parental brain: A critical role of the orbitofrontal cortex. *Social Neuroscience*.

*Authors contributed equally

CHAPTER 1: LITERATURE REVIEW

The vocalisation from an infant stands apart from all other environmental sounds. It is a sound that can uniquely capture an adult's attention, be it a distressed cry or a boisterous laugh. The most conspicuous of these vocalisations is the infant's cry. Unmistakable in origin, familiar and instantly recognisable, the sound of a crying infant confers an urgency to react and provide soothing, nurturing care.

The developing infant relies on vocalisations and facial expressions to communicate with a caregiver. These cues provide important information about the infant's physiological and affective state. Adults typically respond to these cues by approaching the infant and commencing caregiving behaviour (Ainsworth, 1969; Bowlby, 1969). Sensitive caregiving behaviour is characterised by 'responsiveness': prompt and appropriate reactions to infant communicative cues (Ainsworth, Bell, & Stayton, 1974; Bell & Ainsworth, 1972). Responsiveness to infant cues is a fundamental parental capacity present across diverse human cultures and observable in other mammalian species (Konner, 2010; Numan & Insel, 2003).

The parent-infant relationship evolves over time and is vital for infant survival and development (Darwin, 1872; Lorenz, 1943). Parenting is widely recognised to impact on child development and disruptions to responsive caregiving, as observed in postnatal depression for instance, impact on later cognitive and emotional capacities in children (Bakermans-Kranenburg, Van Ijzendoorn, & Juffer, 2003; Murray, Hipwell, Hooper, Stein, & Cooper, 1996). Despite the importance of parental responsiveness for child development, how parents come to understand and respond to their infant's communicative signals is not well understood.

Responses to infant cues can be considered within the broader context of emotional processing. Current theories, informed by neuroscientific research, suggest that there are two levels of processing involved in responding to emotional cues (Barrett & Bar, 2009; LeDoux, 2000; Rolls, 2000). First, rapid analysis of basic stimulus properties promotes the initiation of survival-type behavioural responses (LeDoux, 2000). Second, slower, more detailed analysis of stimulus content provides accurate information for appraisal and decision-making processes (Rolls, 2000). Processing of infant cries, as salient emotional cues, seems likely to follow similar fast and slow routes.

This literature review will consider adults' responses to infant communicative cues as emotional signals, subject to privileged processing. First, behavioural evidence assessing how infant cues can elicit responses in adults will be discussed. Factors shown to affect the interpretation of infant cues and the selection of appropriate caregiving responses will then be reviewed. Potential disruptions to the interpretation of and responses to infant cues will then be described in the context of postnatal depression. The final section of this review will discuss current understanding of the neural mechanisms supporting processing of infant cues in adults. Brain networks involved in the general processing of infant cues will be considered, as well as regions involved in the specific processing of own infant cues. Finally, emerging evidence for the role of one subcortical brain region in promoting responsive behaviour will be discussed. This chapter ends with a description of the main aims and experiments included in this thesis.

1.1 Parental Sensitivity to Infant Communicative Cues

The relationship between parent and infant is inherently complex, characterised by its bidirectional and dynamic nature. As the postnatal period proceeds, interactions become increasingly sophisticated, beginning with simple orienting to infant cues and culminating in prolonged interactions such as play, involving narrative structure and mentalizing aptitude. A behavioural framework has been described in which this relationship can be considered (see Figure 1; Parsons, Young, Murray, Stein, & Kringelbach, 2010). This framework outlines six major components of the parent-infant relationship over the first 18 months: a) orienting system; b) recognition system; c) intuitive parenting; d) attachment relationships; e) intersubjectivity; and f) higher socio-emotional and cognitive functions. The main focus of the work in this thesis, parental responses to infant cues, draws mostly on the first three of these components.

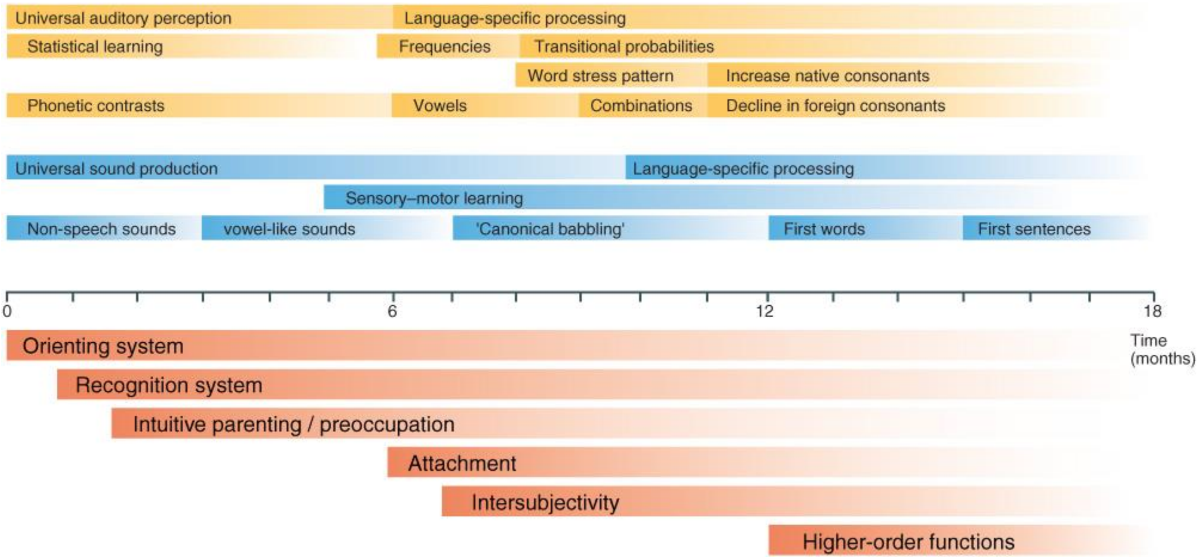

Figure 1. Behavioural framework mapping out key stages of the evolving parent-infant relationship over the first 18 months of life (red). This is compared with the infant's developing auditory (yellow) and verbal (blue) capacities over the same period (Kuhl, 2004).

Considered at the most basic level, parent-infant interactions are built up from parents and infants observing and responding to one another (Squire & Stein, 2003). For parents, the provision of sensitive care requires focussed attention and contingent responses to infant signals. This parental ‘responsiveness’ is evident in early parent-infant interactions. When initiating interactions, parents strive first to make direct eye contact with their infant (Papousek & Papousek, 1983). Once this is established, parents quickly respond by making exaggerated facial expressions and vocalisations in greeting (Papousek & Papousek, 1989). Interactions then tend to proceed with both infant and parent imitating each other’s communicative cues (Field, Woodson, Greenberg, & Cohen, 1982; Kuhl & Meltzoff, 1982; Meltzoff & Moore, 1977; Papousek & Papousek, 1975; Papousek & Papousek, 1989). While the very nature of these interactions is based on complex patterns of behaviour and a high degree of synchronicity, it is possible to scientifically investigate mechanisms supporting the production of these behaviours. In terms of parenting behaviours that are amenable to experimental investigation, a clear starting point concerns the mechanisms through which adults respond with such sensitivity to infant cues.

1.1.1 An Overview of Infant Vocal Cues

Infant communicative cues primarily consist of vocalisations and facial expressions (see Figure 2). From the earliest moments after birth, infants cry. Throughout infancy and into childhood, vocal expressions are a vital source of information for parents about the affective state and current needs of their child. In particular, vocalisations serve an alerting function, attracting the attention of nearby individuals and providing the first available cues as to the survival needs of the offspring. The nature of caregiving behaviour provided may then additionally be influenced by visual cues and contextual factors.

The sound of an infant cry is characterised by its high and variable pitch (ranging between 250-700Hz; Golub & Corwin, 1985). A single cry burst has a duration of one to three seconds and typically has a 'falling' or 'rising-falling' melody (although recent demonstrations suggest that cry melodies can vary across language with contrasting intonation patterns; Mampe, Friederici, Christophe, & Wermke, 2009; Wasz-Höckert, Michelsson, & Lind, 1985). In early life, infant crying is thought to be largely reflexive (Bell & Ainsworth, 1972) often occurring in response to pain, hunger or separation from a caregiver. Beyond the first two to three months after birth, there is a noticeable change in infant vocalisations. Cries become more differentiated, reflecting developmental changes in the vocal tract and a greater degree of control over vocalisation production (Ostwald & Murray, 1985; Soltis, 2004). Around four months of age infants begin to produce a much wider variety of vocal sounds, including positive emotional sounds, such as laughter, as well as pre-verbal vocalisations known as 'babbling' (Kuhl, 2004; Kuhl & Meltzoff, 1982; Oller & Eilers, 1988).

Early theoretical models focussed specifically on adults' responses to infant cries have considered these sounds as either aversive stimuli (Moss & Robson, 1968) or as elicitors of empathy (Hoffman, 1975). The aversive model suggests that infant cries are comparable to other unpleasant stimuli; motivating general behavioural responses to terminate the sound (see also Thompson, Bruzek, & Cotnoir-Bichelman, 2011). The empathy model suggests that hearing a distressed infant elicits sympathetic feelings of distress in the listener, with behavioural responses serving to soothe the infant's (and the listener's) distress. However, neither of these theories seems to be sufficient in accounting for the full range of adults' responses to infant cries.

A more comprehensive model considers infant cries to be ‘motivational entities’, promoting the likelihood of initiating behavioural responses in listeners (Murray, 1979, 1985). This theory originated from earlier hypotheses considering infant cries as an example of an ‘innate releaser’, eliciting stereotyped caregiving behaviours (Lorenz, 1950). The motivational entity model additionally acknowledges the influence of acoustic cues, contextual factors, motivational states and cognitive appraisal in guiding adults’ selection of behavioural responses. This two-stage model is comparable both to general theories of caregiving behaviour (featuring aspects of prompt and appropriate responsiveness; Bell & Ainsworth, 1972) and neuroscientific models of emotional processing (with quick imprecise processing followed by slower detailed processing; LeDoux, 2000; Rolls, 2000).

1.1.2 An Overview of Infant Facial Cues

Like infant vocalisations, infant faces also have distinctive physical features. Infant faces are characterised by the presence of large, rounded foreheads, large low-set eyes, a short and narrow nose, rounded cheeks and a small chin (see Figure 2; Alley, 1981; Glocker et al., 2009). These specific proportions of facial features are frequently referred to as “infant schema”, a term first coined by Konrad Lorenz in 1943. While vocalisations in early life are restricted to more negative expressions, infants are able to make both positive (smiling) and negative (frowning) facial expressions soon after birth (Messinger, 2002). As the infant develops and control over the facial muscles improve, a broader range of facial expressions is observed, including displays of surprise and anger around four months of age (Sullivan & Lewis, 2003).

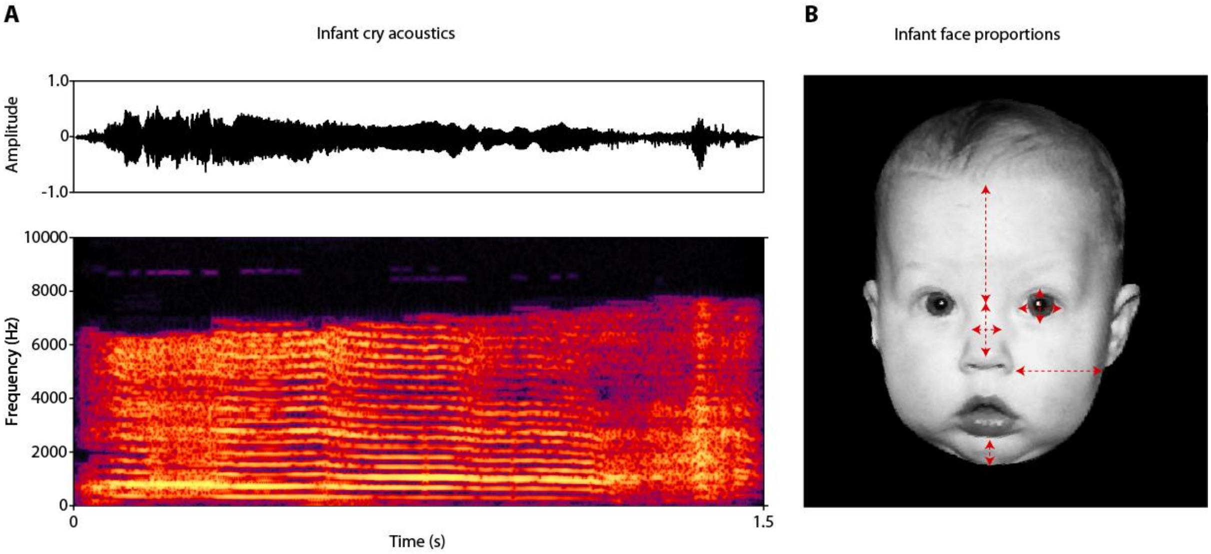

Figure 2. Examples of infant communicative cues. A) Waveform (upper) and spectrogram (lower) of a typical infant cry burst, characterised by high and variable pitch. B) Demonstration of the physical features of a typical infant face, reflecting the ‘infant schema’.

1.2 Prompt Responses to Infant Communicative Cues

1.2.1 Infant Vocal Cues

Both parents and non-parents alike demonstrate sensitivity to infant vocalisations and facial expressions (Kruger & Konner, 2010). Adults are naturally drawn to the sound of a ‘distressed’ infant cry or the image of a ‘cute’ infant face, stimulating the provision of sensitive caregiving behaviour (Ainsworth, 1969; Bowlby, 1969; Darwin, 1872; Lorenz, 1943). A number of experimental studies have investigated how adults react to infant cues presented in isolation.

Infant cries demand attention and have been shown to interfere with performance on simple cognitive tasks (Chang & Thompson, 2011; Morsbach, McCulloch, & Clark, 1986). They are perceived as sounding distressing, piercing, grating and aversive by both parents and non-parents alike (e.g., Dessureau, Kurowski, & Thompson, 1998; Frodi, Lamb, Leavitt, & Donovan, 1978; Gustafson & Green, 1989; Soltis, 2004). This subjective appraisal is associated with a desire to rapidly enact a range of caregiving behaviours (Del Vecchio, Walter, & O'Leary, 2009; Schuetze & Zeskind, 2001).

Both parents and non-parents have also been shown to be remarkably sensitive to subtle features of infant vocalisations. For instance, adults readily interpret variations in pitch as indicating different levels of distress (Dessureau et al., 1998; Zeskind & Marshall, 1988). This sensitivity to pitch in infant cries is thought to be important for determining certain aspects of parental behaviour. For example, the speed of caregiving responses can be affected by the pitch of infant cries (Zeskind & Collins, 1987).

1.2.2 Infant Facial Cues

While infant vocalisations are uniquely capable of attracting the attention of an otherwise inattentive adult, infant faces can also capture adults' attention, as demonstrated by experimental tasks such as the dot-probe (Brosch, Sander, & Scherer, 2007; Pearson, Cooper, Penton-Voak, Lightman, & Evans, 2010; Pearson, Lightman, & Evans, 2011). Infant faces are perceived as 'cute' by adults, who typically respond to pictures of infants by smiling (Hildebrandt & Fitzgerald, 1978) and prefer viewing pictures of infants over pictures of adults (Fullard & Reiling, 1976). Similar to the observed sensitivity to pitch in infant vocalisations, adults are highly sensitive to variations in infant facial cuteness (Glocker et al., 2009; Lobmaier, Sprengelmeyer, Wiffen, & Perrett, 2010; Parsons, Young, Kumari, Stein, & Kringelbach, 2011; Sprengelmeyer et al., 2009). This sensitivity to cuteness is thought to play a role in caregiving behaviour, with positive correlations between perceptions of cuteness and motivation to view faces (Parsons, Young, Kumari, et al., 2011) as well as ratings of motivation to provide care (Glocker et al., 2009).

1.2.1 Effects of gender on sensitivity to infant cues

Across a variety of cultures and species, males and females tend to play different roles in the care of offspring (e.g., Lamb, 2012). In humans, when infants are young, mothers tend to spend more time providing care for infants than fathers do (Lamb, 1977). It has become apparent, however, that interactions with fathers, as well as mothers impact on later child development (Ramchandani, Stein, Evans, & O'Connor, 2005). Nevertheless, traditional views of caregiving roles have led to widespread hypotheses that women and men may be differentially sensitive to infant cues. Indeed, differing qualities of mother-infant compared with father-infant interactions could plausibly have origins in differential processing of infant stimuli in women compared to men. Currently, there is mixed evidence regarding the extent of gender differences in sensitivity to infant cues.

For instance, men and women are similar in their subjective appraisal of infant vocalisations, reporting similar levels of perceived distress and desire to respond (Donate-Bartfield & Passman, 1985; Leger, Thompson, Merritt, & Benz, 1996). Some evidence suggests that men and women may differ in their physiological reactions to these sounds, although the nature of these apparent gender differences remains unclear. When listening to infant cries, increases in heart rate were shown to be greater in women compared to men in one study (Furedy et al., 1989); greater in men compared to women in another (both parents and childless adults; Out, Pieper, Bakermans-Kranenburg, & Van Ijzendoorn, 2010); greater in mothers than fathers (Wiesenfeld, Malatesta, & DeLoach, 1981); or greater in fathers than mothers (Brewster, Nelson, McCanne, & Lucas…, 1998). The lack of a clear pattern of differences between men and women suggests that there may be other factors that impact on physiological reactivity to infant cries. While it may be plausible to suggest that variations in physiological reactivity to infant cry vocalisations could determine motivation or speed of behavioural responses, this hypothesis remains untested.

In contrast with the findings for infant vocalisations, a number of studies have demonstrated gender differences in the subjective appraisal of infant faces, with women being more attuned to infant cuteness (Parsons, Young, Kumari, et al., 2011; Parsons, Young, Parsons, et al., 2011; Yamamoto, Ariely, Chi, Langleben, & Elman, 2009). However, in objective behavioural measures of motivation to view infant faces, women and men do not differ in their responses to infant faces (Parsons, Young, Kumari, et al., 2011; Parsons, Young, Parsons, et al., 2011; Sprengelmeyer, Lewis, Hahn, & Perrett, 2013, but see also Hahn, Xiao, Sprengelmeyer, & Perrett, 2013). The reasons underlying the differential impact of gender on subjective appraisal and motivation to view infant faces are unclear. It is possible that lower explicit ratings of infant attractiveness in males might be a consequence of societal expectations concerning gender roles in parenting (Lamb, 1975). Alternatively, it has been suggested that women have heightened sensitivity to the physical parameters of infant cuteness (Sprengelmeyer et al., 2009). Potential differences in infant face processing between the genders warrants further investigation, but may not be as robust or as great in magnitude, as previously thought.

In summary, infant cues, and infant vocalisations in particular, are powerful attractors of attention and adults are highly sensitive to these stimuli. This sensitivity, present in non-parents and parents, men and women alike, is likely to reflect a universal reaction to ‘infancy’. It is hypothesised that the salience of infant cues initiates prompt caregiving responses. Testing whether infant cues motivate responsive behaviour requires the development of objective behavioural measures assessing adults’ speed and accuracy of physical movements. Whether or not such responses are specific to infant cues requires the assessment of adults’ performance in response to other familiar environmental cues, varying in biological salience.

1.3 Appropriate Responses to Infant Communicative Cues

In addition to prompt reactions to infant cues, responsive caregiving relies on the selection of appropriate behaviours (Bell & Ainsworth, 1972). While the basic desire to respond to infant cues may be universal in nature, the specifics of whether and how caregiving behaviour is expressed is influenced by a number of factors. This is acknowledged in theoretical models of adults' responses to infant cries, which suggest that cries elicit a motivational state and the selection of behavioural responses is modified by additional factors (Murray, 1985). These factors include the specific content of infant communicative cues and the broader context of care (Kruger & Konner, 2010; Wood & Gustafson, 2001). Parental experience also shapes caregiving behaviour as parents become particularly sensitive to their own infant's cues (Cismaresco & Montagner, 1990). In addition, caregiving behaviour has been shown to change over the course of infant development, with parents modifying their responses to communicative cues to be appropriate for the infant's maturational stage (Papousek & Papousek, 1987).

1.3.1 Interpreting Infant Vocalisations

The provision of appropriate care requires accurate assessment of the infant's current needs. Information regarding the infant's current physiological and affective state is available to caregivers through the content of infant communicative cues and knowledge of the broader context of care (Wood & Gustafson, 2001). Taking vocalisations as an example, early infant communicative cues consist of the production of cries in response to a variety of environmental stressors. Much research investigating adults' abilities to provide appropriate care in response to infant cry vocalisations has sought to identify stereotypical cry variants that may be indicative of different causes of distress (e.g., a pain cry, or a hunger cry; see Gustafson, Wood, & Green, 2000 for review). This 'cry types' conceptualisation gained support through studies that demonstrated high degrees of accuracy in naïve listeners when required to match infant cries with specific causes (e.g., Wasz-Höckert, Partanen, Vuorenkoski, Michelsson, & Valanne, 1964). However, these early studies were criticised for the inclusion of extreme examples of infant vocalisations (Soltis, 2004). Studies that used more typical ranges of cries demonstrated low accuracy rates in the identification of cause (Muller, Hollien, & Murry, 1974).

An alternative conceptualisation, dating back to Darwin in 1872, considers infant crying to be a ‘general expression of suffering’. In this respect, an infant’s cry is thought to be a ‘graded signal’, with varying acoustical properties reflecting different levels of distress (Izard, 1971). This has been demonstrated empirically in studies assessing the acoustical parameters and perceived emotional content of infant cries. Among infant cries elicited through medical procedures, a linear relationship was demonstrated between the invasiveness of the procedure and the pitch of the cry (Porter, Miller, & Marshall, 1986). Infant cries increasing in pitch were also found to be perceived by adults as reflecting greater levels of distress (Porter et al., 1986; Zeskind & Marshall, 1988). Considered in the context of evidence in favour of cry ‘types’, these findings indicate that while physical properties of infant vocalisations exist on continuous scales, stimuli at opposite extremes of these scales, when considered in isolation, may appear as discrete ‘types’ (for review, see Murray, 1979). In parallel, adults’ perception of distress seems to vary continuously as a function of physical properties, but there may be qualitative differences in their behavioural responses at extremes of this range (e.g., Zeskind & Collins, 1987).

As a graded signal, infant cries are inherently ambiguous. Further information shaping caregiving responses can be obtained from consideration of broader contextual factors. For instance, knowledge of the infant's current state can be influenced by factors such as time since the last feed or sleep (Murray, 1985). One study that assessed the impact of such factors found that adults were slower to respond to an infant's cry when informed that the infant was in need of sleep, compared to when no contextual information was provided (Wood & Gustafson, 2001). Of note, these differences in behavioural response were independent of ratings of perceived distress in cry vocalisations. This suggests that sensitivity to infant communicative cues and knowledge of contextual factors independently impact on caregiving behaviour.

1.3.2 Specific Sensitivity to 'Own Infant' Cues

Parents' early interactions with their infants lead to the rapid development of recognition abilities and heightened sensitivity to cues from their own infant. Within one week after birth, parents can identify their own infant by its cry, smell or touch alone (Cismaresco & Montagner, 1990; Kaitz, Lapidot, Bronner, & Eidelman, 1992; Porter, Cernoch, & McLaughlin, 1983). A number of studies have demonstrated that mothers have enhanced recognition abilities of their own infant, compared with fathers (Green & Gustafson, 1983; Wiesenfeld et al., 1981). However, more recent studies suggest that the development of recognition abilities does not differ by gender, but is based on the amount of exposure to infant cues. In one study, when the amount of time spent caregiving was taken into account, mothers and fathers did not differ in their ability to recognise their own infant's cry (Gustafsson, Levréro, Reby, & Mathevon, 2013). Further, both mothers and fathers have been shown to be increasingly capable of reliably recognising their own infant's cry or touch with time, indicative of a learning process (Bader & Phillips, 1999; Cismaresco & Montagner, 1990; Kaitz, Shiri, Danziger, Hershko, & Eidelman, 1994).

These learning processes may also impact on adults' perception of infant communicative cues. For example, one study demonstrated that parents have enhanced perceptual sensitivity to infant cries, compared with non-parents, and that this sensitivity was further enhanced to cues from their own infants (Green, Jones, & Gustafson, 1987). In addition, both mothers and fathers were shown to rate cries of their own infants as more unpleasant than cries of unknown infants (Wiesenfeld et al., 1981). It is plausible that this negative appraisal may be an adaptive response to enhanced perceptual sensitivity, perhaps motivating a greater need to respond to these particularly aversive stimuli. There is also evidence for further changes in parental behaviour with additional parenting experience. Experienced parents had increased response latencies to the sound of infant cries, compared to first-time parents (Donate-Bartfield & Passman, 1985).

Recent experimental evidence also demonstrated that perceived infant cuteness and motivation to view infant faces can be affected by experience (Parsons et al., in press). In this study, a computer-based model of interactions with infants demonstrated increased ratings of cuteness and motivation to view faces of infants who were experienced as having a 'happy temperament' (often smiled and laughed) compared to those with a 'sad temperament' (often frowned and cried).

1.3.3 Appropriate Parental Responses Vary Throughout Development

Parents appear to naturally adapt their behaviour to accommodate infants' developing capacities in a process termed 'intuitive parenting' (Papousek, 2007). In early interactions, parents are highly responsive to their infant's affective state, soothing negative affective behaviour and imitating positive affective behaviour (Brazelton, Koslowski, & Main, 1974; Papousek, 2007; Tronick & Gianino, 1986). As the infant develops, parental responses become increasingly delayed in time (Papousek & Papousek, 1987). As a result, the infant becomes increasingly tolerant of this delay, promoting their capacity to regulate their own emotional experiences (Denham et al., 2003; Eisenberg et al., 2003). The nature of parent-infant interactions also evolves with the infant's physical and cognitive capacities, involving fewer direct face-to-face interactions and greater incorporation of objects into play (Trevarthen, Murray, & Hubley, 1981). This natural evolution in parenting behaviour is thought to support infant emotional and cognitive development (Eisenberg et al., 2003; Ruddy & Bornstein, 1982).

Accumulating patterns of interaction over the first six months of life have an impact on the child's emotional capacities and the development of attachment relationships. The sensitivity of parents' responses to infant cues also impacts on the child's sense of efficacy in regulating affective states. By expressing their needs and receiving care infants learn effective strategies of communication (Bell & Ainsworth, 1972; Tronick, Als, & Adamson, 1978). More generally, emotional communication between parent and infant is thought to regulate the synchrony of infant and adult affective experiences, promoting a close relationship and the development of attachment (Murray & Trevarthen, 1985). Attachment to a primary caregiver is first exhibited when the infant is aged around five to six months (Bowlby, 1982). By one year of age, infants with a 'secure' attachment robustly demonstrate a preference towards their primary caregiver over other adults, seeking proximity in times of stress and reacting with distress upon separation (Bowlby, 1969, 1982). When in proximity with their primary caregiver, securely attached infants tend to display greater confidence in their independent exploration of the wider world, confident in the knowledge that parental care will be available when needed (Sroufe, 2005). Maternal sensitivity to infant cues is a key component of responsive interactions that are predictive of secure infant attachment (Ainsworth, Blehar, Waters, & Wall, 1978; Atkinson et al., 2000; Bakermans-Kranenburg et al., 2003). Highly responsive caregiving is thought to promote secure attachment through established expectations of parental behaviour and reduction in uncertainty, promoting emotional stability (Bowlby, 1969).

A secure attachment base supports the infant's exploration of the wider world, aiding the development of higher-order functions. Parent-infant interactions from around six months onwards tend to involve greater inclusion of objects in play, facilitating infant awareness and ability to interact with the wider environment (Bruner, 1975; Carpenter, Nagell, & Tomasello, 1998; Trevarthen et al., 1981; Vygotsky, 1978). Through these processes, infants develop joint attention, the capacity to share a focus of attention with another individual towards external objects (Moore & Corkum, 1994). This in turn leads to an understanding of the subjective states of others (known as intersubjectivity; Trevarthen, 1977) and finally to the developmental landmark of 'theory of mind', the ability to consider another individual's mental state (Frith & Frith, 2003; Gallagher & Frith, 2003).

1.4 Altered Responsiveness to Infant Cues in Postnatal Depression

The importance of parental responsiveness to the evolving parent-infant relationship, and further child development, is best exemplified in situations where responsiveness can become disrupted. There is an established body of evidence associating postnatal depression with difficulties in parent-infant interactions (Field, 1995; Rutter & Quinton, 1984). Postnatal depression affects approximately 13-20% of mothers and 5-10% of fathers (Field, 1995; Gavin et al., 2005; O'Hara & Mc Cabe, 2013; O'Hara & Swain, 1996; Paulson & Bazemore, 2010). Mothers with depression tend to spend less time interacting with their infants than healthy mothers do (Stein et al., 1991). When mothers with depression do interact with their infants, their behaviour is often less responsive, compared with that of healthy mothers (Murray, Fiori-Cowley, Hooper, & Cooper, 1996; Murray, Hipwell, et al., 1996).

The long-term consequences of reduced sensitivity in parent-infant interactions can include poorer child language development (Sohr-Preston & Scaramella, 2006; Stein et al., 2008) and diminished cognitive functioning (Milgrom, Westley, & Gemmill, 2004; Murray, Hipwell, et al., 1996). Less sensitive caregiving also impacts on emotional development with reduced likelihood of developing secure attachment relationships (De Wolff & Van Ijzendoorn, 1997; Mesman, van Ijzendoorn, & Bakermans-Kranenburg, 2009) and increased rates of socio-emotional problems in childhood (Murray et al., 1999; Stein et al., 2013). Maternal sensitivity to infant vocalisations appear to be of particular importance, with reduced sensitivity linked both to poorer language development at 18 months (Lester et al., 1995) and decreased likelihood of the infant developing a secure attachment at one year of age (Bigelow et al., 2010).

The relationship between depression and sensitivity of caregiving behaviour is not well established. Evidence from treatment effects suggests an indirect relationship given that treatment of either symptom alone has not been shown to lead to natural remission of the other (e.g., Poobalan et al., 2007). For instance, an intervention that was successful in reducing maternal depressive symptoms through psychological therapy found no benefit of treatment on infant cognitive development or any measure of child outcomes at five years of age (Murray, Cooper, Wilson, & Romaniuk, 2003). Conversely, an intervention that successfully improved the quality of mother-infant relationships demonstrated only a modest reduction in maternal depressed mood (Cooper et al., 2009). Importantly, not all parents with depression have difficulties interacting with their infants (Nylen, Moran, Franklin, & O'Hara, 2006). It is therefore important to identify protective factors that maintain sensitive caregiving behaviour in individuals with depression (Salekin & Lochman, 2008; Sohr-Preston & Scaramella, 2006). The following section will describe what is known about the impact of depression on both prompt and appropriate responses to infant cues.

1.4.1 Prompt Responses to Infant Cues in Postnatal Depression

Disruptions to the speed of maternal responses to infant communicative cues have been demonstrated in depression and have been shown to directly impact infant social and emotional behaviour. In observational studies, mothers with depression have been shown to respond more slowly to their infant’s distress vocalisations than healthy mothers (Bettes, 1988). Women with depression also report longer delays when responding to infant distress vocalisations than healthy women (Schuetze & Zeskind, 2001). As mentioned above (section 1.3.3) experienced parents were also shown to indicate increased response latencies to infant cries, compared with first-time parents (Donate-Bartfield &

Passman, 1985). While these results may appear inconsistent, it might be expected that the delays reported by experienced parents are minimal compared to the delays occurring in depression. However, this has not yet been empirically demonstrated.

Altered parental responsiveness has a direct impact on infant behaviour during interactions. This has been clearly demonstrated using the ‘still face paradigm’. In this experiment, mothers and infants begin interacting face-to-face and the mother is later instructed to suddenly stop responding to cues from her infant, showing instead a ‘still-face’ expression (Cohn & Tronick, 1983). Infants are sensitive to this change in behaviour, first attempting to re-engage the mother’s attention, then becoming distressed and eventually withdrawing from the interaction (Tronick et al., 1978; Weinberg & Tronick, 1996). Infants also reacted negatively when maternal reactions to infant cues were just delayed in time (rather than stopped completely), highlighting the importance of the speed and contingency of responses (Murray & Trevarthen, 1985).

The mechanisms underlying disruption of prompt responsiveness to infant cues in depression are not well-established. A few studies have assessed levels of physiological reactivity to infant cues. There is evidence linking decreased maternal heart rate reactivity to infant cries with less sensitive caregiving behaviour (Del Vecchio et al., 2009; Joosen et al., 2012). The role of depression in this proposed relationship however is unclear, with one study demonstrating reduced reactivity in depression (e.g., Riem, Pieper, Out, Bakermans-Kranenburg, & van Ijzendoorn, 2011), while another demonstrated enhanced reactivity (Donovan, Leavitt, & Walsh, 1998). While these findings appear contradictory, it is possible that both hypo- and hyper-reactivity to infant emotional cues may interfere with the provision of sensitive and appropriate caregiving behaviour (Stein, Lehtonen, Harvey, Nicol-Harper, & Craske, 2009). The development of objective behavioural studies measuring responsive behaviour may allow greater understanding of these relationships.

Current research on disrupted responsiveness in postnatal depression has focussed on caregiving behaviour. However, general slowing of motor activity and specific reductions in the speed of reactions to environmental cues are a core symptom of depression, known as ‘psychomotor disturbance’ (Taylor et al., 2006). It is plausible that delayed responses to infant cues are related to general psychomotor disturbance, rather than specific reactions to infant cues. Objective measures of adults’ psychomotor abilities in depression would allow investigation of the specificity of impaired speed of responses to infant cues.

1.4.2 Appropriate Responses to Infant Cues in Postnatal Depression

Depression also appears to interfere with the appropriateness of responses to infant cues. Observational studies have demonstrated that mothers with depression are less likely to modify their communications to be appropriate for the infant, exemplified through a reduction in the use of infant-directed speech compared with healthy mothers (Murray, Kempton, Woolgar, & Hooper, 1993). Mothers with depression are also less likely to select appropriate caregiving responses to infant distress vocalisations (Donovan et al., 1998; Zeskind & Collins, 1987).

Investigation of factors affecting the ability to respond appropriately to infant cues has suggested this may be related to disruptions to basic perceptual capacities. Mothers with depression were shown to have a reduced capacity to discriminate small pitch changes in cry vocalisations (Donovan et al., 1998) and were found to rate these sounds as less distressing and aversive compared to healthy mothers (Schuetze & Zeskind, 2001). Studies of infant facial expressions have also shown that mothers with depression have a diminished capacity to recognise positive facial expressions, compared with healthy mothers (Arteche et al., 2011). Other findings are more in line with the established negative interpretation bias widely observed in depression (Gollan, Pane, McCloskey, & Coccaro, 2008). Mothers with depression were found to perceive their infants’ behaviour and negative facial expressions more negatively than healthy mothers (Field, Morrow, & Adlestein, 1993; Stein et al., 2010). Together, these studies suggest that the interpretation of infant communicative cues may be disrupted in depression. As described above (see section 1.3) the systems supporting appropriate responsiveness to infant cues are suggested to be flexible in nature and are affected by learning processes. Disrupted perception of infant cues may therefore impact on such learning processes and consequently affect the appropriateness of caregiving behaviour.

Findings from successful intervention strategies targeting parental sensitivity further support the notion of flexible systems affecting appropriate responsiveness. Brief, focussed training interventions that promote accurate perception of infant cues can improve the sensitivity of parental behaviour (e.g., Juffer, Bakermans-Kranenburg, & van Ijzendoorn, 2007; Wendland-Carro, Piccinini, & Millar, 1999). These interventions in turn have been shown to have a long-term impact on child development, promoting the development of secure attachment relationships (Bakermans-Kranenburg et al., 2003).

The success of intervention strategies targeting perceptual skills should be considered when investigating protective factors that help to maintain sensitive caregiving despite the presence of depressive symptoms. One area of interest may lie in the large body of research that has demonstrated remarkable plasticity in human perceptual systems (e.g., Hawkey, Amitay, & Moore, 2004). A particularly striking example of this is the auditory expertise associated with being a musician (for review, see Kraus & Chandrasekaran, 2010). Musicians’ expertise is not restricted to musical stimuli, with demonstrations of high sensitivity to pure tones, speech sounds and even emotional vocalisations (Lima & Castro, 2011; Magne, Schon, & Besson, 2006; Micheyl, Delhommeau, Perrot, & Oxenham, 2006; Tervaniemi, Just, Koelsch, Widmann, & Schroger, 2005; Thompson, Schellenberg, & Husain, 2004). This is indicative of improvements in fundamental perceptual capacities, although the sensitivity of musicians to properties of infant vocalisations has not been investigated to date.

1.5. Neural Mechanisms Supporting Sensitivity to Infant Cues

Adults' responses to infant vocalisations have been previously considered as biologically-determined stereotyped responses, with infant signals described as 'innate releasing mechanisms' (Lorenz, 1950; Murray, 1979). This idea has been later refined to suggest that infant signals are 'motivational entities', promoting rapid reactions, while also instigating emotional and cognitive processing in adults (Murray, 1979). Such dissociable processing, involving an initial rapid stage, followed by a more detailed, elaborated stage, is the tenet of current major theories of emotional processing in the brain (see Section 1.1; Barrett & Bar, 2009; LeDoux, 2000; Rolls, 2000).

Over the past decade, there has been a concerted effort to examine the neural basis of human caregiving. The majority of research in this field to date has investigated neural circuitry involved in adults' responses to infant faces and voices using functional magnetic resonance imaging (fMRI). In the following section, results from studies in both visual and auditory modalities will be synthesised to propose a network of brain regions recruited by the 'parental brain'. In the existing literature, two approaches to study design have been adopted, investigating neural responses either to cues from unfamiliar infants, or to cues from parents' own infants. This division can be used to associate measured neural activity with different aspects of caregiving behaviour. Responses to unfamiliar infants can be considered as universal reactions to biologically-salient infant cues. Additional responses to own infants can be considered as learned reactions to individual characteristics.

1.5.1 Neural Responses to 'Unfamiliar' Infant Cues

Studies of infant vocalisations have largely focussed on infant cries, demonstrating increased activity in a number of frontal, temporal and subcortical regions compared to a range of physically matched control sounds. These regions include the orbitofrontal cortex (OFC; Laurent & Ablow, 2012a; Laurent & Ablow, 2012b; Lorberbaum et al., 2002), the cingulate cortex (De Pisapia et al., 2013; Laurent & Ablow, 2012a; Lorberbaum et al., 2002), prefrontal cortex (De Pisapia et al., 2013; Laurent & Ablow, 2012b; Montoya et al., 2012) and regions of the temporal lobe involved in processing of emotional stimuli (such as the superior temporal gyrus [STG] and middle temporal gyrus [MTG]; Laurent & Ablow, 2012a; Montoya et al., 2012; Seifritz et al., 2003). Subcortical regions found to be differentially reactive to infant cries compared to control sounds included the periaqueductal gray (PAG), ventral tegmental area (VTA), amygdala, thalamus and putamen (Laurent & Ablow, 2012a, 2012b; Lorberbaum et al., 2002; Montoya et al., 2012; Riem, Bakermans-Kranenburg, van Ijzendoorn, Out, & Rombouts, 2012). A small number of studies have investigated brain responses to infant laughter, highlighting roles for the amygdala and cingulate cortex (Riem, Van Ijzendoorn, et al., 2012; Sander, Frome, & Scheich, 2007; Seifritz et al., 2003).

Comparable studies of adults viewing images of emotional facial expressions of unknown infants have identified an overlapping network of regions. Greater activity to infant faces, compared with a range of control images, was observed in the OFC (Montoya et al., 2012), cingulate cortex (Caria et al., 2012; Montoya et al., 2012) and areas of the temporal lobe (MTG and STG Montoya et al., 2012; Ranote et al., 2004). In addition, activity was observed in the superior/inferior frontal gyri (Caria et al., 2012; Montoya et al., 2012), striatum (Montoya et al., 2012); insula (Caria et al., 2012), and fusiform gyrus (Caria et al., 2012).

Together, these studies highlight a network of frontal, temporal and subcortical brain areas that demonstrate enhanced activity in response to biologically salient infant cues (see Figure 3). The likely roles of these areas in the processing of infant communicative cues will be discussed in detail below (sections 1.5.3, 1.5.4, 1.5.5). In brief, frontal areas identified overlap with regions thought to be involved in social cognition, while temporal lobe areas are known to be involved in the sensory processing of emotional cues (e.g., Belin, Zatorre, & Ahad, 2002; Grandjean et al., 2005). The subcortical regions identified have been shown to be vital for the initiation of survival-related behaviours in rodent studies (Fanselow, 1994).

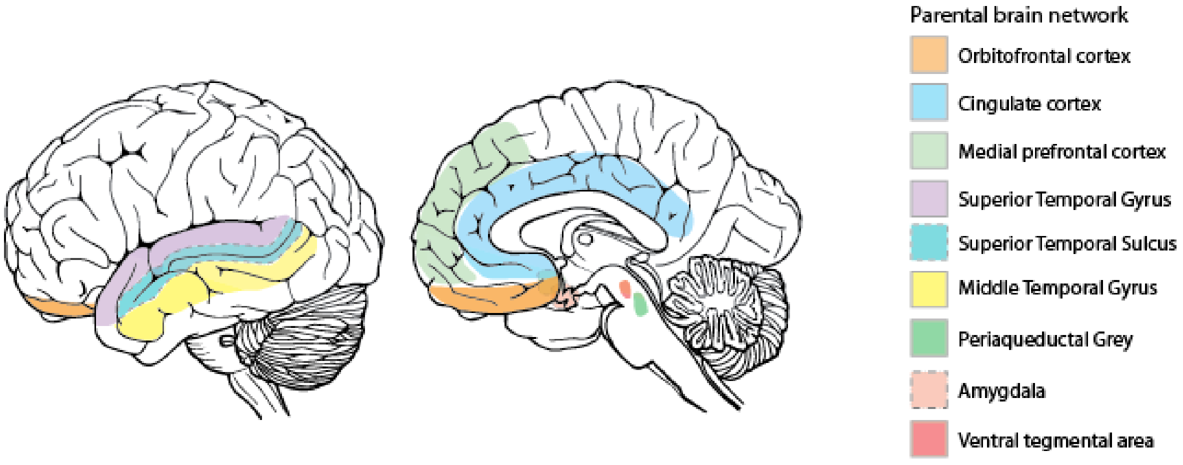

Figure 3. Networks of the 'parental brain'. Findings from fMRI studies demonstrated neural activity in the frontal and temporal lobes and subcortical brain regions when participants heard the sound of infant vocalisations or viewed images of infant faces.

Differences in neural processing between infant faces and voices included recruitment of additional temporal/frontal lobe regions in response to infant faces and greater involvement of subcortical regions in response to infant voices. Clearly, some of these differences are related to basic sensory processing (e.g., fusiform gyrus activity to face stimuli). Other variation may be related to two primary differences in experimental design across studies. First, in auditory studies, control stimuli tended to be matched to the experimental stimuli on physical dimensions (e.g., infant cries versus frequency-matched noise), whereas in visual studies, control stimuli tended to be matched on content (e.g., infant faces versus adult faces). Optimal study design would employ both types of control stimuli, physically matched and content matched, to identify networks of regions uniquely involved in processing infant cues.

1.5.2 Specialised Neural Activity in Response to 'Own Infant' Cues

While studies examining responses to unfamiliar infants may be hypothesised to relate to universal sensitivity to 'infancy', studies examining responses to familiar infants may be hypothesised to relate to learned sensitivity. Studies of own infant cries (compared with cries of unfamiliar infants) demonstrated some overlap in the network of neural regions implicated in responding to unfamiliar infant cries (section 1.5.1). These regions include areas of the frontal cortex (Kim et al., 2011; Musser, Kaiser-Laurent, & Ablow, 2012), the amygdala and a range of subcortical areas implicated in reward processing (Kim et al., 2011; Seifritz et al., 2003; Swain, Leckman, Mayes, Feldman, & Schultz, 2005). While there may be neural overlap in responses to own compared with other infant cries, individual differences across adults may affect the magnitude of activity in these regions. For instance, one study has demonstrated greater activity in the OFC, cingulate cortex, prefrontal cortex (PFC) and PAG in response to own infant cries in mothers with low cortisol responses to these sounds, compared to mothers with high cortisol responses (Laurent, Stevens, & Ablow, 2011).

For infant faces, studies have consistently demonstrated enhanced activity in 'parental brain' (section 1.5.1) regions for own infant faces compared to unfamiliar infant faces.

Heightened activity to own infant faces has been shown in the OFC (Bartels & Zeki, 2004; Nitschke et al., 2004; Noriuchi, Kikuchi, & Senoo, 2008; Ranote et al., 2004; Strathearn, Li, Fonagy, & Montague, 2008), areas of the cingulate and prefrontal cortices (Barrett et al., 2012; Bartels & Zeki, 2004; Noriuchi et al., 2008; Ranote et al., 2004; Strathearn et al., 2008), temporal lobe regions including the STG and MTG (Barrett et al., 2012; Noriuchi et al., 2008), the PAG, VTA and amygdala (Barrett et al., 2012; Bartels & Zeki, 2004; Noriuchi et al., 2008; Ranote et al., 2004; Strathearn et al., 2008). Furthermore, greater activity to own infant faces has been demonstrated in classic visual and face processing regions, including the fusiform gyrus (FFG and occipital areas; Bartels & Zeki, 2004; Nitschke et al., 2004). Combined findings across studies of own infant faces and voices suggest that responding to one's own infant does not involve activity in regions additional to those required for 'universal' responses. Rather, responding to one's own infant demands enhanced activity in those regions already recruited for general responsiveness to infancy.

1.5.3 Role of Frontal Brain Regions in Responding to Infant Cues

As a privileged category of social relationships, parent-infant interactions recruit a number of cortical regions of the social brain, including the OFC, anterior cingulate cortex (ACC) and frontal cortical regions (for reviews, see Frith, 2007; Frith & Frith, 2010; Adolphs, 2003; Amodio & Frith, 2006). Among these regions, the OFC has been described as a crucial cortical relay, uniquely placed as a nexus for sensory integration and participation in learning and decision making for social, emotional and reward-related behaviours (Kringelbach, 2005; Zald & Pardo, 2002). The OFC is thought to play a crucial role in the coordination of both rapid early responses to emotional stimuli and later, detailed processing of these cues (Barrett & Bar, 2009). Early in time, the OFC may be involved in the ‘tagging’ of salient stimuli, using predictive processes to identify stimuli from basic information received from subcortical and sensory brain areas (Bar, 2007; Chaumon, Kveraga, Barrett, & Bar, 2013). Later in time, activity in the OFC represents the reward value of stimuli, important for higher-order processes (Kringelbach & Rolls, 2003).

The strongest evidence for a role for the OFC in responding to infant cues comes from studies of infant face processing. Using magnetoencephalography (MEG) infant faces have been shown to elicit early differential activity in the medial OFC beginning at 130ms, that was not present in response to adult faces (Kringelbach et al., 2008). Further corroborating evidence for OFC involvement in rapid processing of infant face structure comes from a study of adult responses to infant faces with abnormalities (Parsons, Young, et al., 2013). Relative to healthy infant faces, OFC activity in response to infant faces with cleft lip was diminished, suggesting that the minor abnormalities in infant facial structure can alter evoked OFC activity. It has recently been suggested that the OFC is a crucial cortical hub in the overlapping networks implicated in emotional processing and social cognition that facilitate parenting (Parsons, Stark, Young, Stein, & Kringelbach, 2013).

1.5.4 Role of Temporal Lobe Regions in Processing Emotional Vocalisations

Infant vocalisations fall into the general domain of human emotional vocalisations. It therefore seems apparent that their processing would involve similar brain regions required for processing adult vocalisations. Research to date has largely supported this idea, but has suggested additional subcortical involvement in processing infant vocalisations. Studies of human emotional vocalisations have investigated neural responses to adult emotional vocalisations (such as laughs and cries) and adult speech ‘prosody’. Prosody is the intonation or melody of speech, which communicates the emotional state of the speaker, independently of semantic content.

Functional MRI studies have highlighted roles for the primary auditory cortex (A1), superior temporal sulcus (STS), STG and MTG in responding to adult emotional vocalisations and speech prosody (Aziz-Zadeh, Sheng, & Gheytanchi, 2010; Dietrich, Hertrich, Alter, Ischebeck, & Ackermann, 2007; Ethofer, Van De Ville, Scherer, & Vuilleumier, 2009; Fecteau, Armony, Joanette, & Belin, 2005; Fecteau, Belin, Joanette, & Armony, 2007; Grandjean et al., 2005; Meyer, Baumann, Wildgruber, & Alter, 2007; Mitchell, Elliott, Barry, Cruttenden, & Woodruff, 2003; Quadflieg, Mohr, Mentzel, Miltner, & Straube, 2008; Sander & Scheich, 2001, 2005). These regions overlap with findings from studies of infant vocalisation processing, which have also highlighted the A1, STS/STG and MTG (described above, section 1.5.1; Bos, Hermans, Montoya, Ramsey, & van Honk, 2010; Sander et al., 2007; Seifritz et al., 2003; Swain et al., 2008). An established model of vocal emotional processing relates early activity (around 100ms) in primary auditory areas to general sensory processing (Schirmer & Kotz, 2006). The same model also links later activity (around 200ms) in secondary auditory areas to the evaluation of emotional content of vocal cues (Schirmer & Kotz, 2006).

Beyond temporal lobe regions, studies of adult emotional vocalisations have also been shown to recruit a number of frontal areas implicated in the social brain (section 1.5.3) including the OFC, PFC, cingulate cortex and amygdala (Fecteau et al., 2005; Mitchell et al., 2003; Morris, Scott, & Dolan, 1999; Quadflieg et al., 2008; Wildgruber, Ackermann, Kreifelts, & Ethofer, 2006). Tradition models suggest a role for the OFC in the appraisal of the emotional content of vocalisations at around 400ms (Schirmer & Kotz, 2006).

However, as described above in response to visual emotional stimuli (section 1.5.3), it is possible that the OFC may be involved in the coordination of neural activity to emotionally salient stimuli earlier in time.

To date, there have been no studies directly comparing neural processing of infant and adult emotional vocalisations. While findings from studies reported in this section indicate substantial overlap in how infant and adult vocalisations are processed, there was one apparent difference. Infant vocalisations appear to evoke activity in a number of subcortical regions, including the PAG, not apparent in response to adult vocalisations (e.g., Grandjean et al., 2005; Laurent & Ablow, 2012a). Infant vocalisations may necessitate subcortical involvement because of their biological salience and importance for survival.

1.5.5 A Specific Role for the Periaqueductal Gray in Human Caregiving

The neural networks of the ‘parental brain’ identified from fMRI studies largely overlap with regions previously implicated in survival responses and emotion processing. Early neural responses are thought to underlie survival-type responses while slower, detailed analysis of stimulus properties allows refined appraisal and higher-order processing. A number of key regions of the parental brain have been identified, including the OFC and wider PFC, temporal lobe regions and subcortical structures (section 1.5.1). The following will present converging evidence for the involvement of one specific subcortical structure, the PAG.

Caregiving has been shown to recruit a number of subcortical regions involved in fundamental survival behaviours, such as the PAG, VTA and amygdala (Corter &

Fleming, 1990). Among these regions, the PAG of the midbrain has been ascribed roles in maternal behaviour (Lonstein & Stern, 1997) as well as rapid responses to salient environmental cues (Mobbs et al., 2007). The PAG is located in the tegmentum of the midbrain and receives inputs from brainstem regions, the superior and inferior colliculi, as well as a number of frontal cortical areas including the OFC (Cavada, Compañy, Tejedor, Cruz-Rizzolo, & Reinoso-Suárez, 2000; Dujardin & Jürgens, 2005). The PAG sends outputs to frontal brain areas, as well as regions of the midbrain and brainstem areas, involved in the control of autonomic functions and reflexive-type behaviour (Sillery et al., 2005). Together, these inputs and outputs point to the PAG as a unique subcortical interface between sensory and cognitive information processing (Benarroch, 2012).

The PAG has been strongly implicated in animal models of maternal caregiving behaviour. Direct evidence for its role has come from lesion studies of the PAG, demonstrating disrupted expression of maternal behaviour (Numan & Woodside, 2010). For instance, lesioning the female rodent rostro-lateral PAG can lead to diminished pup-retrieval and active nursing, while lesioning the caudal PAG can lead to increased incidence of inappropriate maternal behaviours (Sukikara, Mota-Ortiz, Baldo, Felicio, & Canteras, 2010). Although causal evidence for the PAG’s involvement in human caregiving behaviour is lacking, it has been strongly argued that the brain circuits underlying survival are conserved across species (Corter & Fleming, 1990). Indeed, evidence from human fMRI studies have identified PAG activity in response to infant cues (Bartels & Zeki, 2004; Laurent & Ablow, 2012b; Laurent et al., 2011; Noriuchi et al., 2008). These findings, along with its anatomical connections, suggest a role for the PAG in facilitating human caregiving.

Studying the PAG in humans presents a number of challenges dues to its size and location in the brain. As a small structure, the PAG can easily be overlooked among wider patterns of neural activity. In addition, localisation of any brainstem or midbrain region using fMRI is plagued by motion artefacts caused by pulse-induced brainstem movement (Zhang et al., 2006). The PAG is also not accessible to investigation using MEG or electroencephalography (EEG) as both these methods are primarily sensitive to cortical brain activity (Hillebrand & Barnes, 2002). In spite of these challenges, the studies described above did detect PAG activity in response to infant cues. Future studies should aim to prioritise investigation of this region to further the understanding of its involvement in mediating adults' responses to infant cues.

1.6 Thesis Aims and Structure

The vast majority of studies of the parent-infant relationship have been observational in nature. This approach is essential to investigating the natural course of parenting behaviour. Complementary to the wealth of previous observational work, experimental methods allow investigation and manipulation of key behavioural constructs in parenting. In this work, parental responsiveness is defined throughout in terms of ‘promptness’ and ‘appropriateness’, to provide objective, measurable constructs. Studies presented are focussed on adults’ responses to infant vocalisations as these are often the earliest available cues signalling the infant’s need for care. Infant vocalisations, and infant crying in particular, therefore constitute the primary means of initiating caregiving behaviour. This work also looks to natural changes to how adults respond to emotional vocal cues, as in depression and musical expertise, as a means of investigating the mechanisms typically underlying parenting.

The aim of this thesis is to investigate mechanisms through which adults respond to infant vocalisations. Using a series of novel behavioural tasks, this study will investigate core features of responsiveness to infant vocalisations. These studies will be carried out in adults with and without depression to assess the impact of a specific psychopathology on responsiveness. Finally, direct recordings of neural activity from human subcortical brain regions will be described, assessing the role of these structures in rapid responses to infant vocalisations.

The work in this thesis will further investigation of parental responses to infant vocalisations in five key ways. First, previous work has relied on a general conceptualisation of responsiveness, asking adults to report how they would respond to infant cues (Out, Pieper, Bakermans-Kranenburg, Zeskind, & van Ijzendoorn, 2010). Observational literature, however, specifically defines responsiveness as prompt and appropriate reactions to infant cues (Bell & Ainsworth, 1972). These components of responsiveness may impact on different aspects of parent-infant interactions. For instance, prompt responses may be required for the initiation of interaction, while appropriate responses may be required for synchronicity. Work in this thesis will examine components of responsiveness in novel experimental paradigms. Secondly, previous research primarily consisted of subjective reports of emotional reactions to infant cues. Behavioural studies in this thesis instead employ objective measures of core skills hypothesised to underlie responsive caregiving. This will allow specific aspects of responsive care to be measured that are challenging to approach with typical self-report questionnaires.

Thirdly, observational studies have demonstrated associations between parental postnatal depression and disruptions to responsiveness of caregiving behaviour. The mechanisms through which depression impacts on the quality of caregiving behaviour, however, remain unclear. This thesis will examine the impact of depression on aspects of responsiveness to infant vocalisations. Through comparison of findings from adults with and without depression, disrupted capacities will be identified. Fourthly, previous reporting of neural networks involved in responding to infant cues is largely based on correlational measures of patterns of blood flow in the brain. The final study in this thesis will directly investigate neural activity in subcortical brain regions in response to infant vocalisations using intracranial recordings. This approach will provide relatively localised, temporally-sensitive measurements of brain activity. Finally, the current work uses multiple sound exemplars from a large database of infant, adult and animal vocalisations. This database, designed and constructed as part of this thesis, provides high quality, authentic and recognisable sounds, with which to examine the specificity of effects in relation to infant cues.

Experiments in this thesis aim to investigate features of adults' responsiveness to infant vocalisations, as well as potential neural mechanisms implicated. Systematic investigation of these processes requires the development of a stimulus database of infant vocalisations, which will be described in Chapter 2. A study investigating prompt responsiveness to infant vocalisations and the impact of depression on this process is then described (Chapter 3). The following chapter (Chapter 4) describes a series of experiments investigating appropriate responses to infant vocalisations focussing on the impact of depression and auditory training on sensitivity to features of these sounds. Chapter 5 describes findings from a unique study investigating differential sensitivity of subcortical human brain areas in response to infant vocalisations using direct recordings of neural activity.

CHAPTER 2: DEVELOPMENT OF THE OXFORD VOCAL SOUNDS ('OXVOC') DATABASE

This chapter describes the development and standardisation of a database of affective vocalisations from infants, adults and domestic animals. Spontaneous infant vocalisations varying in affect were acquired from video recordings of infants interacting with a caregiver in their own home. Natural (non-acted) adult 'affective burst' vocalisations and domestic animal vocalisations were collected from online sources. As this thesis is focussed on investigating responses to infant vocalisations, the inclusion of adult and domestic animal vocalisations provided control stimuli varying in developmental stage and species. Stimuli were first physically standardised for duration, onset/offset amplitude envelope and overall intensity. All stimuli were then rated by participants on three dimensions relating to the emotional content of the sounds. These ratings were used to select 15 exemplar stimuli from each of six categories of emotional vocalisation: infant cries, infant laughs, infant neutral vocalisations, adult (female) cries, cat meows and dog whines. Stimuli were then rated by a separate group of participants for valence in two separate sessions to provide a measure of the test-retest reliability. The same participants then categorised each of the sounds (as 'distressed', 'happy' or 'no emotion') to provide a measure of stimulus validity as an exemplar of the predetermined category. Finally, key physical parameters of stimuli (such as pitch and burst duration) were extracted and compared with perceptual ratings to investigate the role of these parameters in emotion perception across different stimulus categories.

2.1 Introduction

Darwin (1872) was the first to describe and classify human emotional facial expressions and vocalisations as biologically determined expressions of inner emotions. Human facial expressions have since been the subject of a vast body of scientific literature, demonstrating that basic categories of emotional facial expressions are highly recognisable and preserved across cultures (for review, see Ekman, 1993). Although studied to a lesser extent, vocal expressions of emotion can also be readily categorised and feature cross-culturally (e.g., Sauter, Eisner, Calder, & Scott, 2010; Scherer, Banse, & Wallbott, 2001). Nonverbal emotional vocalisations emerge at the earliest point of life and are present throughout the lifespan in the majority of mammalian species. These sounds therefore constitute an important category of stimuli, forming the basis of communicative and emotional interactions during infancy and developing into important features of social interaction in later life. The presence of comparable vocalisations in other mammals allows investigation of the extent to which such basic emotional interactions are species-specific.

2.1.1 Infant Vocalisations

In early infancy, vocalisations communicate varying degrees of distress, providing vital information for caregivers about infants' needs (Soltis, 2004). The range of vocalisations that an infant can produce is necessarily constrained both by the physiology of the developing vocal tract and the degree of muscular control over the vocal chords. As the infant develops, the range of vocal expressions becomes more sophisticated, incorporating positive expressions such as laughter at around four months of age (Darwin, 1872; Nwokah, Hsu, Dobrowolska, & Fogel, 1994; Sroufe & Wunsch, 1972). While crying draws infant and parent together for the necessity of care for survival, laughter promotes positive, rewarding social interactions between parent and infant. Around the same age, infants begin to ‘babble’ (Oller & Eilers, 1988). Babbling comprises phoneme-like sounds, often repeated in meaningless strings, and is thought to be the precursor to language development (Petitto & Marentette, 1991). When infants ‘babble’, they tend to be neither distressed nor excited, but in an emotionally neutral state (Rothgänger, 2003). This chapter describes the collection and standardisation of a range of infant emotional vocalisations, selected for being positive, negative and neutral. This is in line with a number of studies of infant facial expressions, using positive, negative and neutral stimuli (e.g., Arteche et al., 2011; Stein et al., 2010).

2.1.2 Adult Emotional Vocalisations

While there is much to be learned from the study of infant vocalisations alone, the presence and development of emotional vocalisations throughout life provides the opportunity to investigate how responses to these sounds change. Previous investigation of changes in responsiveness to emotional vocalisations through development has not been possible due to a lack of suitably controlled and matched stimuli. In the domain of facial research, this question has proven to be of much interest, with studies demonstrating that the degree of ‘baby-ish’ facial features present in later life can impact on perception of attractiveness and behavioural attitudes towards individuals (Cunningham, Barbee, & Pike, 1990). Similar studies investigating the role of juvenile vocal patterns in adults would provide an interesting perspective on the recent findings identifying characteristics of attractive adult voices (Bruckert et al., 2010; Feinberg et al., 2006).

Adult vocalisations, both verbal and non-verbal, communicate a much wider range of emotions than infants, including fear, anger, surprise, disgust, boredom and pleasure (Scherer et al., 2001). Emotion can be communicated through vocalisations in two forms:

verbal prosody and non-verbal affective bursts (Scherer, 1994). Verbal prosody refers to the intonation and melody of speech that communicates the affective state of the speaker on a separate ‘channel’ to linguistic content. Non-verbal affective bursts tend to be more spontaneous, less constrained expressions of emotion that do not contain speech (such as crying, screams, laughter, sighing) and are similar across cultures (Schröder, 2003). These vocalisations typically accompany intense emotional states, and closely parallel affective vocalisations from other species (Scherer, 1995).

2.1.3 Existing Stimulus Databases

The majority of research into human emotional vocalisations to date has focussed on verbal prosody (for review, see Scherer, 2003). Sound databases for this purpose generally contain vocalisations from actors speaking semantically neutral sentences or strings of ‘pseudo-words’ with different types of prosody (e.g., the Danish Emotional Speech Database, Engberg and Hansen, 1996; the Berlin Database of Emotional Speech, Burkhardt, Paeschke, Rolfes, Sendlmeier, & Weiss, 2005). Controlling for semantic content in this manner ensures that there is no interaction between meaning and prosody. However, such sentences have the disadvantage of being unnatural because semantic and prosodic features typically correspond (Scherer, Ladd, & Silverman, 1984). Another common method, using ‘pseudo-sentences’ (sentences made up of non-words), again avoids semantic confounds, but is limited in terms of how natural the vocalisations sound. A second issue is that even pseudo-words often contain features that are language-specific, such as legal phoneme groups. In summary, the study of verbal prosody is, by necessity, entangled with features of the language being studied.

As a consequence, there has been increased interest in the use of nonverbal affective bursts to study emotion processing (Belin, Fillion-Bilodeau, & Gosselin, 2008; Lima, Castro, & Scott, 2013). These bursts constitute a primitive and universal mode of communication (Sauter et al., 2010) and parallel the vocalisations of other species (e.g., Belin et al., 2008; Juslin & Laukka, 2003). Affective bursts are fundamentally different to speech in their underlying production mechanisms (Scott, Sauter, & McGettigan, 2010). Speech requires that specific physical parameters are produced, whereas affective bursts do not, allowing for the presence of a greater range of vocal features. The relative lack of language specific features means that sounds can be used to study universal or cross-cultural perception of emotion. Furthermore, some categories of vocal bursts, such as laughs and cries, are thought to reflect pure, 'raw' emotion and therefore are powerful stimuli to use when investigating responses to emotion (Scherer, 1995).

The available databases of affective bursts contain sounds generated by adult actors (e.g., the Montreal Affective Voices [MAV], Belin et al., 2008; the Corpus of Nonverbal Vocalizations, Lima et al., 2013; the Geneva Multimodal Emotional Portrayal Core Set [GEMEP-CS], Banziger, Mortillaro, & Scherer, 2012). The MAV and GEMEP-CS contain vocalisations from 10 actors consisting of the vowel sound /a/ in different emotional categories. The Corpus of Nonverbal Vocalisations (Lima et al., 2013) contains acted affective bursts representing a range of positive and negative emotions. These databases are comparable with well-established databases of varying facial expressions from the visual domain (e.g., Pictures of Facial Affect, Ekman, 1993; NimStim Face Stimulus Set, Tottenham et al., 2009). A major advantage of these databases is that they include sounds with a range of emotions, which overall, are highly recognisable (Belin et al., 2008; Lima et al., 2013).

The use of actors to obtain affective vocalisation bursts is not without difficulty. Acted vocalisations are, by definition, not spontaneous emotional expressions. Reliance on acted vocalisations assumes that acted and authentic vocalisations are the same, or that differences in authenticity are imperceptible. However, there is growing evidence to suggest that authenticity in emotional vocalisations can be readily detected by adult listeners (Barker, 2013). Acted emotional expressions have been shown to be perceived as sounding more stereotyped and exaggerated (Laukka, Neiberg, Forsell, Karlsson, & Elenius, 2011) and more extreme (Barkhuysen, Krahmer, & Swerts, 2007) compared to authentic ones.

A number of physical differences between acted and authentic vocalisations have also been described, such as acted vocalisations having a higher and more variable pitch (Audibert, Aubergé, & Rilliard, 2010; Jürgens, Hammerschmidt, & Fischer, 2011). In addition, distinct patterns of neuronal activity, including increased activity in frontal brain regions, have been found when participants listen to authentic compared to acted speech segments (Drolet, Schubotz, & Fischer, 2012), further supporting the notion of differential processing and that differences in authenticity may be perceptible.

It has been suggested that some of these features are unavoidable in acted vocalisations, perhaps due to the lack of physiological reactions that accompany genuine emotional vocalisations (Kreibig, 2010; Scherer, 2003). These physiological reactions are likely to be difficult to imitate in acted vocalisations as they are not fully under voluntary control (Juslin & Laukka, 2001; Scherer, 1986). Given this limitation, there is a clear advantage in using authentic affective vocalisations in studies of emotional processing. In the current stimulus database, non-acted adult vocalisations were included.

Only negative adult affective vocalisations are described below, as these were the only stimuli employed in studies included in this thesis. These vocalisations consisted of female adult cries (sobs, rather than screams). They are a useful category of control stimuli when investigating responses to infant cries as they provide an additional category of negative emotional vocal stimuli from conspecifics.

2.1.4 Domestic Animal Vocalisations

Vocalisations from animals, particularly of the domestic kind, share a number of similarities with infant vocalisations. They convey information about an animal’s current needs and often serve to initiate caregiving responses. In particular, domestic cat meows and dog whines are familiar, easily recognisable sounds that tend to elicit sympathetic and caregiving responses from human hosts (Pongrácz, Molnár, Miklósi, & Csányi, 2005). For instance, domestic cats can effectively use their characteristic vocalisations, purrs, to solicit human care (McComb, Taylor, Wilson, & Charlton, 2009). Despite some functional overlap between animal and infant vocalisations (in eliciting care from human adults), these sounds are clearly distinctive and easily differentiated from human vocalisations. For these reasons, distress vocalisations from domestic animals provide a suitable category of comparison stimuli to investigate the species-specific effects of processing of emotional vocalisations.

2.1.5 Physical Parameters of Human Emotional Vocalisations

Vocalisations varying in emotional content differ on a number of physical parameters, such as pitch, harmonic content, intensity, timbre and temporal characteristics. Further understanding of the mechanisms underlying adults’ perception and interpretation of emotional vocalisations can be gained through a systematic analysis of the physical properties that make up these sounds. The dynamic nature of auditory stimuli presents a number of challenges for the measurement of acoustic vocal features of interest. At a single point in time, sounds are defined by their frequency content. Over time, the distribution of power in frequency content can vary affecting parameters such as the pitch, intensity and timbre of a sound. Despite these challenges, robust associations have been observed between physical parameters and perceived emotional content.

One physical parameter that has been the focus of much research is the fundamental frequency (F_0) of vocalisations. Fundamental frequency is related to the perceptual property of a sound's pitch. Higher F_0 in adult vocalisations has been reliably linked to perception of greater stress (Protopapas & Lieberman, 1997), higher emotional arousal (Bänziger & Scherer, 2005) and more negative valence (Sauter et al., 2010). Similar associations have been demonstrated in infant vocalisations, with greater F_0 linked to higher levels of perceived distress (Donovan et al., 1998; Schuetze & Zeskind, 2001). The speed and intensity of emotional vocalisations also vary with emotion. In adult speech, the length of voiced periods varies with emotion: sadder speech tends to involve longer voiced bursts (Huttar, 1968) while happy, angry or excited speech involves shorter, faster voiced bursts (Banse & Scherer, 1996). In terms of loudness, increased range of intensity is generally perceived as increased 'emotional involvement', or more intense emotion (for review, see Murray & Arnott, 1993). Comparison of physical parameters with perceived emotional content of sounds across different types of emotional vocalisations could inform key processes involved in the interpretation of vocal emotion.

2.1.6 Aims and Hypotheses

The database described in this chapter (the Oxford Vocal Sounds, ‘OxVoc’ database) contains clips of infant, adult and domestic animal vocalisations. These categories were selected for use in studies of the processing of infant distress vocalisations. Infant laughter and neutral vocalisations provide comparative infant vocal emotional and non-emotional stimuli. Adult distress cries were selected as a different class of distress stimuli from conspecifics that do not elicit parental caregiving responses. Finally, domestic animal distress sounds were utilised as a category of familiar distress stimuli from another species. Methods of collecting infant cry, laughter and babble vocalisations are described, as well as the acquisition of adult female cry vocalisations and distress vocalisations from domestic animals. The physical standardisation procedures of these stimuli are then described, followed by details on the investigation of the perceived emotion in these sounds. Investigation of the reliability and validity of these perceptual ratings is also detailed. Finally, a comparison of physical acoustic properties and psychological ratings is described.

2.2 Stimulus Collection and Physical Standardisation

2.2.1 Infant Vocalisation Stimuli

Infant vocalisations were obtained from video recordings of infants filmed in their own homes during a play and feeding session with their primary caregiver. This play and feeding session included periods of interactive play, a mealtime and brief separation of the infant and caregiver. These video recordings were collected as part of a previous project (The Oxford Parent Project) with the approval of the Oxford Research Ethics Committee. Recordings from nine infants were used for the current database. All infants were full-term, healthy and aged between six and eight months at the time of recording (*M* = 6.70 months, *SD* = 0.91). Audio recordings were extracted from these videos and 84 clips of vocalisations, free from background noise, were selected (see Appendix A for example waveforms and spectrograms).

2.2.2 Adult and Animal Distress Vocalisation Stimuli

Adult and animal distress vocalisations were obtained from freely available internet sources (e.g., www.freesound.org). Individuals who uploaded content used in this database were contacted and their permission obtained for the anonymous use of short clips of distress vocalisations for research purposes. Adult distress cries were obtained from video blogs consisting of individuals speaking directly to a camera. Video blogs were searched for emotional segments and clips of crying were extracted. No videos of male adults crying were found, so all 23 cry clips extracted came from videos posted by females, aged approximately 18-30 years. The cause of crying was variable across individuals, but often occurred when describing upsetting life events. Animal distress vocalisations were obtained from recordings of pet cats and dogs posted by their owners on various online forums. Twenty-three samples of cat meows and 22 samples of dog whines were obtained (see Appendix A for example waveforms and spectrograms).

2.2.3 Physical Standardisation of Stimuli

The duration of stimuli was standardised using Audacity software (version 1.3.4-Beta, http://audacity.sourceforge.net/) and matching of all other physical parameters was completed using Adobe® Audition® software (CS5 v4.0, Adobe System Corporation, San Jose, CA; see Figure 4). All selected stimuli were free from background noise and were cropped to 1500ms in length. The onset and offset amplitude envelope was standardised across all stimuli by applying linear rise and fall times of 150ms to the start and end of each clip ('Volume Envelope' function). The rise-time or 'onset attack' of a sound is known to affect its timbre – the tonal quality or 'production' of a sound. While different types of voice production tend to vary little in timbre, acoustic differences in rise time can be readily detected and have been demonstrated to affect the latency of early brain responses to sounds (e.g., Caclin et al., 2006). In addition, any unnatural effects of stimulus clipping on the onset of stimuli were corrected by applying standardised rise/fall times.

Matching Stimulus 'Loudness'

Stimuli were matched for total root-mean-square (RMS) amplitude for each clip to -25dBFS (decibels full scale). This method matched the overall intensity of the clip, while retaining variations in amplitude envelope across time. While sound stimuli can be readily matched for intensity, an objective measure of the energy within a sound, there is some degree of variability in individuals' sensitivity to sound intensity. Frequently, different individuals can perceive the same sound as varying in ‘loudness’. The subjective nature of stimulus loudness must be accounted for in auditory studies of emotional stimuli as loudness can impact on perceived intensity of the emotion within a sound (Murray & Arnott, 1993) and especially loud sounds can be uncomfortable and unpleasant to listen to.

The intensity of stimulus presentation was adjusted for each individual participant in the experiments described in this thesis. To achieve comparable ‘loudness’ across participants, each individual’s threshold of hearing was measured and sounds were presented at 70dB above this level. A number of psychoacoustic methods for testing absolute threshold of hearing are available, all involving presentation of sounds at varying intensities and assessing an individual’s ability to hear each. The method implemented here was an adapted version of a ‘staircase’ procedure (Levitt, 1971), using a sample 1.5s infant cry stimulus. This implementation was designed to rapidly assess hearing thresholds. While standard measures of psychoacoustic abilities involve extensive testing, the short version implemented here was sufficient for the purpose of establishing an accurate threshold to account for inter-individual variability in sensitivity to intensity. This task was programmed and presented using Presentation® software (Neurobehavioral Systems Inc.) on a laptop computer with Sony in-ear earphones (MDR-EX77LP).

This task consisted of a two-alternative forced-choice, two-down, one-up design. On each trial in this procedure, a red cross appeared centrally onscreen while the sound clip was presented. Participants were then prompted to press a button to indicate whether they had heard it or not. After each positive response (when participants reported that they could hear a sound) the intensity on the following trial was decreased by 10dB. This procedure continued until the participants reported that they could no longer hear the sound. From this point (the ‘reversal’ point), subsequent sounds were presented at increasing intensities (5dB increase on each trial), until the participant reported that they could hear the sound. This also constituted a reversal point, after which sounds were presented at decreasing intensities again. Criterion for absolute threshold of hearing was set at two consecutive ‘up/down’ reversals (the point at which participants report they could hear a sound, after not hearing the previous sounds) at the same value. The task continued running until this criterion was reached, lasting typically for 20-25 trials.

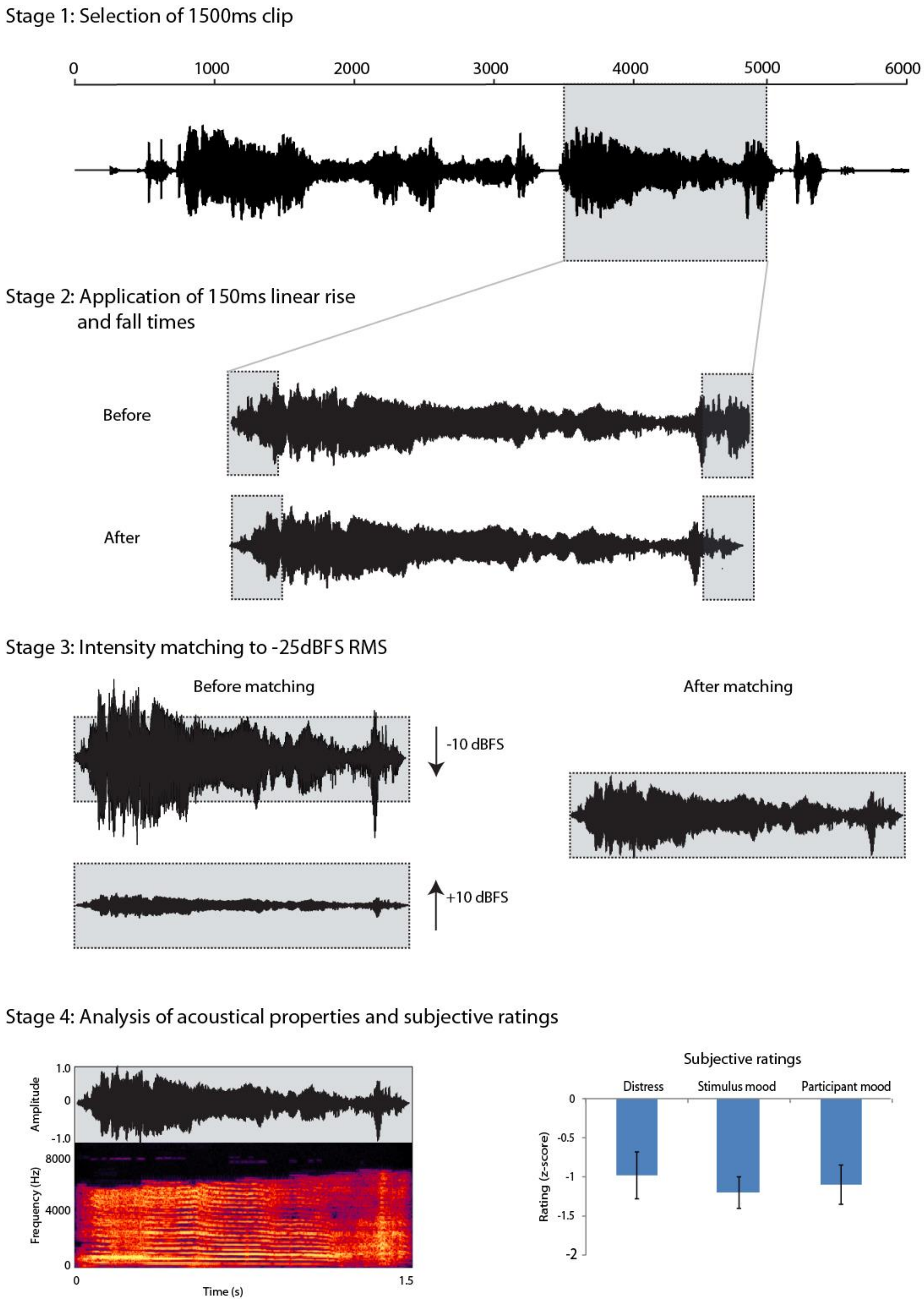

Figure 4. Pipeline of stimulus physical standardisation procedure. 1500ms stimuli were first selected from original recordings. Each selected affective burst had 150ms linear rise and fall times applied to change beginning and end of amplitude envelope, ensuring a smooth 'fade in' and 'fade out' of sounds. Stimuli were then matched for overall intensity and finally, physical parameters were assessed and subjective ratings collected. (Error bars represent mean +/- standard error)

Validity of Intensity Changes

The integrity of intensity changes implemented through Presentation® software was assessed using an ‘artificial ear’ (Bruel & Kjaer, Model 4153). An artificial ear contains a microphone to detect auditory stimuli and an acoustic coupler with a number of additional cavities that mimic the transduction effects of the human outer ear and auditory canal. This allows precise electroacoustic measurements of sound intensities as they would actually be experienced at the eardrum. Measurements using this type of device allow ecologically valid assessments of the intensity of sounds experienced by the listener in dBSPL (dB sound pressure level). Stimuli varying in intensity (adjusted using Presentation’s ‘global attenuation’ setting) were presented to the artificial ear and a sound pressure level meter was used to record peak intensity values. This analysis demonstrated that linear changes in amplitude settings in Presentation software (in dBFS) translated directly into linear changes in intensity at the eardrum in dBSPL.

2.3 Perceptual Ratings of Vocalisation Stimuli

Selection and validation of stimuli for use in the current ‘OxVoc’ database was carried out in a two-stage process (comparable to Bänziger et al., 2012). First, groups of participants provided ratings on a large set of stimuli on three dimensions in order to select stimuli representative of categories of emotional vocalisations. As the majority of stimuli were distress vocalisations, participants rated all stimuli on the level of perceived distress. Participants also rated the perceived valence of each sound stimulus and their own emotional reaction to the sound. These ratings aimed to provide measures of both perceived mood of the individual producing the vocalisation, and a measure of how much this affected the participants’ own subjective mood. Based on these ratings, groups of 15 stimuli representing different emotional categories were selected. Following this procedure, an independent sample of participants then rated all stimuli for valence on two occasions as a measure of reliability of perceived emotional content. The same participants also completed a categorisation task with these same stimuli, providing a measure of validity of stimulus labelling. All participants were recruited through poster and email adverts in University Departments and Colleges.

2.3.1 Experiment 1: Methods of Rating Dimensions of Subjective Emotional Experience

Ratings of Infant Vocalisations

Sixty-one participants (30 male; aged *M* = 27.58 years, *SD* = 6.87) rated 84 infant vocalisation stimuli. These stimuli were pre-categorised by two researchers using the original audio-visual clips as examples of cries (*n* = 21), laughs (*n* = 18), neutral ‘babbles’ (*n* = 25) and ‘difficult to classify’ (*n* = 20). For each stimulus, participants rated “How distressed do you think the baby was?” on a 7-point Likert scale, with 1 as ‘not distressed’ and 7 as ‘very distressed’. Participants also responded to “Please rate the mood of the baby” and “How did you find the sound?” on 9-point Likert scales, with -4 as ‘very negative’ and +4 as ‘very positive’. Participants listened to each stimulus once, after which they completed all three ratings before moving on to the next stimulus. The order of stimuli was pseudo-randomised across participants. Stimuli were presented using WinAmp software (Winamp v 5.61 Nullsoft, Inc., http://www.winamp.com), through Sony in-ear earphones (MDR-EX77LP). Responses were made by circling the appropriate value on the Likert scale using pen and paper. This procedure took approximately 15 minutes in total. For ease of comparison, all data were converted to standardised scores (z-scores) and de-trended to ensure that a z-score of zero equated to a neutral rating.

Ratings of Adult and Domestic Animal Vocalisations

An independent sample of 25 participants (nine male; age *M* = 27.67 years, *SD* = 11.72) rated 68 adult and animal distress vocalisations. This set of stimuli consisted of female adult distress cries (‘sobs’ as opposed to screams, *n* = 23), cat meows (*n* = 23) and dog whines (*n* = 22). A similar set of ratings to that described above for the infant vocalisations was collected. Participants responded to: i) “How distressed do you think the adult/animal was?” from ‘very distressed’ to ‘not distressed’; ii) “Please rate the mood of the animal”, from ‘very negative’ to ‘very positive’; and iii) “How did you find the sound?” from ‘very negative’ to ‘very positive’.

Participants responded using computerised visual analogue scales (VAS), implemented as vertical onscreen bars. A grey bar indicated the total range of possible responses, while a slightly narrower white bar, centred on top of the grey bar indicated the participants’ current response. The height of the white bar could be altered by using the ‘UP’ and

‘DOWN’ arrows on the keyboard. At the beginning of each rating, the white bar was half of the height of the grey bar, indicating a neutral response.

After each stimulus was presented, individual scales appeared on the screen along with the question to be rated. This task was self-paced; participants moved from one rating to the next by pressing the spacebar. It lasted approximately 12 minutes in total. After a single rating was completed, the next scale appeared onscreen along with the new question to be rated (for an individual stimulus, previous ratings remained visible onscreen). Upon completion of all three ratings for a single stimulus, ratings and questions were removed from the screen and replaced with a fixation cross while the next stimulus was presented. Participants’ responses were converted to standardised scores and de-trended as described above.

2.3.2 Experiment 1: Results of Ratings

Ratings of Infant Vocalisations

Inter-participant reliability was demonstrated to be very high (intra-class correlation, ICC(2, 61) = .99) so data from all participants were combined. Ratings were used to identify three groups of 15 stimuli that fell into distinct affective categories: negative (infant cries), positive (infant laughs) and neutral (infant ‘babbles’). Cries were identified as those rated as most distressed (*M* = -1.07, *SD* = .27), with lowest perceived infant mood (*M* = -1.75, *SD* = .23), and most negative listener response (*M* = -1.72, *SD* = .24). Laugh stimuli were those rated as low in distress (*M* = 1.24, *SD* = .21), with positive infant mood (*M* = 1.10, *SD* = .37) and most positive listener response (*M* = .98, *SD* = .36). Neutral stimuli were those rated as low in distress (*M* = 0.76, *SD* = .24), with perceived neutral infant mood (*M* = .01, *SD* = .32) and neutral listener response (*M* = .02, *SD* = .31). Gender differences were investigated for ratings on all stimulus dimensions for the final set of 45 stimuli. Independent samples t-tests demonstrated no significant gender differences in ratings of infant distress (*t*(59) = 1.71, *p* = .09, *r* = .22). While not significant, comparison of means demonstrated that males tended to rate infant vocalisations as sounding more distressed (*M* = 1.12; *SD* = .59) than females (*M* = .80; *SD* = .84). There were no significant gender differences, or differences approaching significance in measures of perceived infant mood (*t*(59) = 0.66, *p* = .51, *r* = .09) or participant mood (*t*(59) = 0.25, *p* = .80, *r* = .03).

Ratings of Adult and Domestic Animal Vocalisations

Fifteen stimuli in each of the adult cry, cat meow and dog whine categories were also selected. These stimuli were selected by ranking stimuli on response values on each dimension. Ranks were summed across dimensions and the most negatively rated stimuli in each category selected. Inter-participant reliability (*n* = 25) was very high so data were combined (ICC(2, 25) = .91). As distress vocalisations, stimuli were selected as those with: i) highest perceived distress (adult, *M* = -1.49, *SD* = .18; cat, *M* = -.84, *SD* = .39; dog, *M* = -1.19, *SD* = .18), ii) lowest adult/animal mood (adult, *M* = -1.78, *SD* = .20; cat, *M* = -.95, *SD* = .47; dog, *M* = -1.04, *SD* = .20) and iii) most negative listener response (adult, *M* = -1.83, *SD* = .25; cat, *M* = -1.04, *SD* = .49; dog, *M* = -1.54, *SD* = .25).

As with the infant vocalisation ratings, gender differences were also assessed for the adult and domestic animal vocalisations ratings using independent samples t-tests. These analyses revealed no significant gender effects in ratings of distress (*t*(23) = 1.55, *p* = .14, *r* = .31), perceived stimulus mood (*t*(23) = 1.53, *p* = .14, *r* = .30) and participant mood (*t*(23) = 1.56, *p* = .26, *r* = .31). While differences were not significant at an α-level of 0.05, inspection of effect sizes demonstrated medium-sized effects of gender (*r* > .30). Investigation of the trends in this data demonstrated that, compared with males, females rated adult and animal vocalisations: i) as sounding more distressed (females: *M* = -2.32, *SD* = .61; males: *M* = -1.91, *SD* = .68); ii) with more negative stimulus mood (females: *M* = -2.29, *SD* = .61; males: *M* = -1.90, *SD* = .63), and iii) with more negative participant mood (females: *M* = -2.10, *SD* = .51; males: *M* = -1.84, *SD* = .57).

Comparison of Dimensions of Subjective Emotional Experience

Mean ratings within the selected stimuli were compared across all stimulus categories using a one-way ANOVA. Results demonstrated a significant main effect of stimulus category for ratings of distress (*F*(5, 89) = 301.50, *p* < .001, *r* = .88), stimulus mood (*F*(5, 89) = 198.29, *p* < .001, *r* = .83) and participant mood (*F*(5, 89) = 176.45, *p* < .001, *r* = .82).

Post-hoc Bonferonni-corrected pairwise comparisons demonstrated significant differences (*p* < .05) in distress ratings for all human vocalisation categories (see Figure 5), with lowest distress for infant laughter, followed by infant neutral vocalisations, infant cries and finally adult cries. Perceived distress in infant cries did not differ from perceived distress of either cat meows (*p* = .22) or dog whines (*p* = 1.00). Distress in the animal vocalisations was significantly higher than that for infant laughs and neutral vocalisations, and significantly lower than distress in adult cries (*p* < .05). For the stimulus mood ratings, there were significant differences across all stimulus categories (*p* < .05) apart from infant cries and adults cries (*p* = 1.00). The adult and infant cry vocalisations were perceived as reflecting similarly negative mood and were significantly more negative than animal distress vocalisations (*p* < .05). Finally, for the 'participant mood' ratings, responses were similar for infant cries, adult cries and dog whines (all *p* > .25). All other pairwise comparisons for the participant mood ratings were significantly different (*p* < .05).

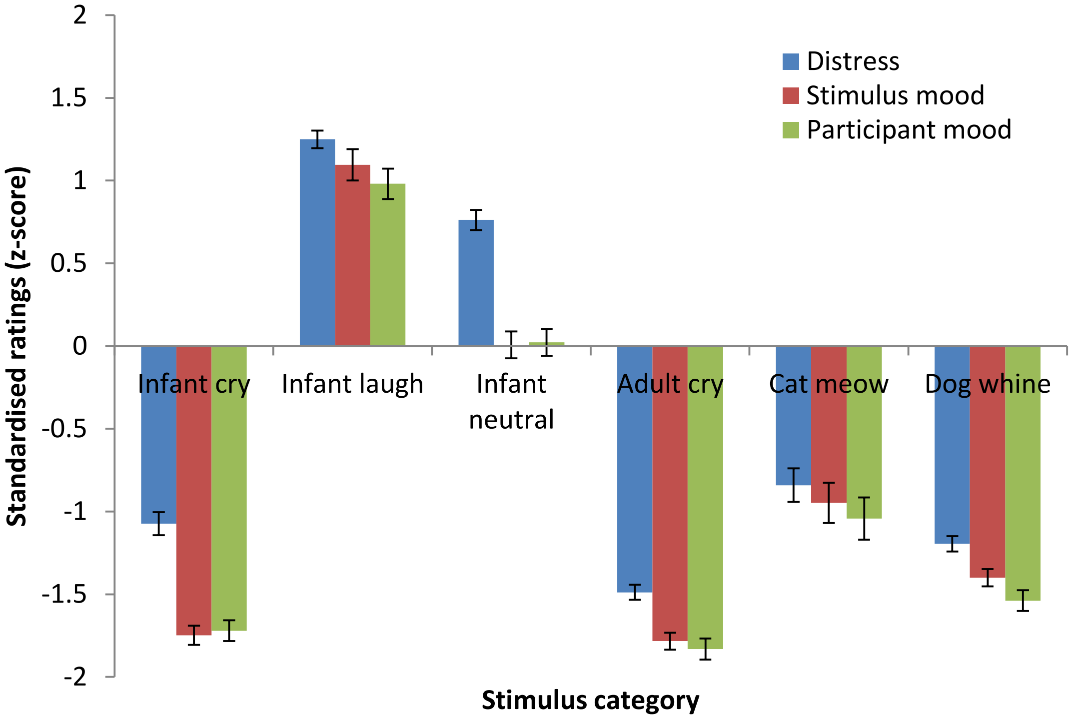

Figure 5. Mean ratings of distress, stimulus mood and participant mood across stimulus categories. Negative ratings reflect greater distress, more negative stimulus valence and more negative emotional response to stimuli. Infant cries, adult cries, cat meows and dog whines were all rated as high in distress and low in stimulus valence and participant emotional response. Infant laughter stimuli were low in distress, with positive stimulus valence and participant emotional response. Infant neutral stimuli were low in distress, with neutral stimulus valence and participant emotional response. Error bars represent mean +/- standard error.

Relationships between perceptual dimensions measured were assessed using Pearson correlations. All measures were found to be significantly highly correlated (all $r > .95$, $p < .001$). Given this finding, a factor analysis was performed to investigate whether these measures loaded onto a single underlying variable. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was 0.73 and Bartlett's test of sphericity was significant, ($\chi^2(3) = 641.12$, $p < .001$) justifying dimension reduction. A principle components analysis revealed the first factor explained 98.52% of the variance across ratings. Comparison of stimulus categories was then carried out using this general 'valence' dimension, using a one-way ANOVA with stimulus category as a factor. This analysis demonstrated a significant main effect of stimulus category on perceived 'valence' ($F(5, 89) = 225.60$, $p < .001$, $r = .85$). Post-hoc Bonferroni analyses demonstrated that perceived 'valence' was similar for infant cries, adult cries and dog whines (all $p > .95$). Cat meows were perceived as significantly less negative than infant cries ($p < .001$). Perceived valence was significantly different for infant neutral vocalisations and infant laughter compared with all other sound categories ($p < .001$).

2.3.3 Experiment 2: Methods of Assessing Reliability and Validity

The following measures of reliability and validity were collected in accordance with procedures described for the widely-used 'NimStim Face Stimulus Set' (Tottenham et al., 2009). The set of 90 stimuli selected in Experiment 1 were rated and categorised by an independent sample of 38 participants (21 male; aged *M* = 30.29 years, *SD* = 17.01). Given the findings from Experiment 1 demonstrating an underlying factor of stimulus 'valence', participants were asked to rate each stimulus on a single 'valence' dimension ("Please indicate how happy or sad you think the baby/adult/animal is", from 'very sad' to 'very happy'). These ratings were implemented using computerised VASs described above. After rating each of the stimuli once for valence, there was a 20-minute break and participants then rated each of the stimuli for valence again. A 20-minute break between ratings has previously been used to assess reliability of other comparable stimulus databases (e.g., Tottenham et al., 2009). The reliability of stimulus ratings were assessed using correlation analyses of ratings at different time points.

Following this, participants were also asked to categorise each of the stimuli to provide a measure of the validity of stimulus labelling. Participants heard each stimulus again and were asked to select one label to describe the sound from four possible alternatives using pen and paper: 'cry/distressed', 'laugh', 'sound with no emotion', or 'could not categorise'. Accuracy and chance-corrected accuracy scores (Cohen's kappa) were then calculated for each stimulus. Chance-corrected accuracy scores provide a conservative, standardised measure of performance, taking into account the likelihood of correct answers occurring by chance. The order of stimulus presentation was randomised during the reliability ratings and pseudo-randomised across participants for the validity ratings. In total, this procedure lasted around 55 minutes (including the 20 minute break).

The validity of stimulus categorisation was assessed by examining how well participants could categorise the sounds, using probability correct scores and chance-corrected scores (Cohen's kappa coefficient). Cohen's kappa coefficient is calculated as:

$$\mathrm{K} = (\mathrm{Pr}(a) - \mathrm{Pr}(e))/(1\text{-}\mathrm{Pr}(e)),$$

where $\mathrm{Pr}(a)$ is the observed proportion correct, and $\mathrm{Pr}(e)$ is the probability of a correct answer by chance. In this study, there were four possible categorisations, so $\mathrm{Pr}(e)$ was equal to .25.

2.3.4 Experiment 2: Results of Reliability and Validity Analyses

Inter-participant reliability was found to be very high, so data were averaged across all 38 raters (ICC(2, 28) = .98). For the infant vocalisations, cries were perceived as negatively valenced (infant cries, *M* = -1.30, *SD* = .37), laughs were perceived as positively valenced (infant laughter, *M* = 1.06, *SD* = .32) and non-emotional 'babble' vocalisations were perceived as neutrally valenced (*M* = .04, *SD* = .31). Adult and animal vocalisations were perceived to have negative valence (adult cry, *M* = -1.84, *SD* = .29; cat meows, *M* = -.81, *SD* = .29; dog whines, *M* = -1.05, *SD* = .16).

A one-way ANOVA with 'stimulus category' as a factor was used to compare valence ratings. Results demonstrated a significant main effect of stimulus category (*F*(5, 89) = 186.92, *p* < .001, *r* = .82). Post-hoc Bonferroni pairwise comparisons revealed significant differences in valence ratings across all stimulus categories, except between infant cries and dog whines (*p* = .31), and between cat meows and dog whines (*p* = .48). Adult cries were perceived to be significantly more negative than either infant cries or animal distress vocalisations.

Reliability of Perceived Emotional Valence

First and second ratings of all 90 stimuli were compared using Pearson's correlation coefficients to assess the reliability of perception of emotional content of vocalisations over time. Ratings between time one and time two were significantly highly correlated (*r*(90) = .99, *p* < .001). Ratings for subcategories of stimuli were also significantly positively correlated between time one and time two (all *r* > .72).

Validity of Stimulus Categorisation

Overall, the proportion of correctly classified sounds was high (*M* = .76, *SD* = .29), with chance-corrected scores greater than K = .65 (*M* = .68, *SD* = .38). Assessment by stimulus category revealed high accuracy scores for all stimulus types (see Table 1), except for infant neutral vocalisations, accuracy for which was very low (proportion correct, *M* = .20, *SD* = .11). Chance-corrected scores for infant neutral vocalisations demonstrated performance was lower than chance. Closer inspection revealed that responses were split evenly across all possible categories (cry, *M* = .32, *SD* = .18; laugh, *M* = .25, *SD* = .13; sound with no emotion, *M* = .20, *SD* = .11; cannot categorise, *M* = .23, *SD* = .06).

Table 1. *Mean stimulus categorisation performance by vocalisation type*

Vocalisation Type	Proportion correct (*M*, *SD*)	Chance-corrected probability (*M*, *SD*)
Infant Cry	0.93, 0.10	0.91, 0.14
Infant Laugh	0.86, 0.15	0.81, 0.20
Infant Babble	0.20, 0.11	-0.07, 0.15
Adult Cry	0.98, 0.04	0.97, 0.05
Cat Meow	0.73, 0.15	0.64, 0.19
Dog Whine	0.85, 0.10	0.80, 0.13

Effects of Gender on Perceived Emotional Valence

Ratings of stimulus valence were compared between men and women. An independent samples t-test demonstrated no significant difference in ratings by male and female participants ($t(36) = .27$, $p = .79$, $r = .04$). Splitting data by sub-category also did not reveal any significant gender effects (all $p > .09$). Independent t-tests comparing accuracy in categorisation between male and female participants demonstrated no significant differences in any of the stimulus subcategories (all $p > .05$).

2.4 Comparison of Physical Parameters

All vocalisation stimuli were assessed on a number of different physical parameters to investigate the association between their physical properties and the perceived emotional qualities of sounds. The properties extracted were fundamental frequency (F_0), the average duration of each vocal burst within the standardised 1500ms clip, the average number of bursts in each clip and the peak amplitude in each vocalisation. Differences were first assessed across stimulus categories to investigate basic categorical differences in physical parameters. The values of physical properties were then compared with the perceptual ratings described in the previous section in order to investigate the relationship between physical features of sounds and the emotion contained therein.

2.4.1 Methods of Assessing Physical Properties

Physical properties of stimuli were measured using Adobe® Audition® (AA) software. Accurate measurement of F_0 in natural vocalisations is restricted due to the varying frequency content of these sounds over time. Stimulus F_0 was approximated by calculating Fast Fourier Transforms (FFTs) using a Hamming window (sampling rate 48000Hz, window size 4096 samples, approx. 85ms; AA ‘Frequency Analysis’ tool). The first major peak was extracted from resulting frequency plots as a measure of the F_0 of the stimulus. Typically, this first peak was equivalent to an amplitude greater than -45dB (an arbitrary value selected on the basis of visual inspection of a number of frequency plots). A preliminary test of this procedure was carried out by manually identifying the first peak from the spectrogram. Where the method proved inaccurate (e.g. if there was no major peak greater than -45dB), an approximation to a more accurate value was calculated by lowering the amplitude cut-off until a major peak was apparent within the 100-700Hz F_0 range of human vocalisations (Gustafson & Green, 1989; Lederman, 2010).

Peak amplitudes were extracted (AA ‘Match Volume’ tool) and number of bursts and average burst length were measured through inspection of individual stimulus waveforms. A single burst was defined as a visibly distinct amplitude wave, separated by a period of 10ms or more of low amplitude sound (less than 2% of maximum amplitude). The number of bursts per stimulus was recorded as well as the duration of each burst, from which the mean burst duration was calculated.

2.4.2 Results of Comparisons of Physical Properties

The differences in physical parameters across stimulus categories were assessed using one-way ANOVAs, with stimulus category as a factor. Differences between pairs of categories were then investigated using post-hoc Bonferroni pairwise comparisons.

Fundamental Frequency

In general, the F_0 of infant cries and animal distress sounds were higher than those of infant laughter and neutral vocalisations or adult cries. A one-way ANOVA revealed a significant main effect of stimulus category ($F(5, 89) = 13.90$, $p < .001$, $r = .37$; see Table 2). Post-hoc comparisons demonstrated that infant cries had significantly higher F_0 than either infant laughter ($p = .001$) or neutral vocalisations ($p < .001$). Infant cries were also higher pitched than adult cries ($p = .001$), but not significantly different in their frequency range from animal sounds ($p = 1.00$). Infant neutral and laughter vocalisations were in a comparable frequency range to adult cries (all $p > .05$).

Burst Duration

There were significant main effects of stimulus category for average burst length ($F(5, 89) = 10.36$, $p < .001$, $r = .32$). Average burst length was longest for cat meows and infant cries, followed by infant neutral vocalisations and dog whines and finally adult cries and infant laughter (see Table 2). Cat meows, infant cries and infant neutral vocalisations had significantly longer burst lengths than infant laughter ($p < .05$). Additionally, burst length for cat meows and infant cries were significantly longer than adult cries ($p < .05$). All other pairwise comparisons were not significantly different ($p > .05$).

Number of Bursts

There was also a significant main effect of average number of bursts ($F(5, 89) = 8.84$, $p < .001$, $r = .30$) across stimulus categories. Infant laughter had significantly more bursts than any other stimulus category (all $p < .05$). There were no significant differences in the number of bursts in infant cries, infant neutral vocalisations, adult cries, cat meows or dog whines (all $p > .05$; see Table 2).

Peak Amplitude

Finally, there was also a significant main effect of stimulus category on peak amplitude ($F(5, 89) = 5.54$, $p < .001$, $r = .24$). Peak amplitudes were significantly higher for infant laughter compared to cat meows ($p < .01$) or dog whines ($p = .02$). Similarly, adult cry peak amplitudes were also significantly higher than that for cat meows or dog whines ($p < .01$). Peak amplitudes for infant cries and infant neutral vocalisations were not significantly different from any other category ($p > .05$; see Table 2).

Table 2. *Key physical parameters of vocalisations by category*

Vocalisation Type	F_0 (Hz) (*M, SD*)	Burst duration (s) (*M, SD*)	No. of bursts (*M, SD*)	Peak amplitude, dBFS (*M, SD*)
Infant Cry	444.30, 43.16	1.06, 0.47	1.73, 0.88	-5.97, 1.77
Infant Laugh	337.50, 30.75	0.42, 0.23	3.33, 0.88	-4.94, 1.80
Infant Babble	307.81, 45.92	0.95, 0.43	1.80, 1.50	-5.75, 2.80
Adult Cry	339.84, 64.32	0.57, 0.13	2.13, 0.35	-4.77, 1.30
Cat Meow	406.98, 126.88	1.21, 0.35	1.27, 0.46	-7.77, 1.74
Dog Whine	471.58, 54.93	0.82, 0.41	2.00, 0.85	-7.43, 2.56

Note. F_0 = fundamental frequency.

Comparison of physical parameters of infant vocalisations with perceptual ratings

Pearson correlation analyses were performed to compare the physical parameters of infant vocalisation stimuli with perceived valence. Results demonstrated significant negative correlations between perceived valence and F_0 ($r(45) = -.68$, $p < .001$; see Figure 6) as well as between perceived valence and burst duration ($r(45) = -.47$, $p = .001$). With increasing pitch and burst duration, infant vocalisations were perceived to have more negative valence. There was a positive correlation between perceived valence and number of bursts ($r(45) = .44$, $p = .003$) indicating that shorter, more frequent bursts were perceived as having more positive valence. There was no significant association between valence and peak amplitude ($r(45) = -.14$, $p = .37$).

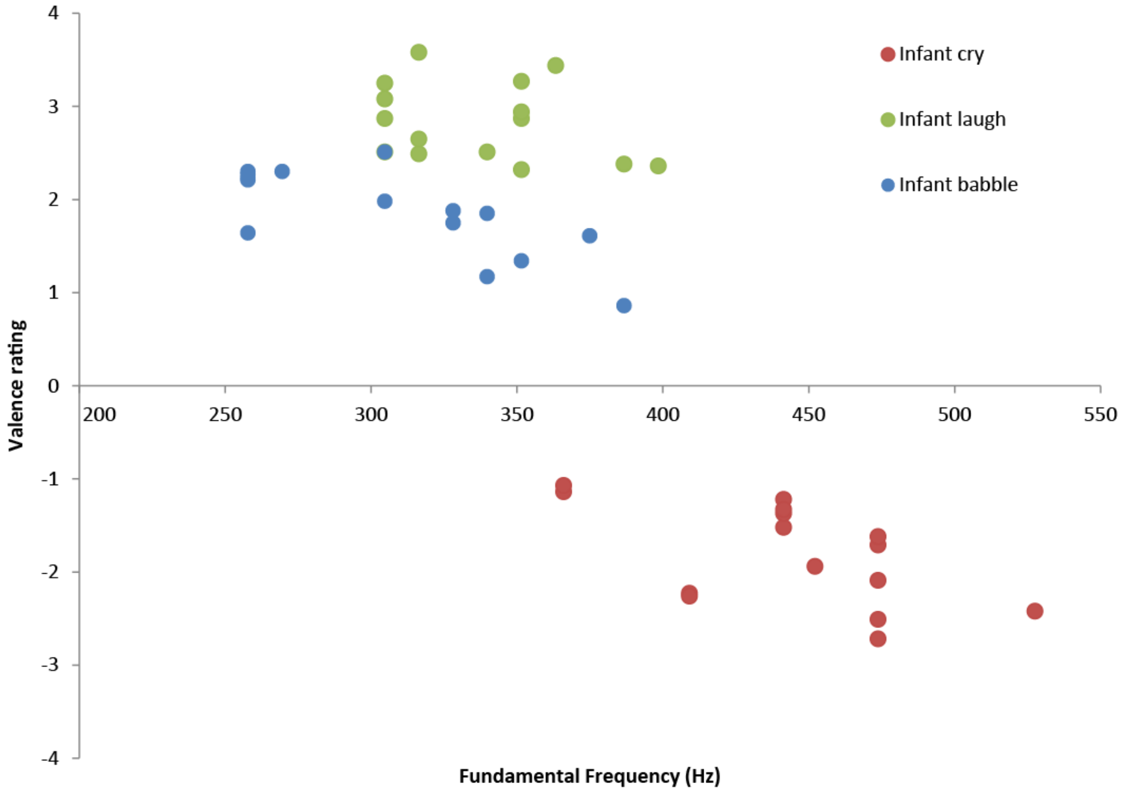

Figure 6. Comparison of fundamental frequency with ratings of valence in infant vocalisations. The overall correlation between fundamental frequency and valence was significant ($p < .001$). Infant laugh and babble vocalisations generally had lower fundamental frequencies and were rated with more positive valence compared to infant cry vocalisations. The valence rating scale spanned from ‘very negative’ (-4) to ‘very positive’ (+4).

2.5 Summary of Standardisation Procedures

A standardised database of authentic emotional vocalisation stimuli from infants, adults and domestic animals was established. Based on perceptual ratings, statistically discrete categories of negative, positive and neutral infant vocalisations were identified. The same procedure was used to identify groups of negatively valenced adult and domestic animal vocalisations. Exploratory analysis of the perceptual dimensions used, measuring perceived distress, perceived mood of the vocalisation and participant’s own mood after listening to the stimulus revealed a single underlying factor reflecting stimulus ‘valence’. Comparison of stimuli in different categories based on this factor demonstrated that infant cries, adult cries and dog whines were all perceived as similarly negative. Cat meows were perceived as significantly less negative than these categories, but more negative than either infant laughter or infant neutral vocalisations. Infant laughter and infant neutral vocalisations also differed significantly from one another in their perceived valence.

In a second experiment, stimuli were assessed for the reliability of perceived emotional valence over time and the validity of emotional categorisation. Valence ratings performed by the same participants at two time-points demonstrated high reliability of perceived valence. Comparison of stimulus categories based on these ratings demonstrated significant differences between all categories, except infant cries, cat meows and dog whines. Adult cries were perceived as significantly more negative than these categories of sounds. Infant laughs and neutral vocalisations were rated significantly differently from each other and were both more positively rated than the other stimulus categories. Stimulus categorisation validity was shown to be high, except for the neutral infant vocalisations which were correctly classified on only 20% of trials overall. There were no significant effects of participant gender in either of these experiments.

Comparison of stimuli based on perceptual properties revealed a number of effects. Infant cries, cat meows and dog whines were found to be characterised by high pitch, in comparison to infant laughter, infant neutral vocalisations and adult cries. Shorter burst duration differentiated infant laughter and adult cries from other stimuli, while infant laughter additionally had more bursts per clip. Finally, infant laughter and adult cries were found to vary most in amplitude, while cat meows and dog whines varied the least. Comparison of physical parameters of infant vocalisations with ratings of valence demonstrated that infant vocalisations with higher pitch and with longer, fewer vocalisation bursts were associated with most negative valence ratings.

2.6 Discussion

A standardised stimulus database of affective vocalisations from infants, adults and domestic animals was developed. Stimuli were selected based on ratings of perceived emotion and were shown to demonstrate high reliability of ratings over two sessions and high overall validity of emotional categorisation. Key physical properties of stimuli were then compared across stimulus categories and with perceptual ratings in order to investigate the physical parameters involved in perceiving emotion in vocalisations.

This is the first standardised database containing infant vocalisations varying in affect and is therefore comparable to current databases of adult vocalisations (Lima et al., 2013), infant facial expressions (Kringelbach et al., 2008) and adult facial expressions (Tottenham et al., 2009). These standardised infant vocal stimuli will allow future investigation of how both parents and non-parents perceive and respond to infant vocal emotion. On the basis of perceptual ratings of distress, perceived mood and listener mood, stimuli were categorised as negative sounds (cries), positive sounds (laughs) and neutral sounds (babbles).

Comparison of physical parameters of infant vocalisations demonstrated that cries were characterised by higher pitch than laughs or neutral vocalisations. Previous work has demonstrated that within infant cry vocalisations, pitch indicates the level of perceived infant distress (Schuetze & Zeskind, 2001). The current work extends these findings to incorporate a wider range of infant vocal emotion. Infant laughs were differentiated from neutral vocalisations on the basis of temporal characteristics (shorter, more frequent bursts). These findings are comparable to other studies demonstrating shorter, more frequent vocal bursts in happy, compared with sad, adult speech (Banse & Scherer, 1996).

Together, these findings indicate that generally in human vocalisations, high pitch indicates distress, while rapid temporal dynamics indicates excitement and happiness.

Assessment of the physical properties of adult cries and animal distress vocalisations demonstrated that adult cries were characteristically low in pitch, with short burst duration. Animal distress sounds had comparable pitch and burst duration to infant cries. These patterns of pitch differences across infants, adults and animals are in line with previous findings demonstrating that voice pitch scales with the size of the animal (e.g., Smith, Patterson, Turner, Kawahara, & Irino, 2005). Adult cries were among the most negatively rated stimuli. This is in direct contrast with the findings from infant vocalisations, demonstrating that stimuli with higher pitch and longer burst duration were rated more negatively. This may be representative of a developmental change in the production of vocal emotion, perhaps in part related to changes in the length and muscular control over the vocal chords (Kuhl & Meltzoff, 1982). Perception of emotion in vocalisations is therefore more complex than simple associations between physical properties and perceived emotional content, and is likely to be highly dependent on context, varying perhaps both developmentally and across species.

Notably, the infant neutral vocalisations were the only category of stimuli with low accuracy scores on a stimulus categorisation task. Participants' performance was at chance levels when categorising these stimuli, with approximately equal responses identifying these sounds as positive, negative, neutral or 'could not categorise'. This suggests that infant neutral vocalisations are an inherently ambiguous category of stimuli. Other studies have found comparable effects, demonstrating that babbles can comprise a full range of vocal emotions, including positive, neutral and negative affect, perhaps as a precursor to prosody (Oller et al., 2013). The ambiguity in these vocalisations would be of much interest in future studies looking at changes in sensitivity to infant cues through parental experience, or in groups likely to demonstrate disrupted sensitivity, such as in the case of postnatal depression.

The inclusion of negatively valenced adult cry and animal distress vocalisations allowed comparisons between responses to stimuli varying in developmental stage and species. Perceptual ratings of perceived emotion in these sounds had considerable overlap with the perception of infant cries. The use of adult cries and animal distress vocalisations as suitable control stimuli in studies of infant emotion is therefore warranted. From Experiment 1, ratings combined into a single ‘valence’ factor demonstrated similarities in the ratings of infant cries, adult cries and dog whines, while cat meows were perceived as less negative than these sounds. In Experiment 2, where a single ‘valence’ rating was performed by participants, similar valence was perceived in infant cries, dog whines and cat meows, while adult cries were perceived more negatively. The difference in findings between these two studies is likely due to procedural differences between the two studies, with participants in Experiment 1 rating either infant sounds or adult/animal sounds, while in Experiment 2 all participants rated all stimuli. Ratings for Experiment 2 are therefore thought to be more robust.

In Experiment 2, adult cries were perceived to be more negative than infant cries. The cause of this difference in perceived valence was not investigated. It might be hypothesised that these differences are related to individuals’ attributions regarding the cause of the vocalisations. For example, the cause of infant crying is often related to readily controllable environmental factors, such as hunger or separation from a caregiver.

The cause of adult crying on the other hand, is often related less controllable factors, such as loss or failure. It might be anticipated that, in line with attribution theory (Weiner, 1985), the greater ‘controllability’ of infant crying compared with adult crying would impact on the appraisal of these sounds. Indeed, previous research has demonstrated that even within infant vocalisations, individuals’ perceived efficacy in caregiving, and presumably their ability to terminate infant distress, was related to how negatively infant cries were rated (e.g., Verhage, Oosterman, & Schuengel, 2013).

While there were no significant gender differences in ratings of emotional content of affective vocalisations, results from some analyses suggested trends towards significant effects. In Experiment 1, males tended to rate infant vocalisations as sounding more distressed and negative than females did. Conversely, females tended to rate adult and animal vocalisations as sounding more distressed and negative than males did. The ability to draw conclusions from these findings is restricted by the use of different samples of participants to rate different stimuli and relatively small sample sizes. Comparison of gender differences in Experiment 2 (in which all participants rated all stimuli on a single ‘valence’ dimensions) demonstrated no significant gender differences in the perceived valence of infant, adult or domestic animal vocalisations. Together, these findings present no reliable evidence for gender differences in perceived emotional content of affective vocalisations.

2.6.1 Limitations of Experimental Procedures

The main limitations of the experiments described here concern the range of stimuli included. First, while infant vocalisation stimuli included a range of positive, negative and neutral affect, adult and animal vocalisations were limited to negative affect. The range of negative adult vocalisations was further limited to female cry sounds only. Future inclusion of adult vocalisations from males and females varying in affect would increase the versatility of this stimulus database.

A second issue concerns the extent to which stimuli are representative of a full range of emotional vocalisations. While efforts were made to collect authentic emotional expressions, the method used was naturally one of convenience. In addition, by selecting stimuli that represent discrete categories of emotional vocalisation, the extent to which these stimuli reflect continuous, natural variation in vocal cues is further limited. These limitations would be addressed by the addition of a broader range of stimuli, obtained from a greater number of individuals varying in age and from different linguistic backgrounds.

Inclusion of a greater range of stimuli would also allow studies investigating developmental changes in the production and perception of emotional vocalisations. In the context of infant vocalisations, addition of stimuli produced at different points during the first 18 months of life would be of particular value. During this period, infant vocalisations vary dramatically (Kuhl, 2004; Ostwald & Murray, 1985; Soltis, 2004), as do parental responses to these sounds (Papousek, 2007; Papousek & Papousek, 1989). Systematic analysis of these vocalisations and how they are perceived by adults with and without parenting experience would provide valuable insight into the processes by which infant vocal cues may guide caregiving behaviour.

Finally, investigation of perceived stimulus properties was limited to perception of one dimension, namely valence. Classic models define emotional experiences along orthogonal dimensions of valence and arousal (e.g., Watson, Wiese, Vaidya, & Tellegen, 1999).

Future work might benefit from the inclusion of ratings of arousal to allow more comprehensive modelling of perceived emotion within human and non-human vocalisations. The relationship between perceived arousal and physical properties of vocalisation stimuli may also provide a useful future avenue to enhance the understanding of mechanisms involved.

2.6.2 Conclusions and Future Directions

Stimuli presented in this chapter represent a range of spontaneous infant affective vocalisations and comparable negative adult and domestic animal vocalisations. Future addition of other comparable stimuli would allow for a range of studies investigating the development of mechanisms involved in the production and perception of affective vocalisations. A database of this type would also be of use to investigate how different groups of individuals might vary in their perception of emotional vocal cues. Clear starting points would be to investigate how perception of emotional vocalisations varies across cultures, or in individuals with altered experiences of emotion, such as in mood disorders. In the context of infant vocalisations specifically, comparison of individuals varying in parental experience would be of particular interest.

CHAPTER 3: THE IMPACT OF INFANT VOCALISATIONS ON EFFORTFUL MOTOR PERFORMANCE

The sound of a crying infant is a powerful attractor of attention, serving to alert nearby adults to the needs of an infant. Responsive caregiving behaviour requires reactions to infant cues that are both prompt and appropriate (Ainsworth et al., 1974). In the context of postnatal depression, findings from observational studies have suggested that there are specific impairments in the speed of parents' responses to their infant's vocalisations (Bettes, 1988). However, the causal pathway linking depression with an impaired quality of caregiving is not well understood. This aim of this chapter was to investigate effortful features of behavioural responsiveness following exposure to infant vocalisations, which may be disrupted in depression. A novel behavioural task was used to assess the speed and effort of motor performance after listening to infant cries and other sounds, in adults with and without depression.

3.1 Introduction

A distressed infants’ cry is characterised by high and variable pitch, with higher pitch, longer duration and greater dysphonation generally signalling greater distress (Soltis, 2004). Adults often report the sound of a crying infant as irritating, annoying, distressing, urgent and aversive (e.g., Frodi et al., 1978; Soltis, 2004). In addition, healthy adults generally report a desire to carry out a variety of caregiving behaviours upon hearing an infant cry (Schuetze & Zeskind, 2001). There is some evidence to suggest that in depression, perception of infant cries is somewhat altered.

Women with depression have been shown to rate cries as less perceptually salient (Schuetze & Zeskind, 2001) and less likely to initiate a caregiving response, compared with healthy women. In depression, processing of emotional stimuli, such as social cues (e.g., Emerson, Harrison, & Everhart, 1999; Gollan et al., 2008; Leppänen, Milders, Bell, Terriere, & Hietanen, 2004) and music (e.g., Naranjo et al., 2011) has been shown to be negatively biased. Responses to negative emotional stimuli are often reported to be particularly disrupted. Individuals with depression have been shown demonstrate attentional biases towards negative stimuli in general and disrupted interpretation of negative facial expression specifically (e.g., Gotlib, Krasnoperova, Yue, & Joormann, 2004; Gur et al., 1992) . Mothers with depression have been shown to rate sad infant facial expressions more negatively than healthy mothers (Stein et al., 2010).There is evidence to suggest that processing of positive infant emotional stimuli might also be disrupted, as mothers with depression have been shown to be less accurate at identifying happy infant facial expressions than healthy mothers (Arteche et al., 2011). It is possible that some of the difficulties that parents with depression experience when caring for infants are related to the disrupted processing of emotional cues.

Previous studies investigating behavioural responses to infant vocalisations have demonstrated reduced speed of responses with either with parental experience or the presence of mood disorders (Bettes, 1988; Schuetze & Zeskind, 2001). The nature of these changes clearly deserves further exploration, although it might be hypothesised that the reduction in speed related to parental experience may be substantially less than that occurring in depression. These studies, however, are limited by their use of self-report measures of the speed of caregiving responses. Further understanding of the mechanisms supporting effortful caregiving behaviour may be obtained through the use of objective behavioural measures.

3.1.1 Physiological Reactivity to Infant Cries

Studies investigating mechanisms of rapid responding to infant vocalisations to date have focussed on the role of physiological arousal in response to the sound of infant crying. There is a widely held assumption that rapid responding to infant cues is achieved partly through eliciting physiological reactions in adult listeners (Frodi et al., 1978; Out, Pieper, Bakermans-Kranenburg, & Van Ijzendoorn, 2010). A variety of measures have been used to study physiological changes in response to infant cries, including: heart rate, blood pressure, skin conductance, respiratory sinus arrhythmia and hand grip force (e.g., Bakermans-Kranenburg, van Ijzendoorn, Riem, Tops, & Alink, 2012; Frodi et al., 1978; Joosen et al., 2012; Out, Pieper, Bakermans-Kranenburg, & Van Ijzendoorn, 2010).

The most common physiological measure used to assess the effects of hearing infant cry vocalisations is heart rate, or heart rate reactivity (a measure of heart rate during a period of interest relative to baseline). A number of studies have shown significant increases in heart rate in mothers listening to their own or other infants' cries (Bleichfeld & Moely, 1984; Del Vecchio et al., 2009; Frodi & Lamb, 1980; Furedy et al., 1989; Wiesenfeld et al., 1981). Other studies have shown either no change in heart rate (Brewster et al., 1998) or heart rate deceleration (Wiesenfeld et al., 1981) in mothers listening to the sound of unknown infants.

Two studies aiming to test the link between heart rate reactivity and measures of caregiving suggest that increases in heart rate promote adaptive caregiving responses. In one study, heart rate increases were found to be greater in mothers rated as 'highly sensitive' during an observed play session with their infants, compared with 'less sensitive' mothers (Joosen et al., 2012). Similarly, mothers with greater heart rate reactivity showed decreased latency to initiate a caregiving response to the sound of an infant crying (Del Vecchio et al., 2009). In line with these findings, a study of childless women found that heart rate reactivity was decreased in a group of women with subclinical depressive symptoms relative to healthy women (Riem et al., 2011). Other studies have demonstrated decreased heart rate reactivity in adults with depression, relative to healthy participants, in response to other stressful stimuli, such as a public speaking task (Salomon, Clift, Karlsdóttir, & Rottenberg, 2009). This perhaps suggests that diminished physiological reactivity might account for some of the problems adults with depression experience when providing care for an infant.

In contrast, another study found that heart rate reactivity was actually greater in a group of mothers who were more 'depression-prone' compared to healthy mothers (Donovan et al., 1998). These studies of women with depressive symptoms did not include any behavioural measure of caregiving responses, limiting the conclusions that can be drawn from these apparently contradictory findings. In addition, a few studies have suggested that increased heart rate reactivity can be observed in adults who physically abuse children (for review, see McCanne & Hagstrom, 1996). The link between increased heart rate reactivity and more appropriate, sensitive caregiving behaviour is far from clear cut.

The pattern of responses is further complicated when considering parental status and gender differences. One study found increased heart rate in mothers and decreased heart rate in childless women (Banse & Scherer, 1996), while another demonstrated increased heart rate in all adults, but larger increases in childless adults than in parents while listening to infant cries (Out, Pieper, Bakermans-Kranenburg, & Van Ijzendoorn, 2010). Across the genders, heart rate increases have been observed in childless women but not men (Furedy et al., 1989) and also in mothers but not fathers (Wiesenfeld et al., 1981). In contrast, other studies have shown heart rate increases in fathers but not mothers (Ekman, 1993) and larger heart rate increases in men compared to women (both parents and childless adults; Out, Pieper, Bakermans-Kranenburg, & Van Ijzendoorn, 2010). One recent study suggested a potential genetic role in physiological reactivity to infant cries. Childless women with one variant of the oxytocin receptor gene (OXTR GG) showed greater increases in heart rate during exposure to infant cries than women with the other (AA/AG) alleles (Riem et al., 2011). This group difference was not apparent in women with depressive symptoms. Together, these studies suggest that individual characteristics and genetics can play a role in the anticipated physiological effects of exposure to infant distress cries. The interaction of these characteristics with depressive symptoms has not been fully investigated.

Other measures of arousal, such as skin conductance, may provide more specific assessment of ‘arousal’ states (e.g., Berntson, Boysen, & Cacioppo, 1993; Craske et al., 2008). Studies of responses to infant cries using these methodologies, while limited in number, do seem to tell a more consistent story. Exposure to infant cries has been shown to result in increased skin conductance in mothers and fathers (Frodi et al., 1978; Wiesenfeld et al., 1981). Other results suggest that there are increases in blood pressure in parents (Frodi et al., 1978), stronger respiratory sinus arrhythmia withdrawal in mothers (Joosen et al., 2012) and increased hand grip force in childless women (Bakermans-Kranenburg et al., 2012) during exposure to infant cries. Together, these findings suggest that hearing the sound of a crying infant can initiate a state of heightened physiological arousal. The involvement of changes in heart rate in this response, and consequent caregiving behaviour, however, requires further careful analysis.

3.1.2 Aims and Hypotheses

While it has been suggested that these physiological changes may enhance sensitive caregiving behaviour (e.g., Joosen et al., 2012), the mechanisms of this change at a behavioural level have not been investigated. It might be expected that a state of physiological arousal would have a general ‘priming’ effect on behavioural responses, not restricted to caregiving behaviours. In the current study, a novel behavioural task was used to directly test this hypothesis. Participants’ performance on a motor task, requiring rapid, coordinated movements, was compared after listening to infant cries and control sounds. In addition, a group of adults with depression were tested to investigate whether depression might interfere with the physical ability to make prompt behavioural responses.

The behavioural task used was selected to isolate the physical demands of responsiveness required to initiate and sustain responsive behaviour in a standardised laboratory task. The key component of caregiving behaviour of interest in this study was rapid, coordinated, effortful motor movements. Rather than test immediate reactions to an individual infant cry, which might be highly dependent on physical properties of the stimulus and individuals' reaction times, this study was designed to investigate the sustained physical state after exposure to infant cries.

It was hypothesised that exposure to the sound of infant cries would promote rapid, effortful motor movements in adult listeners. To assess the specificity of infant cries for inducing this effect, motor performance was compared after listening to infant cries, adult cries and bird sounds. Adult sounds were included as a comparative category of emotional vocalisations from conspecifics that do not induce the same caregiving responses. Bird sounds were included as a 'baseline' condition of sounds in a similar frequency range to infant vocalisations, but which convey no particular emotion.

The performance of a group of adults with major depressive disorder was also compared to the performance of healthy adults on this task. It was hypothesised that two features of depression would be most likely to interfere with motor performance. First, disrupted emotional processing might lead to a lack of differential motor performance after listening to highly emotional infant cries, compared to other stimuli. Studies suggest that in depression, there is a negative bias in the interpretation of neutral or negative facial expressions (Gotlib et al., 2004; Gur et al., 1992; Leppänen et al., 2004) and slower reaction times when identifying sad facial expressions (e.g., Gollan et al., 2008). Responses to auditory emotional cues have been studied to a lesser extent, but some findings indicate that processing of vocal and musical stimuli is similarly disrupted (Emerson et al., 1999; Naranjo et al., 2011; Uekermann, Abdel-Hamid, Lehmkämper, Vollmoeller, & Daum, 2008). Altered emotional processing on this task would likely result in similar performance after exposure to infant cries, adult cries and bird sounds.

Secondly, ‘psychomotor disturbance’, defined as a general slowing of motor activity and a specific difficulty in responding quickly to the environment (Taylor et al., 2006) features within the diagnostic criteria for major depressive disorder (DSM-IV). Studies of motor skills have reported that adults with depression have slower reaction times and slower, less effortful movements compared to healthy adults (Cornell, Suarez, & Berent, 1984; Lawrie, MacHale, Cavanagh, O'Carroll, & Goodwin, 2000; Sabbe, Hulstijn, van Hoof, Tuynman-Qua, & Zitman, 1999; White, Myerson, & Hale, 1997). In the current study, psychomotor disturbance would be reflected in an overall reduction in motor performance in adults with depression, compared to healthy adults. Crucially, it would be expected that enhanced performance as a result of listening to infant cries would be retained, albeit at a relatively reduced level. Finally, if both altered processing of emotional stimuli and psychomotor disturbance affect the motor performance of adults with depression after listening to infant cries, it would be expected that performance would be both reduced relative to healthy adults, and exhibit no differentiation by stimulus type.

A secondary aim was to investigate the relationship between physiological reactivity to infant cry vocalisations and behavioural performance on an effortful motor task. As the most commonly-used measure of physiological reactions to infant cries, heart rate was recorded during task performance. It was hypothesised that enhanced motor performance following exposure to infant cry vocalisations would also be related to relative increases in participants’ heart rate.

3.2 Methods

3.2.1 Participant Demographics

Forty healthy adults (20 males, 20 females) and twenty adults with current major depressive disorder (7 males, 13 females) participated in this study. Adults with depression had significantly higher levels of depressive and anxious symptoms than healthy adults, as measured using the Edinburgh Postnatal Depression Scale (EPDS) and Generalised Anxiety Disorder Questionnaire (GAD-Q-IV; described below, section 3.2.2). Participants were aged between 18 and 45 years (*M* = 26.82, *SD* = 7.91) with no significant differences in age between groups (age of healthy participants, *M* = 26.50, *SD* = 8.21; age of participants with depression, *M* = 27.66, *SD* = 7.42). Participants reported no hearing impairments and were not taking any medications affecting the brain. At the time of the study, six of the participants (three healthy adults, three adults with depression) were parents, but their children were all aged over 18 months. Table 3 shows the demographic details of participants.

Table 3. *Demographic characteristics of participants with and without depression*

Demographic measure	Healthy participants	Participants with depression
n (*n* male)	40, 20	20, 7
Age in years (*M, SD*)	26.50, 8.21	27.55, 7.42
EPDS score (*M, SD*)	4.28, 3.24	18.85, 2.56*
GAD-Q score (*M, SD*)	2.30, 2.24	9.51, 1.21*

Note. EPDS = Edinburgh Postnatal Depression Scale, GAD-Q-IV = Generalised Anxiety Disorder Questionnaire, * denotes significant difference between groups (*p* < .001) using independent samples *t*-tests.

Ethical Approval and Recruitment

Ethical approval for this study was granted by the Oxfordshire Research Ethics Committee B (12/07/2010). Participants were recruited through internet advertisements on a local information website with a specialised section devoted to recruitment of volunteers (www.dailyinfo.co.uk) and posters in University departments and colleges. Participants with depression were recruited using targeted adverts (e.g., “Feeling down? Participate in research study?”). Individuals who responded to adverts were sent full information about the nature of the study and procedures involved.

At this stage, individuals were also screened for age, hearing impairments and medication status. Healthy participants fulfilling these criteria were then invited to attend an experimental session. Participants responding to adverts relating to depression were additionally informed: “We are looking for people currently suffering from depression or that have depressive symptoms (e.g., feeling down, tired, not enjoying life anymore, not sleeping well, not eating food as normal). If you decide to come in, we would like to go through an interview to discuss how you have been feeling before we go on to do the tasks.” Participants reporting these types of symptoms were then invited to attend an experimental session

Participation was voluntary and written informed consent was obtained prior to participation. Participants were informed that they were free to leave the study at any point without giving a reason and were debriefed at the end of the experimental session. Those who took part in a clinical interview were offered information about local mental health support services.

3.2.2 Assessment of Psychological Symptoms

Symptoms of depression were initially assessed using the EPDS. Given the known high rates of comorbidity of anxiety and depression (67% in a recent large cohort study using clinical interviews; Lamers et al., 2011), all participants were also assessed for current anxiety symptoms using the GAD-Q-IV. These self-report scales were selected to provide quantifiable measures of depressive and anxious symptoms respectively. The 10-item EPDS was originally developed as a community screening tool for postnatal depression (Cox, Holden, & Sagovsky, 1987). It has since been validated for use with women outside the postnatal period (Cox, Chapman, Murray, & Jones, 1996) and fathers (Matthey, Barnett, Kavanagh, & Howie, 2001). It has also been used in a number of studies with men outside the postnatal period (e.g., Buist, Morse, & Durkin, 2003; Deater-Deckard, Pickering, Dunn, & Golding, 1998; Ramchandani et al., 2008). The EPDS asks participants to respond to statements related to a range of depressive symptoms including feelings of unhappiness, enjoyment of normal activities, guilt, sleep disruption and ability to cope. Participants respond by ticking the appropriate level of a four-point Likert scale (from 'No, not at all' to 'Yes, most of the time'). Responses are then numerically coded to reflect severity of symptoms (i.e. '0' for no symptoms, '3' for the highest level of symptoms). The maximum score on the EPDS is 30, and 13 or above is a cut-off with high specificity and sensitivity indicating the likelihood of high levels of depressive symptoms (Cox et al., 1996).

The GAD-Q-IV is a self-report measure with high sensitivity and specificity for detecting symptoms of generalised anxiety disorder (GAD), based on the DSM-IV criteria (Newman et al., 2002). It is the most commonly used self-report measure of GAD. The GAD-Q-IV consists of 9-items that assess symptoms of anxiety including five 'yes/no' questions about current symptoms, a checklist of somatic symptoms and two 8-point Likert scales to indicate the severity of disruption to daily activities and distress. The maximum score on the GAD-Q-IV is 13, and a cut-off of 5.7 is recommended to optimise the sensitivity and specificity of this tool (Sauter et al., 2010). However, it is acknowledged that there is a likelihood of including false-positives with this cut-off and a more stringent cut-off of 9 is suggested to ensure that all diagnostic criteria are endorsed (Sauter et al., 2010). Within the sample of participants with depression included in the current study, 100% scored above the low cut-off of 5.7 on the GAD-Q-IV while 65% scored above the more stringent cut-off of 9.

To confirm diagnosis of major depressive disorder, participants scoring 13 or above on the EPDS were invited to take part in a semi-structured clinical interview (Structured Clinical Interview for DSM-IV Axis 1 [SCID]; First, Spitzer, Gibbon, & Williams, 2002). Interviews were conducted by a trained researcher and additional ratings of responses were completed by a second trained researcher. Only participants with a primary diagnosis of depression were included in the study. Participants who received a primary diagnosis of generalised anxiety disorder, or who failed to reach diagnostic criteria of current major depression were excluded from the study.

3.2.3 Stimuli: Infant Cries, Adult Cries and Bird Sounds

Infant cry stimuli came from the standardised database of affective sounds (the ‘OxVoc’ database) described in the previous chapter. In brief, these sounds were obtained from videotaped interactions of infants and caregivers recorded in their own homes during a play and feeding session (approved by Oxford Research Ethics Committee). Infants were all aged between six and eight months (*M* = 6.70 months, *SD* = 0.91) at the time of recording and were healthy and full-term at birth. Parents provided consent for these sounds to be used in research studies. Stimuli were all 1500ms in duration, were free from background noise and were matched for root mean square intensity. Fifteen exemplars of infant cries were selected on the basis of independent ratings completed by 61 adults (as described in Chapter 2, section 2.3).

The sounds selected were those rated as most distressed (on a Likert scale from 1, not distressed, to 7, very distressed: *M* = 5.31, *SD* = 1.52), with most negative mood (on a scale from 1, most negative, to 9, most positive: *M* = 2.56, *SD* = 1.30) and most negative participant mood after hearing the sounds (on a scale from 1, most negative, to 9, most positive, *M* = 2.92, *SD* = 1.49). Adult cry stimuli also came from the OxVoc database. Adult cries were used to compare effects from another stimulus also communicating human distress. These sounds were obtained from videoblogs posted by adults online. All adults provided consent for non-identifiable vocal clips to be used for research. Bird sound stimuli were obtained from a free online database of sounds (www.freesound.org). Bird sounds were used to compare motor performance after exposure to sounds with a comparable frequency range to infant cries, but that carried no emotional distress. Fifteen clips were selected that could be easily cropped to last 1500ms and demonstrated a range of bird vocal noises.

Presentation of sound stimuli during the main experiment and the hearing threshold test (see below, section 3.2.4) were performed using Presentation® software (version 14.4, Neurobehavioral Systems Inc., www.neurobs.com) on a laptop computer. The computer contained a 24-bit Realtek High Definition Audio sound card and sounds were played through Sony in-ear earphones (MDR-EX77LP).

3.2.4 Experimental Procedure

In a within-subjects design, participants completed three rounds of a brief, effortful motor task after listening to infant cries, adult cries and bird sounds (see Figure 7). Recordings were made of motor performance (score on the game), effort (minimum, maximum and mean amount of force applied over the course of a game) and heart rate (both mean heart rate and peak heart rate during each game).

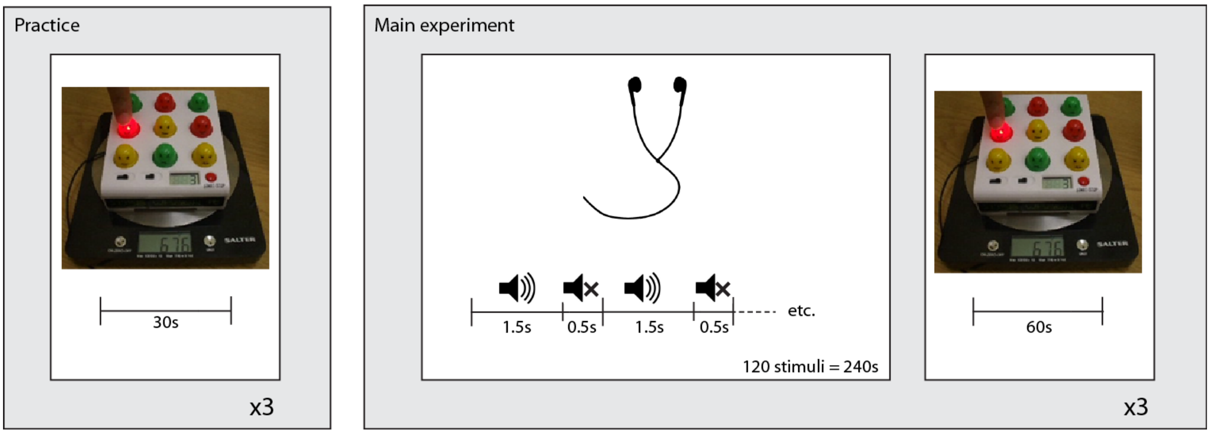

Figure 7. Experimental procedure for the ‘Whack-a-mole’ game. Participants first completed three 30-second practice rounds in order to become familiarised with the task. During the main experiment, participants listened to four minutes of sound clips then played a full 60-second game. This was repeated three times, listening to different categories of sound stimuli each time.

Assessing Thresholds of Hearing

Before beginning the task, each participant’s hearing threshold was established in order to ensure participants heard sounds at a comparable ‘loudness’. Loudness is the perceptual experience of the intensity of a sound and it is important to match on this dimension rather than physical intensity to account for differences in sensitivity to sound intensity between participants. Each individual’s threshold of hearing was identified using a computerised ‘staircase’ method, based on a 2-alternative forced choice, two-down, one-up design (described in Chapter 2, section 2.2.3). Briefly, this method presents participants with sounds of decreasing intensity, each time asking the participant to indicate whether they could hear the sound or not. After each trial when participants indicated they could hear a sound, the intensity on the subsequent trial decreased (by 10dB). After the first time the participant indicated they could not hear a sound, the intensity of the next sound increased (by 5dB). This is known as the first ‘reversal’ point. When participants next indicated that they could hear a sound, the successive stimuli became quieter again. This procedure continued until the participant made two ‘down-up reversals’ (where they go from ‘being able’ to hear a sound, to ‘not being able’ to hear a sound) at the same intensity value. This value was taken as the participants’ hearing threshold and stimuli during the main task were presented at 70dB above this level.

The ‘Whack-a-mole’ Game

Participants were first familiarised with the effortful motor task. This task was a miniature version of the classic arcade game ‘Whack-a-mole’ (‘Whack-it’, USB version). During the game, nine buttons arranged in a 3x3 grid illuminated individually in a random order. Participants pressed the target button (indicated by the currently illuminated light) in order to score points. A successful press had to occur within a specified amount of time (the duration of illumination) and required minimum amounts of force (equivalent to approx. 300g in weight) in order to score points. Over the course of a game, the duration of illumination decreased, requiring successively quicker responses. Participants were instructed to respond using the index finger of their dominant hand. They were familiarised with the task by completing three practice rounds, each lasting 30 seconds, before completing three full games, each lasting one minute.

Each of the full games was preceded by a ‘listening phase’ during which participants listened to sound stimuli with their eyes closed (see Figure 7). Each listening phase lasted for four minutes and consisted of a series of 1500ms sound stimuli separated by 500ms of silence. Participants heard 120 stimuli during each listening phase, consisting of eight repetitions of 15 exemplars of stimuli from the same category (e.g. infant cries), presented in a random order. The order of presentation of the sounds categories was randomised between participants.

Data consisted of scores on each game (including practice games) and a measure of force applied during the game. The measure of force applied was obtained by attaching the game to a set of digital scales (Salter 1036 BKDR, calibrated by manufacturer). A digital camera was used to video record performance during the task and was later reviewed to obtain the pressure applied during each button press. This data was then used to calculate the mean pressure applied by each participant during each game, and to identify minimum and maximum pressure values over the course of a game. During the practice rounds and main trials, average and peak heart rate was also recorded using a heart rate monitor with a chest strap transmitter (Polar T31 Coded^{TM}) and digital wrist unit (Polar FT1, Polar Electro (UK) Ltd., www.polaruk.co.uk). Peak and mean heart rate responses were recorded at the end of each game, encompassing the period during stimulus presentation and the duration of the game itself.

3.2.5 Preliminary Analysis Procedures

Data consisted of participants’ scores on the game, pressure data reflecting the amount of force applied over the course of the game (minimum, maximum and mean pressure) and heart rate data (mean heart rate, peak heart rate). This data was collected for each of the three experimental conditions as well as three practice rounds.

Performance on the 'Whack-a-mole' Game

Participants' mean scores during the main experiment were normally distributed (*D*(60) = .08, *p* = .20). Mean scores from practice rounds were not normally distributed (assessed using Kolmogorov-Smirnoff [K-S] tests), so log transformations were performed. Following log transformation, average scores from practice rounds were still not normally distributed. Data from practice games were analysed using non-parametric statistical tests.

Measures of Pressure Applied During the Game

Data were converted to standardised scores. Minimum, maximum and mean pressure scores during the practice rounds were not normally distributed. After log transformation, minimum and mean pressure scores were normally distributed (minimum: *D*(52) = .08, *p* = .20, mean: *D*(52) = .07, *p* = .20). Following removal of outliers falling outwith two standard deviations of the mean (M+/-2SD), log transformed maximum pressure scores were normally distributed (*D*(52) = .08, *p* = .20). Analyses described below were repeated with and without inclusion of outliers and results reported include outliers, unless otherwise stated. Standardised scores from the main experiment were not normally distributed so scores were converted to positive numbers and a log-transformation was performed. This resulted in normally distributed data as confirmed by K-S tests (minimum pressure: *D*(175) = .05, *p* = .20; maximum pressure: *D*(175) = .06, *p* = .20; mean pressure: *D*(175) = .07, *p* = .07).

Factor analysis was performed to examine whether the three measures of pressure score (minimum, maximum and mean) reflected a single underlying variable. Analyses were performed separately for the practice rounds and main experiment. For the practice rounds, correlations between variables were found to all be >.45, the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was 0.64 (above the recommended cut-off of 0.50; Field, 2005) and Bartlett's test of sphericity was significant ($\chi^2(3) = 68.41$, $p < .001$). A principal components analysis was then conducted and showed that the first factor explained 74.25% of the variance, with the second factor explaining 19.33% and the third explaining 6.42%. Only the first factor had an eigenvalue greater than one and so this factor was computed to provide a composite 'pressure' score for the practice rounds. Composite pressure data were normally distributed ($D(52) = .07$, $p = .20$).

A similar procedure was completed for data from the main experiment with correlations > .40, KMO = .59 and Bartlett's test was significant ($\chi^2(3) = 322.02$, $p < .001$). Principal components analysis showed three factors with the first explaining 75.29% of the variance and the second and third explaining 21.12% and 3.60% respectively. The first factor was computed to create a 'composite pressure' score for the main experiment. Data in the composite pressure variable were normally distributed ($D(175) = .06$, $p = .20$).

Measures of Heart Rate

From the practice rounds, heart rate measures were normally distributed after log transformation (mean heart rate: $D(59) = .07$, $p = .20$; peak heart rate: $D(60) = .07$, $p = .20$). During the main experiment, mean and peak heart rate measures were normally distributed without transformation (mean heart rate: $D(60) = .05$, $p = .20$; peak heart rate: $D(60) = .08$, $p = .20$).

3.3 Results

3.3.1 Performance on the 'Whack-a-mole' Game

Individual performance on the game was highly variable with scores on the practice rounds varying between 4 and 50 (*M* = 22.91, *SD* = 12.77) and scores on full games varying between 9 and 104 (*M* = 51.22, *SD* = 27.57).

Practice Games

Healthy participants' scores on the practice games (*Mdn* = 25.00) were higher than the scores of participants with depression (*Mdn* = 13.00). As this data was not normally distributed, an independent-samples Mann-Whitney U test was performed. This analysis confirmed that healthy participants' scores were significantly higher than scores of participants with depression (*U* = 215.50, *z* = -2.67, *p* = .008, *r* = -.35).

Full-length Games

Figure 8 presents participants' scores on the game, separated by participant group and stimulus category. This figure shows that healthy participants' scores were higher than scores of participants with depression. It also shows that scores were highest after listening to infant cries and lowest after listening to bird sounds.

Participants' scores on the main task were normally distributed (see section 3.2.5). A repeated-measures mixed-design 2x3 ANOVA was carried out with participant group (healthy participants and participants with depression) as a between-subjects factor and stimulus category (infant cries, adult cries and bird sounds) as a within-subjects factor. This analysis showed a significant main effect of participant group

($F(1, 58) = 11.32$, $p = .001$, $r = .40$), a significant main effect of stimulus category ($F(2, 116) = 8.3$, $p < .001$, $r = .26$) and no significant interaction between participant group and stimulus category ($F(2, 116) = 1.15$, $p = .32$, $r = .11$). Mean scores of healthy participants ($M = 59.03$, $SD = 26.93$) were significantly higher than mean scores of participants with depression ($M = 35.60$, $SD = 22.05$). Across all participants, scores were highest after listening to infant cries ($M = 53.68$, $SD = 28.66$), and lowest after listening to bird sounds (adult cries: $M = 50.52$, $SD = 28.34$; bird sounds: $M = 49.47$, $SD = 27.13$). Post-hoc paired t-tests with an alpha level of .017 (Bonferroni correction for multiple comparisons) were used to test significant differences between pairs of stimulus conditions. Results demonstrated significant differences between scores after listening to infant cries and scores after listening to adult cries ($t(59) = -2.56$, $p = .01$, $r = .32$) and between scores after listening to infant cries and scores after listening to bird sounds ($t(59) = 4.08$, $p < .001$, $r = .47$). There was no significant difference between scores after listening to adults cries and scores after listening to bird sounds ($t(59) = .88$, $p = .39$, $r = .11$).

Further investigation of the link between depressive symptoms and performance on the game was carried out using bivariate correlation analyses. The link between anxious symptoms and game performance was also assessed (although all participants had comorbid depression and anxiety with a primary diagnosis of depression). There was a significant correlation between scores on the practice games and EPDS scores ($r = -.34$; $n = 60$; $p < .01$) with higher levels of depressive symptoms related to lower scores. There was no significant correlation between practice scores and GAD-Q-IV scores (although this did approach significance, $r = -.24$; $n = 60$; $p = .06$). With data from the main experiment, there were significant correlations between scores on the game and both

EPDS scores (*r* = -.39, *n* = 60, *p* = .002) and GAD-Q-IV scores (*r* = -.32, *n* = 60, *p* = .02). Participants with higher depressive and anxious symptoms had lower scores on the game.

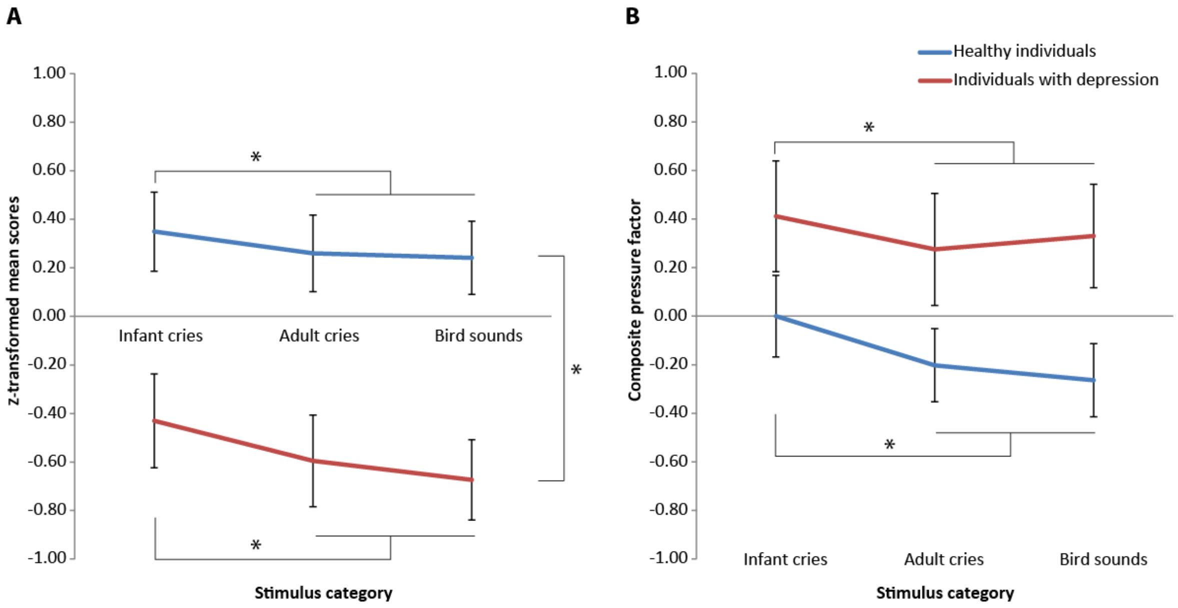

Figure 8. Effects of stimulus category and participant group on game performance and amount of pressure applied. A) Scores on the 'Whack-a-mole' game were significantly higher after listening to infant cries compared to listening to adult cries or bird sounds. Healthy participants' scores were significantly higher than scores of participants with depression. B) Composite pressure scores were significantly higher after listening to infant cries compared to listening to adult cries or bird sounds. There were no significant differences in the amount of pressure applied by the two groups. * denotes significant differences (Bonferroni corrected α = .017), error bars represent mean +/- standard error.

3.3.2 Measures of Pressure Applied During Game

Table 4 presents minimum, maximum and mean pressure scores for groups of healthy participants and participants with depression during the practice games and during the games of the main experiment. All statistical analyses were carried out using the composite pressure factor described above, data from which was normally distributed for both the practice games and the full-length games.

Practice Games

During the practice rounds, there were no significant differences in the amount of pressure applied by healthy participants and participants with depression ($t(50) = -1.18$, $p = .25$, $r = .16$).

Full-length Games

A repeated-measures mixed-design 2x3 ANOVA demonstrated a significant main effect of stimulus category ($F(2, 108) = 3.44$, $p = .04$, $r = .18$) on composite pressure scores. There was no significant main effect of participant group ($F(1, 54) = 2.35$, $p = .13$, $r = .30$), and no significant interaction of participant group and stimulus category ($F(2, 108) = .55$, $p = .57$, $r = .07$, see Figure 8). Post-hoc paired t-tests (with a Bonferroni corrected alpha level of .017) showed that pressure applied was significantly higher after listening to infant cries ($M = 172.66$, $SD = 71.66$) compared to after listening to adult cries ($M = 123.81$, $SD = 59.00$; $t(57) = 2.41$, $p = .02$, $r = .30$) or bird sounds ($M = 128.39$, $SD = 69.93$; $t(56) = 2.07$, $p = .04$, $r = .27$). There was no significant difference in pressure applied after listening to adult cries or bird sounds ($t(55) = .07$, $p = .95$, $r = .01$). Across both participant groups, responses were more forceful after listening to infant cries than other sounds.

Table 4. *Pressure measures of participants with and without depression during the 'Whack-a-mole' game*

Pressure Scores	Healthy participants	Participants with depression
	Practice Games	
Minimum (*M*, *SD*)	132.83, 68.68	214.28, 92.74*
Maximum (*M*, *SD*)	1061.84, 411.38	862.07, 138.69*
Mean (*M*, *SD*)	581.31, 242.08	451.14, 99.11*
	Full-length Games	
Minimum (*M*, *SD*)	132.14, 51.78	179.70, 78.83*
Maximum (*M*, *SD*)	833.74, 224.31	1004.18, 334.64*
Mean (*M*, *SD*)	455.29, 161.11	534.07, 191.24

Note. * indicates significant difference between groups ($p < .05$).

3.3.3 Measures of Heart Rate Reactivity

Practice Games

Mean heart rate was similar in healthy participants and participants with depression during the practice games (healthy participants, $M = 76.75$, $SD = 16.65$; participants with depression, $M = 80.40$, $SD = 10.73$). An independent t-test of log-transformed data demonstrated no significant difference between groups ($t(57) = -.589$, $p = .56$, $r = .08$). Similarly, there were no differences in peak heart rate of healthy participants ($M = 88.16$, $SD = 12.65$) and participants with depression during practice games ($M = 89.85$, $SD = 12.67$; $t(58) = -.49$, $p = .63$, $r = .06$).

Full-length Games

Mean heart rate

A 3x2 repeated-measures mixed-design ANOVA with stimulus category as a within-subjects factor and participant group as a between subjects factor was carried out. There were no significant main effects of stimulus category (*F*(2, 116) = 2.57, *p* = .08, *r* = .15) or participant group (*F*(1, 58) = .03, *p* = .86, *r* = .02) on mean heart rate. There was also no significant interaction effect of stimulus category and participant group (*F*(2, 116) = 1.72, *p* = .18, *r* = .12). Mean heart rate was similar after listening to infant cries (*M* = 76.37, *SD* = 10.59), adult cries (*M* = 76.32, *SD* = 9.96) and bird sounds (*M* = 74.87, *SD* = 14.18). Mean heart rate was also similar in healthy participants (*M* = 76.03, *SD* = 10.44) and participants with depression (*M* = 75.50, *SD* = 12.00).

Peak heart rate

Using the same analysis for peak heart rate measures, there was no significant main effect of stimulus category (*F*(2, 116) = 1.96, *p* = .15, *r* = .13), no significant main effect of participant group (*F*(1, 58) = .03, *p* = .86, *r* = .02), and no significant interaction effect (*F*(2, 116) = 1.08, *p* = .35, *r* = .10). Peak heart rate measures were similar after listing to infant cries (*M* = 90.67, *SD* = 12.43), adult cries (*M* = 89.27, *SD* = 12.19) and bird sounds (*M* = 88.42, *SD* = 18.46). Peak heart rate measures of healthy participants (*M* = 89.68, *SD* = 12.92) were similar to peak heart rate measures of participants with depression (*M* = 89.00, *SD* = 14.49).

3.3.4 The Effects of Gender on Performance Measures

To investigate the impact of gender on results, analyses from the main experiment described above were repeated with an additional between-subjects ‘gender’ variable. Repeated-measures mixed-design 3x2x2 ANOVAs were carried out with stimulus category (infant cries, adult cries and bird sounds) as a within subjects factor and participant group (with or without depression) and gender (male or female) as between subjects factors. There were no significant main effects of gender on scores on the game, any of the pressure measures or heart rate measures (all $p > .05$). There were also no significant interactions of gender with stimulus category, gender with participant group or gender with stimulus category and participant group on any of these measures (see Appendix B for statistics).

3.3.5 The Effects of Parenthood on Performance Measures

At the time of testing, six participants (three healthy adults and three adults with depression) were parents, although all children were aged over 18 months at the time of testing. Analyses of game scores, pressure measures and heart rate measures were repeated with data from the parents removed and no differences in the pattern of results was observed (see Appendix C for full details).

3.3.6 Summary of Results

Healthy participants had significantly higher scores on the game than participants with depression. This was apparent for both the main experiment and practice rounds. Across all participants, game scores were significantly higher after listening to infant cries compared to either adult cries or bird sounds. There were no significant differences in the pressure applied between healthy participants and participants with depression during either the practice games or the main experiment. In all participants, a significantly greater amount of pressure was applied after listening to infant cries compared to after listening to either adults cries or bird sounds. There were no significant differences in mean or peak heart rate between participant groups or after listening to different stimuli. There were also no significant associations between heart rate and amount of pressure applied. There was no significant main effect of gender, or interaction effect of gender and participant group, gender and stimulus category or gender, participant group and stimulus category on any of the experimental measures. Finally, removal of data from the six participants who had children at the time of testing had no effect on the pattern of results observed.

3.4. Discussion

This study demonstrated that effortful motor performance was significantly enhanced after listening to infant cries in adults with and without depression. Overall, adults with depression demonstrated impaired motor performance when compared with healthy adults, but retained reactivity to infant distress cries. These findings suggest that slowed responsiveness in depression is more related to features of psychomotor disturbance than impaired differential processing of emotional stimuli. Further, psychomotor disturbance in depression was limited to the ability to respond rapidly in a coordinated manner and did not affect the force of motor movements, suggesting a selective disruption to the speed and accuracy of motor movements in depression. Based on these results, it might be hypothesised that disruptions to the sensitivity of caregiving behaviour in depression would be affected by psychomotor disturbance.

Enhanced motor performance after exposure to infant distress cries was reflected in scores on an effortful motor game and also in measures of force applied while playing the game. Performance was better after listening to infant cries compared with performance after listening to adult cries or bird sounds. This suggests that infant cries are differentially effective in enhancing effortful motor performance. This is likely to be related to their function as highly alerting signals that promote the initiation of prompt caregiving responses (e.g., Del Vecchio et al., 2009; Schuetze & Zeskind, 2001). Other emotional human vocalisations (adult cries) do not produce the same effect and neither does 'emotionally neutral' birds sounds. More rapid, effortful and coordinated movements are clearly advantageous for the provision of prompt responsiveness to an infant in distress.

Motor performance of healthy adults was significantly greater than that of adults with depression (scores on the game were approximately 66% higher overall in healthy adults). In addition, higher levels of depressive symptoms were associated with poorer task performance. This is in line with the well-known symptom of psychomotor disturbance in depression, a slowing of motor movements and impairment in responding rapidly to environmental stimuli (e.g., Cornell et al., 1984). Impaired psychomotor performance was evident in the lower scores of adults with depression compared to healthy adults and was apparent even in the practice games before the main task. This is indicative of a pervasive psychomotor disturbance in adults with depression, such as that observed in other studies (e.g., Sabbe et al., 1999; White et al., 1997). In contrast, there was no difference in the amount of pressure applied by healthy adults and adults with depression during the game. This suggests that there are two dissociable components of effortful motor movement: the ability to make rapid, coordinated motor movements and the amount of effort or force applied within those movements. Only the former appears to be disrupted in this sample of adults with depression.

3.4.1 Limitations of Experimental Procedures

No differences were observed in heart rate after listening to infant cries compared to listening to adult cries or bird sounds. There were also no differences apparent in heart rate during the experiment between healthy adults and adults with depression. There are a number of explanations for these findings. First, as outlined in the introduction to this chapter (section 1.3) previous findings in heart rate reactivity to the sound of infant cries suggest that a complex interaction of individual characteristics can impact on heart rate measures (e.g., Bleichfeld & Moely, 1984; Out, Pieper, Bakermans-Kranenburg, & Van Ijzendoorn, 2010). These factors are not fully taken into account in the current study.

Secondly, heart rate reactivity is not the most sensitive measure of physiological arousal, as it is known to be affected by a multitude of other factors (e.g., Berntson et al., 1993). Inclusion of ‘purer’ measures of arousal, such as skin conductance might be more effective in the investigation of this issue.

Thirdly, the method of measuring heart rate in this study may have limited the potential to find differences in reactivity. The measure used was a simple heart rate monitor, which allowed a measure of maximum and mean heart rate over a given recording time. Assessing concurrent physiological changes was a secondary aim of this study and this method was chosen as a basic, easy-to-use measure that would minimise interference with task performance. Heart rate was recorded at the end of each ‘round’ of the game which comprised both a listening phase and a game-playing phase. It is likely that heart rate increased during game play due to the motor movements involved. Perhaps then, differences in heart rate during the listening phases were masked by the greater changes in heart rate during the game-playing phase. The design of the experiment did not allow separate recordings as it was deemed necessary for participants to move directly from the listening phase into the game-playing phase of the task without a pause in which to record heart rate data.

A different paradigm with continuous recording of heart rate data with millisecond resolution throughout the experiment would have allowed a more sensitive analysis to be performed. More sensitive assessment of changes in heart rate could also be achieved by using an explicit baseline period in which no sound stimuli were presented. In the present design, exposure to bird sounds was intended as a baseline condition, involving presentation of sound stimuli with no particular emotional content. It is plausible that presentation of these stimuli may have resulted in heart rate changes that masked the ability to detect further changes following exposure to infant vocalisations.

Analyses of differences in performance affected by gender or parental status found no significant effects. The current findings may be indicative of a rapid, universally induced state that promotes the ability to move quickly with greater effort and coordination. This study was not designed to specifically look at the effects of parenthood and the small number of parents included had children who were aged more than 18 months at the time of testing. It is possible that parents with young infants might show a different pattern of behaviour on this task.

3.4.2 Conclusions and Future Directions

In conclusion, exposure to the sound of distressed infant cries promotes rapid, effortful motor performance. This may be an adaptive mechanism facilitating sensitive caregiving behaviour in general and physical aspects of responsiveness to infant cues in particular. In adults with depression, impaired performance on this task seemed to be related to a general psychomotor disturbance affecting the ability to respond rapidly to environmental stimuli. Importantly, enhanced motor performance following exposure to infant cries was present in all adults, regardless of depressive symptoms. If these findings were to be extended to a caregiving context, it might be expected that disruption to sensitive caregiving in depression might be more related to general symptoms of depression rather than specific difficulties in responding to infant cues.

A central future aim would be to assess the link between this laboratory-based experimental measure of motor performance with observational measures of behavioural responsiveness to infant cues. Investigation of this link would provide a measure of the validity of this task as a strictly-controlled proxy of prompt responsiveness to infant vocal distress. Current findings suggest that general impairments in psychomotor capacities in depression might affect capacities to make rapid behavioural responses following exposure to infant cries. It may be that psychomotor disturbances in depression are related to reductions in sensitivity of caregiving behaviour, but as yet this possibility remains untested. The increasing use of virtual reality suites in psychological research and therapies (e.g., Powers & Emmelkamp, 2008) to create tightly controlled, ecologically valid experimental settings might be an avenue of interest for the further investigation of these behaviours.

In addition, future work aiming to investigate differences in psychomotor performance related to parental experience would be of much interest to establish whether this enhanced behavioural performance is altered by current experience with caring for young infants. It would also be important to assess aspects of physical responsiveness to infant cues in parents experiencing postnatal depression. This would allow further investigation of mechanisms that link postnatal depression with impaired sensitivity in parent-infant interactions.

Another future direction might be to assess further the specificity of these responses to infant cries. Current findings suggest that infant cry vocalisations enhance motor responding, relative to adult cry vocalisations or neutral bird sounds. Future work might investigate whether this response to infant cues is specific to humans, by assessing performance after listening to domestic animal distress vocalisations and whether it is specific to distress vocalisations by assessing performance after listening to infant laughter vocalisations. Finally, further investigation of the physiological state that might accompany this enhanced effortful motor performance would be of much interest to understand the physiological mechanisms by which exposure to an infant cry can lead to such a behavioural advantage. These investigations should include different physiological measures that are more specific to detecting arousal responses, such as skin conductance (e.g., Craske et al., 2008).

CHAPTER 4: INTERPRETING INFANT VOCALISATIONS - THE IMPACT OF DEPRESSION, MUSICAL EXPERIENCE AND PERCEPTUAL TRAINING

The developing infant relies on vocalisations to communicate with a caregiver at a distance, serving to promote proximity between the pair (Bell & Ainsworth, 1972; Bowlby, 1969, 1982). Chapter 3 demonstrated that infant cries promote rapid, coordinated motor responses in adults, indicative of heightened motivation to respond. Sensitive caregiving requires responses that are not only prompt, but also appropriate (Bell & Ainsworth, 1972). In order to provide appropriate care, an adult must accurately perceive and interpret infant cues. Previous work, including findings from Chapter 2, has demonstrated robust associations between the pitch of an infant cry and its perceived distress (Soltis, 2004; Zeskind & Marshall, 1988). This chapter describes three experiments investigating factors that affect adults' ability to interpret distress in infant vocalisations. Experiment 1 is a cross-sectional study investigating the roles of depression and previous musical training on sensitivity to pitch as a marker of distress. Experiments 2 and 3 assess the effectiveness of short training interventions aimed at improving sensitivity to infant vocal distress.

4.1 Introduction

As demonstrated and reviewed in Chapter 2 (section 2.1.5), physical parameters of infant vocalisations provide cues as to an infant's current physiological and affective state. While there are a number of parameters that contribute to the interpretation of infant vocalisations, robust associations have been demonstrated between the pitch of infant cries and perceived distress. Higher pitched cries are readily interpreted as sounding more distressed than lower pitched cries (Soltis, 2004; Zeskind & Marshall, 1988). The physiological mechanisms linking changes in affective state with changes in fundamental frequency (an acoustical measure related to pitch) of infant cries are relatively well-established (Soltis, 2004). In brief, autonomic arousal in emotive situations confers a range of physiological adaptations, such as changes in heart and breathing rates. Among these adaptations, muscle tension in the lungs and vocal tract is altered, resulting in changes in the pitch of vocalisations. In response to aversive stimuli, greater distress is associated with increased tension in the vocal chords, producing higher pitched vocalisations (Porges, 1995; Porter, Porges, & Marshall, 1988). This association may not be confined to aversive situations, laughter is also characterised by increased and variable pitch indicative of physiological arousal (Amoss, Martin, & Owren, 2011; Truong & van Leeuwen, 2007).

The association between infant physiological arousal and pitch of cry vocalisations has also been demonstrated in observational studies. Cries recorded during a surgical procedure that was increasingly invasive and (presumably) painful as it progressed were found to be increasingly high pitched, and independently rated as sounding more distressed (Porter et al., 1986). A similar association between pitch of infant cries and ratings of distress was demonstrated in Chapter 2 (section 2.4.2) from recordings made during a play and feed session with a caregiver. Other studies directly manipulating the pitch of infant cries also found a strong relationship between pitch and perceived distress (Dessureau et al., 1998; Protopapas & Lieberman, 1997). Together, these findings indicate robust associations between an infant's physiological arousal, the pitch of its cry and perceptions of distress.

4.1.1 Altered Interpretation of Infant Vocal Distress

Interpretation of infant vocal distress requires processing of both sensory and affective features of auditory stimuli. The mechanisms underlying appropriate interpretation of infant vocal cues can be assessed in individuals who demonstrate altered sensitivity to such stimuli. Responsiveness to infant vocalisations is known to be disrupted in postnatal depression, with depressed mothers less likely to initiate appropriate caregiving responses to their infant's cries (Bettes, 1988; Murray et al., 1993; Schuetze & Zeskind, 2001). It has been suggested that impairments in the appropriateness of maternal responsiveness to infant vocalisations are related to decreased sensitivity to basic perceptual features of these sounds, such as pitch (Donovan et al., 1998).

However, not all mothers with depression have difficulties interacting with their infants (Nylen et al., 2006). Identification of protective factors that account for the individuals who, despite being exposed to some of the same risk factors, continue to make healthy behavioural adaptations is of clinical and theoretical importance (Salekin & Lochman, 2008; Sohr-Preston & Scaramella, 2006). In the context of sensitivity to pitch in infant cries, one group of individuals who may demonstrate increased auditory sensitivity are those who are musically trained. As auditory stimuli, music and infant cries share a number of physical features and it is possible that individuals who have enhanced sensitivity to musical stimuli show a similar sensitivity to pitch in infant cries. Previous studies have shown that musicians are better than non-musicians at detecting small pitch changes in pure and complex sounds (e.g., Micheyl et al., 2006). In addition, a few recent studies suggest that musicians also have enhanced abilities in identifying emotion in speech prosody (Lima & Castro, 2011; Thompson et al., 2004). It is unknown whether the demonstrated advantages associated with musical training could extend to interpreting infant affective vocalisations. Furthermore, no studies to date have addressed whether musicians with affective disorders such as depression retain their auditory processing advantages.

4.1.2 Experiment 1: Aims and Hypotheses

The first experiment in this chapter aimed to investigate sensitivity to distress in infant cries in adults with and without depression and musical training. Sensitivity to distress was assessed using a two-alternative forced-choice (2-AFC) paradigm in which the pitch of infant cries was artificially manipulated. Pairs of stimuli differing in pitch were presented and participants were required to choose which ‘sounded more distressed’. Accuracy on this task was taken as a measure of sensitivity to infant vocal distress. It was hypothesised that adults with depression would show decreased sensitivity to distress relative to healthy adults and musicians would show enhanced sensitivity relative to non-musicians. Lowest sensitivity was predicted in adults with depression and no musical training.

4.1.3 Auditory Perceptual Training

Cross-sectional studies of musicians are limited in some regards because it is not possible to identify the extent of training required to improve perceptual capacities. Experiments 2 and 3 investigated whether it is possible to improve sensitivity to infant vocal distress through short, focussed training interventions. Plasticity in adults' perceptual systems has been the focus of a large body of research, demonstrating the capacity to improve fundamental auditory and visual perceptual abilities through short, focussed training sessions (e.g., Amitay, Irwin, & Moore, 2006; Karni & Sagi, 1993). Among other abilities, training has been demonstrated to improve the discrimination of pitch in sounds (Demany, 1985). While work in this field has focussed primarily on training of pitch with pure tones, some studies have also demonstrated effective perceptual learning with complex sounds (Miyazono, Glasberg, & Moore, 2010). One study demonstrated that just four to eight hours of focussed perceptual training in non-musicians resulted in improved pitch discrimination abilities to a level comparable to that observed in musicians (Micheyl et al., 2006).

Musicians, compared with non-musicians, have been shown to display a range of auditory advantages including the perception of pitch in pure tones, speech and emotional vocalisations (Lima & Castro, 2011; Magne et al., 2006; Tervaniemi et al., 2005; Thompson et al., 2004). In parallel, widespread changes in functional activity of auditory brain regions, including early stages of auditory processing in the brainstem, have been observed in musicians, compared with non-musicians (see Kraus & Chandrasekaran, 2010 for review). This broad range of perceptual advantages observed in musicians suggests that training occurs through modification of fundamental auditory capacities. While the effect of focussed perceptual training is likely to be more restrictive than that observed with musical training, it is possible that some of the neural plasticity associated with musicianship can be effected through short training interventions.

In more applied settings, both musical and perceptual training has been shown to impact on children's linguistic capacities. The impact of musical training in childhood was assessed in one longitudinal study, which demonstrated that children who were randomly allocated to one year of music lessons had improved abilities in detection of anger and fear prosody compared to their non-trained peers (Thompson et al., 2004). Short-term focussed training interventions targeting auditory discrimination abilities have been shown to have long-term benefits for wider perceptual capacities (Bradley & Bryant, 1983). In the domain of child language development, children with language learning impairments were shown to demonstrate improvements in speech discrimination and language comprehension after four weeks of phonemic contrast training (Tallal et al., 1996). Further studies have demonstrated comparable effects in typically developing children (Moore, Rosenberg, & Coleman, 2005), that were found to be maintained up to five to six weeks after the end of training. Perceptual training therefore appears to have broad, long-term impacts on auditory capacities.

4.1.4 Design of a Training Paradigm to Enhance Sensitivity to Infant Vocalisations

The optimal training parameters for enhancing perceptual learning remain an elusive field of study. Learning can be achieved through a variety of simple paradigms requiring participants to discriminate small differences between stimuli. Experimental design varies on the specific task involved as well as the amount of difference between stimuli (i.e. how easy or difficult the task is), the range of stimuli involved and the amount of feedback provided. Measures of successful training focus both on the change in ability to discriminate stimuli of interest as well as the 'generalisation' of this training, that is the extent to which improvements in perceptual skills extend to new stimuli and new tasks.

While the specific demands of a training task appear relatively unimportant for perceptual learning, the selection of appropriate stimuli can determine the efficacy of training. Similar rates of learning were observed using two-alternative forced choice ‘match-to-sample’ designs or three-alternative forced choice ‘oddball’ tasks (Amitay, Irwin, Hawkey, Cowan, & Moore, 2006). However, the choice of a ‘standard’ stimulus (the baseline stimulus, which other stimuli are compared to) in both of these task designs can impact the rate of learning. Designs employing identical standards on successive trials resulted in faster rates of learning, compared to when the standard varied across trials (Amitay, Hawkey, & Moore, 2005).

Participant motivation, influenced by task difficulty and level of feedback, has also been shown to impact on learning. While tasks that are too easy reduce the rate of learning, even impossibly difficult tasks can produce robust improvements in perceptual capacities (Amitay, Irwin, & Moore, 2006). Related to this, appropriate and accurate feedback can also enhance learning rates (Amitay, Halliday, Taylor, Sohoglu, & Moore, 2010). While the provision of a challenging task with accurate feedback can enhance the rate of learning, these features are not essential for improving perceptual skills as observed in one study demonstrating substantial improvements in pitch discrimination using a passive-listening design, with no task demands (Amitay, Irwin, & Moore, 2006).

Amount of Training Required to Effect Changes in Perceptual Performance

The amount of training required to observe significant changes in perceptual ability is not well-established. In general, there does appear to be a critical number of trials per training session below which there are no significant improvements (360 trials per session as reported in one study; Wright & Sabin, 2007). At the same time, it appears that training beyond a critical amount during a single session (or on a single day) may also confer no further training advantage (Roth, Kishon-Rabin, Hildesheimer, & Karni, 2005). There is much evidence demonstrating that performance accumulates and improves over multiple days (Delhommeau, Micheyl, Jouvent, & Collet, 2002; Demany, 1985; Irvine, Martin, Klimkeit, & Smith, 2000). Multi-session improvement is thought to be indicative of memory consolidation and transfer of short-term into longer-term memory (see Wright & Sabin, 2007). The most efficient training paradigms therefore might consist of short training sessions, repeated over multiple days.

Generalisation of Training Effects to Other Contexts

One primary goal of most training paradigms is to improve perceptual abilities not just to the stimuli used during training, but to a broader range of acoustical cues. The principles underlying generalisation of learning are not clear. Studies to date indicate that while learning on many tasks does generalise in part, complete generalisation to other tested conditions does not occur (Delhommeau et al., 2002; Miyazono et al., 2010). The majority of studies investigating generalisation have focussed on pure tone stimuli, although comparable performance was observed in a paradigm involving complex sound stimuli (Grimault, Micheyl, Carlyon, & Collet, 2002).

4.1.5 Experiments 2 and 3: Aims and Hypotheses

Experiments 2 and 3 aimed to investigate whether discrimination of pitch and sensitivity to distress in infant cries can be enhanced through short perceptual training interventions. Experiment 2 investigated the efficacy of a single training session on perceptual learning. Experiment 3 assessed the effects of two identical training sessions on consecutive days.

Participants were randomly allocated to auditory or visual training sessions, both consisting of a three-alternative forced choice (3-AFC) task. Participants in the auditory condition were required to discriminate infant vocalisations varying in pitch, while participants in the visual condition were required to discriminate infant faces varying in expression. Perceptual tests, involving discrimination of pitch and sensitivity to distress in infant cries were completed before and after training to assess the impact of the intervention.

It was hypothesised that adults participating in the auditory training intervention would demonstrate enhanced pitch discrimination thresholds after training, while participants in the visual training intervention would not. In addition, it was expected that this ability would generalise to increased sensitivity to infant vocal distress. Improvements in perceptual abilities were predicted to be greater after two training sessions, compared to after a single session.

4.2 Experiment 1: Investigating Sensitivity to Infant Vocal Distress

4.2.1 Methods: Experiment 1

Participant Demographics

Participants were 57 adults (28 male, 29 female) aged between 18 and 54 years (*M* = 27.16, *SD* = 8.31; see Table 5 for demographic details). All reported no hearing problems and were not currently taking medications affecting the brain. Current mood was assessed using the EPDS for symptoms of depression and the GAD-Q-IV for symptoms of anxiety (as described in Chapter 3, section 3.2.2). Healthy participants (*n* = 30; 18 male, 12 female) were those who scored below the established questionnaire cut-offs for depression and anxiety (see Table 5). Those scoring above questionnaire thresholds were invited to take part in a semi-structured interview (SCID) after which only those receiving a primary diagnosis of major depressive disorder were included in the study. Twenty-seven participants (10 male, 17 female) met criteria for current major depressive disorder.

In this study, being a musician was defined as having received four or more years of formal, extra-curricular musical training. Any participant with less than four years of musical training, or no training at all, was defined as a non-musician. This was a lenient definition of musicianship in comparison with other studies (e.g., 10 or more years; Musacchia, Sams, Skoe, & Kraus, 2007), but was chosen to reflect a common level of musical training in the general population. Of the 27 participants with depression, 13 were musicians. Of the 30 healthy participants, 15 were musicians.

Ethical Approval and Recruitment

Ethical approval for this study was granted by the Oxfordshire Research Ethics Committee B (12/07/2010). Participants were recruited through posters displayed in University Colleges and Departments and online advertising on local information websites with specialised sections devoted to recruitment of volunteers (www.dailyinfo.co.uk, www.gumtree.co.uk). Groups of participants with and without depression and musical training were recruited through targeted advertising. Participants with depression responded to adverts containing the phrase "Feeling down?" (e.g., "Feeling down? Participate in research study?"). Participants with musical training responded to adverts containing the word "Musicians" (e.g., "Musicians: participate in a sounds study?"). Adverts targeted at musicians with depression contained both of these phrases (e.g., "Musicians: feeling down? Participate in sounds study?"). Full information regarding the nature of the study and procedures involved was sent to all individuals who responded to adverts.

Individuals were screened at this stage for age, hearing impairments, medication status, and musical experience. The presence of depressive symptoms in individuals responding to depression-targeted adverts was confirmed at this stage by informing individuals that "We are looking for people currently suffering from depression or that have depressive symptoms (e.g., feeling down, tired, not enjoying life anymore, not sleeping well, not eating food as normal). If you decide to come in, we would like to go through an interview to discuss how you have been feeling before we go on to do the tasks" (identical to the procedure described above, section 3.2.1). Individuals who passed this initial screening stage were invited to attend an experimental session.

Written informed consent was obtained and individuals were informed that their participation was voluntary and that they could leave the study at any point. Debriefing at the end of the experimental session included the opportunity to ask any further questions and those who received a clinical interview were offered information about mental health support services in the local area.

Table 5. *Demographic characteristics of participant in different groups*

Demographic Measure	Healthy non-musicians	Healthy musicians	Non-musicians with depression	Musicians with depression
n (*n* male)	15, 9	15, 9	14, 5	13, 5
Age (*M, SD*)	31.00, 12.40	25.07, 3.26	27.79, 8.51	24.46, 4.45
EPDS (*M, SD*)	3.00, 2.51	3.27, 1.98	19.00, 2.83*	19.00, 2.58*
GAD-Q (*M, SD*)	1.29, 1.60	1.49, 1.67	9.52, 1.49*	9.11, 2.21*
Years music (*M, SD*)	.33, 0.90	8.87, 3.58*	0.00, 0.00	8.15, 3.53*

Note. ‘Years music’ = number of years of musical training, EPDS = Edinburgh Postnatal Depression Scale (cut off for depression > 13), GAD-Q = Generalised Anxiety Disorder Questionnaire (cut off for anxiety > 5.7). * denotes significant difference between groups ($p < .05$).

Infant Vocal Distress Stimuli

Fifteen digital recordings of infant cries from the OxVoc database (see Chapter 2) were digitally altered to increase and decrease overall pitch by 0.25, 0.5, 1 and 2 semitones (using the ‘Pitch Shift’ effect, Adobe® Audition® CS5.5, v4.0). The resulting stimuli consisted of 8 variants on each of the original 15 stimuli, 120 stimuli in total. These stimuli were subsequently presented in pairs differing in 0.5 semitones (+0.25 vs. -0.25 semitones), one semitone (+0.5 vs. -0.5 semitones), two semitones (+1 vs. -1 semitone) or four semitones (+2 vs. -2 semitones). The efficacy of this manipulation was confirmed using measures of fundamental frequency (see Chapter 2, section 2.4.1 for details of calculation of fundamental frequency). Table 6 displays the mean fundamental frequencies of the original and manipulated stimuli.

Table 6. *Mean fundamental frequencies of original and pitch-manipulated stimuli*

Pitch Manipulation (semitones)	Fundamental Frequency (Hz) *M (SD)*
Original pitch	475.88 (78.62)
Raised Pitch	
+ 2.00	532.59 (91.89)
+ 1.00	505.31 (82.85)
+ 0.50	489.52 (81.68)
+ 0.25	484.50 (79.95)
Lowered Pitch	
- 0.25	469.42 (76.22)
- 0.50	459.37 (74.22)
- 1.00	449.33 (75.09)
- 2.00	424.20 (68.56)

Note. Mean values across all stimuli (*n* = 15) are presented at each level.

'Sensitivity to Distress' Task

Participants performed a two-alternative forced choice task in which they were asked to choose which of each stimulus pair sounded 'more distressed'. Clips of the same cry at different pitches were presented in pairs that differed by 0.5, 1, 2 and 4 semitones. All 15 different cry stimuli were presented at each of the different pitch levels, resulting in 60 trials in total. Digitally altered stimuli were not presented with original vocalisations to avoid potential over-exposure to the original sound clips which could bias responding. The perceived 'loudness' of stimuli was matched across participants by presenting stimuli at 50dB above each individual's threshold of hearing, assessed using a staircase method as described in Chapter 2 (section 2.2.3).

Participants were instructed: “You are going to hear pairs of infant cries. Each cry is represented by a shape on the screen. After listening to both cries once, you then have to choose which you think sounded more distressed. Press the ‘up’ button if you think the first cry sounded more distressed, and the ‘down’ button if you think the second cry sounded more distressed.” Each trial began with a 700ms fixation cross. An image of a fractal then appeared in the upper half of the computer screen paired with the first of two infant cry stimuli. This was then replaced with an image of a fractal in the lower half of the computer screen, paired with the second infant cry stimulus. Images were on screen for 1600ms and audio stimuli (duration 1500ms) were presented 100ms after the onset of each image. Following this, both fractals appeared simultaneously on screen, cuing participants to respond. Participants responded by using the ‘UP’ or ‘DOWN’ arrows on the keyboard to select their chosen fractal, representing either the first or the second cry stimulus. Their selection was represented onscreen by the appearance of a white box around their chosen fractal image for 800ms (see Figure 9). The order of trials and whether the higher or lower pitched sound was presented first was randomised.

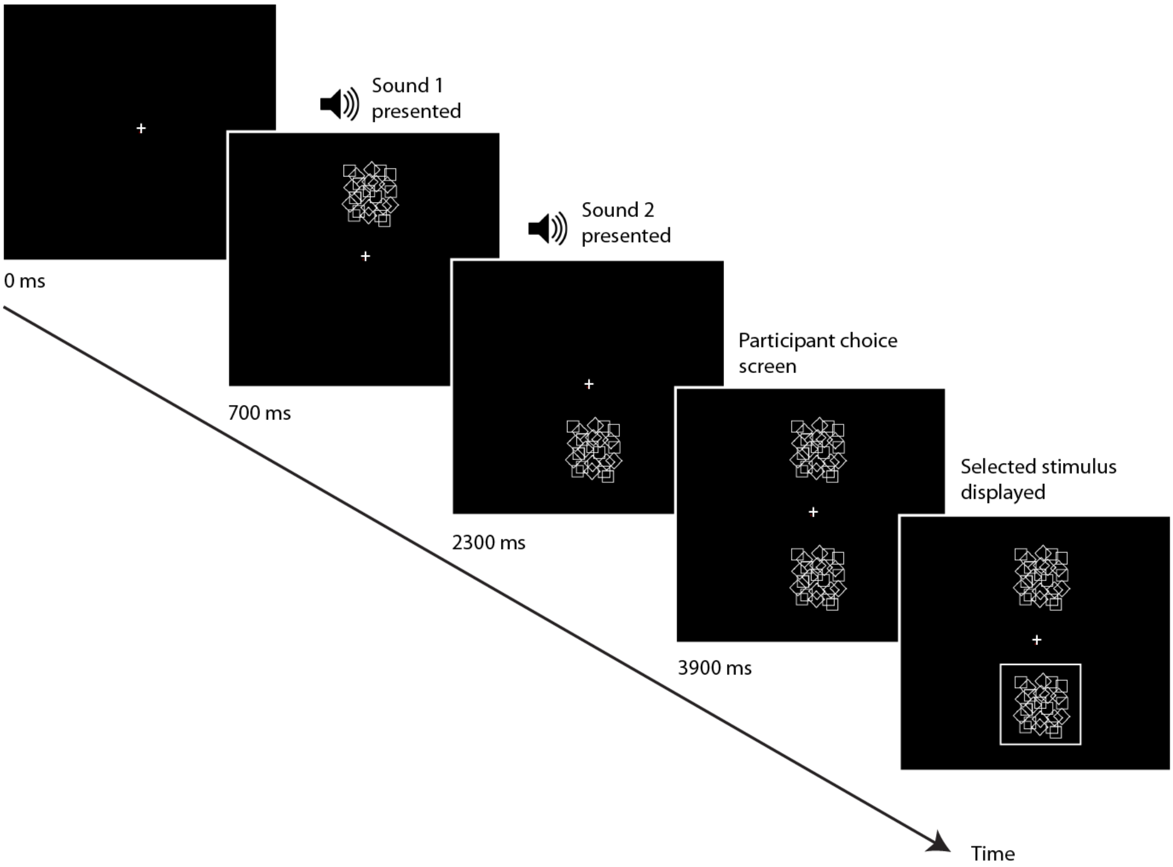

Figure 9. Overview of the 'sensitivity to distress task'. Each trial began with a fixation cross, followed by presentation of the first cry stimulus, paired with an image of a fractal in the upper half of the screen. Following this, the second stimulus was presented, paired with an image of a fractal in the lower half of the screen. Both fractals were then presented together, cuing participants to select the image representing the infant cry that 'sounded more distressed'. Finally, participants' chosen stimulus was indicated with a white box, before the next trial commenced.

Sound stimuli were presented through Sony in-ear earphones (MDREX77LP) from a laptop computer (24-bit Realtek High Definition Audio sound card). This experiment was programmed and administered using Presentation® software (version 14.4, Neurobehavioral Systems Inc., www.neurobs.com). Data recorded consisted of a measure of the proportion of 'correct' trials at each level of difficulty across the task. Correct trials were defined as trials in which participants selected the higher pitched cry as sounding more distressed. Accuracy in the selection of higher pitched cries as sounding more distressed was compared between participant groups.

4.2.2 Results: Experiment 1

In general, accuracy in the selection of higher pitched cries as sounding ‘more distressed’ was greater than chance (percentage of higher pitched sounds chosen, *M* = 65.25, *SD* = 9.84). Accuracy data on the sensitivity to distress task were normally distributed, as confirmed by a K-S test (*D*(57) = .08, *p* = .20). Associations between the presence of depression, previous musical experience and performance on the task were investigated using a mixed design repeated-measures ANOVA. Level of difficulty on the task was used as a within-subjects factor (4 levels) and participant group was used as a between-subjects factor. This analysis demonstrated a significant main effect of task level (*F*(3, 159) = 33.60, *p* < .001, *r* = .42). Performance was more accurate at easier levels of the task (i.e. the percentage of higher pitched sounds chosen was greater when differences in pitch between stimuli were also greater; see Figure 10). There was also a significant main effect of participant group (*F*(3, 53) = 4.45, *p* = .007, *r* = .28). There was no significant interaction of task level by participant group (*F*(9, 159) = 1.85, *p* = .06, *r* = .11). Post-hoc Bonferroni-corrected pairwise comparisons were used to investigate the nature of between-group differences. These analyses revealed significantly lower scores for non-musicians with depression (*M* = 56.59, *SD* = 7.87) compared either with musicians with depression (*M* = 68.33, *SD* = 12.59; *p* = .04, *r* = .27) or healthy musicians (*M* = 67.22, *SD* = 11.56; *p* = .02, *r* = .30 respectively). Healthy non-musicians’ scores (*M* = 62.50, *SD* = 7.42) were not significantly different from any other group.

Specific differences in performance at different levels of the task were also investigated using one-way ANOVAs. These analyses demonstrated significant differences in performance across participant groups at the two semitone (*F*(3, 38) = 3.92, *p* = .02, *r* = .31) and four semitone levels (*F*(3, 38) = 13.38, *p* < .01, *r* = .51), but not at the 0.5

semitones ($F(3, 38) = 1.15$, $p = .34$, $r = .17$) or one semitone levels ($F(3, 38) = .74$, $p = .54$, $r = .14$). Further post-hoc Bonferroni pairwise comparisons demonstrated that non-musicians with depression achieved significantly lower scores than either group of musicians at both the two and four semitone levels (all $p < 0.05$, see Figure 10).

The associations between musical training, depressive symptoms and task performance were investigated using Pearson correlation analyses. There was a significant positive correlation of overall score with years of musical training ($r = .55$, $p < .001$) and no significant relationship between overall score and EPDS score ($r = -.05$, $p = .71$). There were no differences in overall performance between men and women ($t(55) = .01$, $p = .99$, $r = .01$). Five participants in this study were parents. A sensitivity analysis was carried out removing these participants, and the same pattern of results was observed.

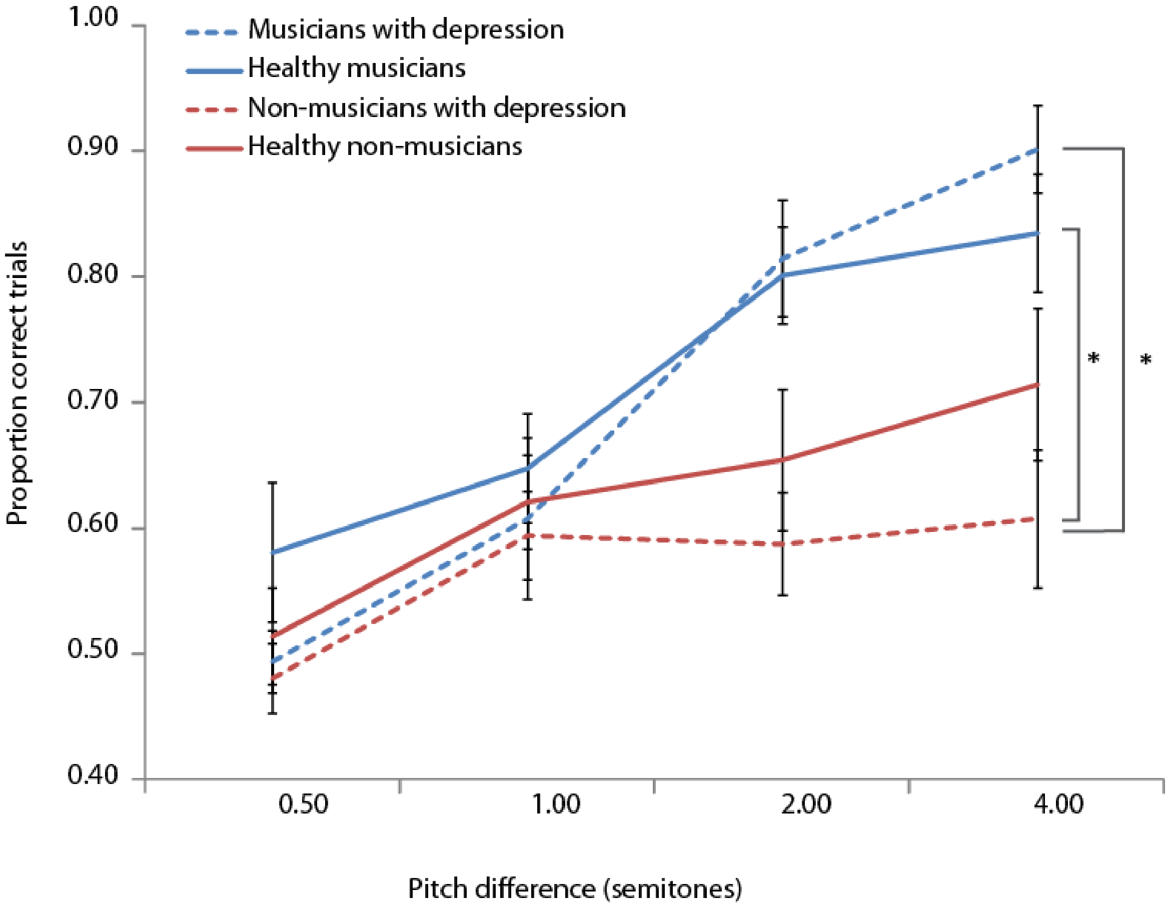

Figure 10. Mean sensitivity to distress in infant cries in groups of participants with and without depression and musical training. The proportion of higher pitched cries chosen as sounding more distressed (proportion of correct trials) for each group at each level of pitch difference is displayed. The scores of non-musicians with depression were significantly lower than both groups of musicians at two and four semitones. Error bars represent mean +/- standard error, *indicates significant difference (*p* < .001)

4.3 Experiments 2 and 3: Training to Improve Sensitivity to Infant Vocalisations

4.3.1 Methods: Experiments 2 and 3

Participant Demographics

Participants in Experiment 2 were 24 adults (11 male, 13 female), aged 20-39 years (*M* = 26.29, *SD* = 5.79). All participants reported no hearing problems and were not taking medication that affects the brain. As in Experiment 1, all participants completed EPDS and GAD-Q-IV screening questionnaires. Only participants scoring below the cut-off for depressive symptoms on the EPDS (*M* = 4.75, *SD* = 3.23) and anxiety symptoms on the GAD-Q-IV (*M* = 1.72, *SD* = 1.82) were invited to take part in the study. These 24 individuals were allocated to either the auditory training condition (*n* = 12, 3 male, 9 female) or the visual training condition (*n* = 11, 7 male, 4 female).

Participants in Experiment 3 were thirty-eight adults (15 male, 23 female), aged 18-49 years (*M* = 22.57, *SD* = 6.52). The same exclusion criteria were used as in Experiment 2. All participants in Experiment 3 scored below the cut-offs on measures of depression (EPDS; *M* = 3.61, *SD* = 2.18) and anxiety (GAD-Q-IV; *M* = 1.37, *SD* = 1.96). These 38 individuals were also allocated to either the auditory training condition (*n* = 19, 9 male, 10 female) or the visual training condition (*n* = 19, 6 male, 13 female).

Ethical Approval and Recruitment

Ethical permission for these studies was approved by the Medical Sciences Interdivisional Research Ethics Committee (IDREC, ref: MSD-IDREC-C1-2012-148). Participants were recruited through posters and mailing list adverts in University Colleges and Departments. Recruitment procedures were similar to those described for Experiment 1 (see section

4.2.1), but without aspects related specifically to the recruitment of individuals with depression or musical training.

Stimuli

Two infant cry stimuli from the OxVoc database were used in this study. Stimuli were cropped to 1000ms in length in order to shorten the total task duration. Each stimulus was manipulated to increase or decrease overall pitch (using the ‘Pitch Shift’ effect, Adobe® Audition® CS5.5, v4.0). Pitch manipulations ranged from 0.05 semitones to 2 semitones in 0.05 steps. This resulted in 80 new stimuli for each cry clip, differing from the original by up to +/-2 semitones.

For the visual training control condition, infant face stimuli varying in facial expression were used. These stimuli were developed using morphing software to create a range of stimuli that systematically changed from a neutral to happy facial expression or a neutral to sad expression. Each series was created using images of two expressions from the same infant, morphing at 2% intervals from neutral to full emotion (resulting in 100 stimuli per face). These stimuli have previously been used in studies assessing maternal sensitivity to emotional facial expressions in infants (e.g., Arteche et al., 2011).

Overview of Experimental Procedures

In Experiment 2, participants were randomly allocated to visual or auditory training conditions. At the beginning of the experimental procedure, all participants completed the ‘sensitivity to distress’ task described in Experiment 1 (referred to as ‘Distress task’ in Figure 11 below). All participants then completed a pitch discrimination threshold test (this is referred to as ‘Voice 1’, in Figure 11 below) for one infant cry stimulus (vocalisation clip 1). For participants allocated to the auditory training condition, this was the stimulus they subsequently received training on. Participants in the auditory training condition completed an additional pitch discrimination test (‘Voice 2’ in Figure 11) for a second infant cry stimulus (vocalisation clip 2), while participants in the visual training condition completed a test of facial expression discrimination (‘Face 1’). Training then consisted of 400 trials on a 3-AFC oddball task (of either pitch discrimination or facial expression discrimination), lasting approximately 40 minutes. After training, the threshold discrimination and sensitivity to distress tasks were completed again.

In Experiment 3, participants completed the same procedure described for Experiment 2 on two consecutive days, with one exception: the sensitivity to distress task was completed only before training session one and after training session two (see Figure 11). All tasks were computer-based and were programmed and run using Presentation® software.

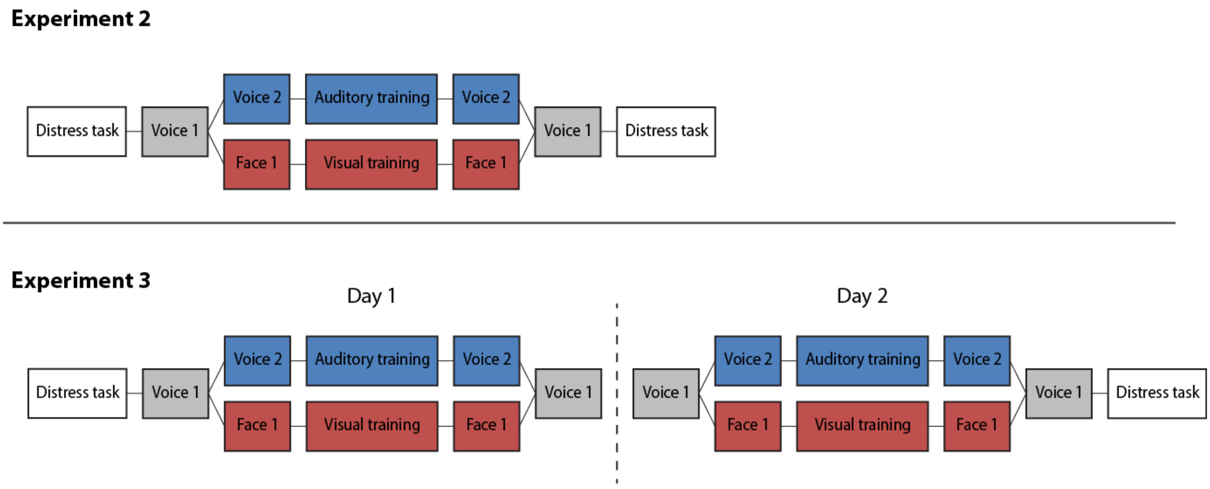

Figure 11. Overview of experimental procedure for perceptual training in Experiments 2 and 3. All participants completed the sensitivity to distress task at the beginning and end of the full training protocol, while discrimination testing for individual stimuli was completed before and after each training session. Differences in training protocol are highlighted for participants in the auditory training group (blue) and visual training group (red). 'Voice 1' refers to the pitch discrimination threshold test for vocalisation clip 1, 'Voice 2' refers to the same test for vocalisation clip 2 (participants in the auditory training group only), 'Face 1' refers to the facial expression discrimination test (participants in visual training group only).

Testing of Perceptual Discrimination Thresholds: Pitch in Infant Cries and Expression in Infant Faces

This task had a 3-AFC 'oddball' design, implemented using an adaptive staircase to rapidly establish participants' threshold of perceptual discrimination. Similar procedures were used both for assessing discrimination of pitch in infant cries and discrimination of expressions in infant faces. The procedure described below relates to testing of pitch discrimination thresholds, but the procedure for testing facial expression was identical, apart from the use of visual rather than auditory stimuli.

On each trial of the 3-AFC task, three stimuli were presented sequentially, one of which was slightly higher (or lower) in pitch than the other two (which were the unmodified cry stimulus). Participants were asked to select which stimulus was the 'odd-one-out'. Before completing this task for the first time, participants completed five practice trials to familiarise them with the procedural demands of the task. The use of the adaptive staircase meant that with every correct response, the pitch difference on the subsequent trial decreased, making the task more difficult. With every incorrect response, the pitch difference on the subsequent trial increased, making the task easier.

The adaptive staircase was implemented using three different levels. At the first level, the interval between trials varying in difficulty was large, meaning that with subsequent trials, the pitch difference between stimuli rapidly became smaller (in steps of 0.3 semitones with each correct response). When participants made their first incorrect response (the reversal point), the task progressed to level two and the difference in pitch increased on subsequent trials, but by a smaller interval than level one (0.15 semitones with each incorrect response). After participants made another correct response, the task progressed to level three, during which correct responses increased task difficulty by 0.10 semitones and incorrect responses decreased difficulty by 0.05 semitones. The task continued for a total of 40 trials, and the mean of the last four reversal points was taken as a measure of an individual's perceptual threshold (as used in Amitay, Irwin, Hawkey, et al., 2006).

Presentation of each auditory stimulus was paired with a letter onscreen located on the left, the middle or the right of the screen (for the first, second and third stimuli respectively). Each letter was on screen only for the duration of the auditory stimulus with which it was paired (1s). The three letters were then presented on screen simultaneously, prompting the participant to choose which was the 'odd one out'. The letters presented onscreen corresponded to the keys on the keyboard with which participants should respond (the 'z', 'v' and 'm' keys). Each trial ended when participants made a key-press response. After each trial, participants received feedback about whether their performance was correct or not, in the form of a happy or sad cartoon face (presented for 1s).

Training Procedure

Training consisted of a 3-AFC oddball task, similar to the method employed in the testing phase, but with a fixed difference in pitch/face morphing across all trials. The difference between stimuli was set at each individual's discrimination threshold. The timing and feedback of trials was identical to that described above. Each training session consisted of 400 trials, separated into blocks of 100 trials, with breaks between each block. Auditory and visual training procedures were identical, apart from the use of either auditory or visual stimuli. The training sessions were largely self-paced but lasted for approximately 40 minutes.

4.3.2 Results: Experiment 2

Effects of Training on Perceptual Discrimination Thresholds

Pitch discrimination threshold testing was completed before and after each training session. All participants completed tests for one infant cry stimulus (vocalisation clip 1), which was the stimulus used during training in the auditory group. Participants in the auditory training group were additionally tested on an unfamiliar cry stimulus (vocalisation clip 2). Participants in the visual training group completed a test of facial expression discrimination.

Data from all measures of threshold of pitch and facial expression discrimination were not normally distributed when tested with K-S tests (all $p < .05$). All measures were significantly positively skewed, which could not be corrected using transformations or outlier removal. This positive skew indicated a ‘ceiling effect’ of performance with participants’ scores clustering at the highest end of the measurable range of sensitivity. The majority (74.2%) of participants scored in the top 10% of the range of performance (discrimination of 0.2 semitones or less) before the first training session. This ceiling effect was not apparent during piloting of the task. Data from these tasks were therefore analysed using non-parametric tests.

Discrimination threshold of pitch in infant vocalisation clip 1 (all participants)

Wilcoxon signed rank tests were used to investigate changes in pitch discrimination threshold in the different training groups. On average, thresholds were very low, with participants reliably detecting pitch differences of as little as 0.1 semitones (*M* = 0.23 semitones, *SD* = 0.21). The median pitch discrimination threshold of vocalisation clip 1 was compared before and after one session of training for the auditory and visual training groups separately. There was a significant difference in pitch discrimination threshold in the auditory training group (*z* = -2.40, *p* = .02, *r* = -.69; see Figure 12, A). In this group, the median threshold was significantly lower after training (*Mdn* = 0.13 semitones, *IQR* = .03) compared to before training (*Mdn* = 0.16 semitones, *IQR* = .13). There was also a significant difference in pitch discrimination threshold in the visual training group (*z* = -2.44, *p* = .02, *r* = -.70), again with a significant decrease in median pitch discrimination threshold after training (*Mdn* = .13, *IQR* = .05), compared with before training (*Mdn* = .17, *IQR* = .13).

Discrimination threshold of pitch in vocalisation clip 2 (auditory training group)

The generalisation of improvements in auditory training was investigated in two ways: first by assessing pitch discrimination thresholds on a second auditory stimulus and secondly by assessing sensitivity to distress in a separate task (see below). As with thresholds for vocalisation clip 1, pitch discrimination thresholds for vocalisation clip 2 were low (M = .17 semitones, SD = .18). Wilcoxon signed rank tests demonstrated that there were no significant differences in pitch discrimination threshold for vocalisation clip 2 after a single training session (z = -.45, p = .66, r = -.13; see Figure 12, B).

Discrimination threshold of expression in infant faces (visual training group)

Accuracy scores on the facial discrimination task also showed that participants performed at ceiling. This task had 50 levels of difficulty and all participants' performance fell within the most difficult 10% of this range (M = 2.17, SD = .45). Wilcoxon signed rank tests found no significant difference in facial expression discrimination ability after training (Mdn = 2.32, IQR = .50), compared with before training (Mdn = 2.12, IQR = .64; z = .85, p = .40, r = .25).

Effects of Training on Sensitivity to Distress in Infant Cries

Generalisation of auditory training to other stimuli and other tasks was assessed by measuring participants' ability to detect differences in pitch and interpret these as indicating distress (the 'sensitivity to distress' task described in Experiment 1). Mean scores on this task were found to be normally distributed, as confirmed by K-S tests (all $p > .10$). The mean proportion of trials correct (proportion of higher pitched stimuli chosen as sounding more distressed) was above chance (M = .60, SD = .09). Paired-samples t-tests were performed to assess changes in performance on the ‘sensitivity to distress’ task in each of the training groups. There were no significant differences in performance after training, compared to before training, in either the auditory training group (*t*(11) = -1.45, *p* = .18, *r* = .30) or the visual training group (*t*(10) = -.84, *p* = .42, *r* = .26; see Figure 12C-D).

The Effect of Gender on Perceptual Abilities Before and After Training

Independent samples t-tests demonstrated no significant effects of gender on pitch discrimination thresholds before (*t*(22) = -.59, *p* = .56, *r* = .12) or after training (*t*(22) = -.70, *p* = .49, *r* = .15) on vocalisation clip 1. There were also no significant gender differences on performance for vocalisation clip 2 in the auditory training group before (*t*(10) = -.58, *p* = .57, *r* = .18) or after training (*t*(10) = -.38, *p* = .71, *r* = .12). Finally, there were no significant gender differences in performance on the distress task before (*t*(21) = -.69, *p* = .50, *r* = .15) or after training (*t*(21) = .37, *p* = .72, *r* = .08).

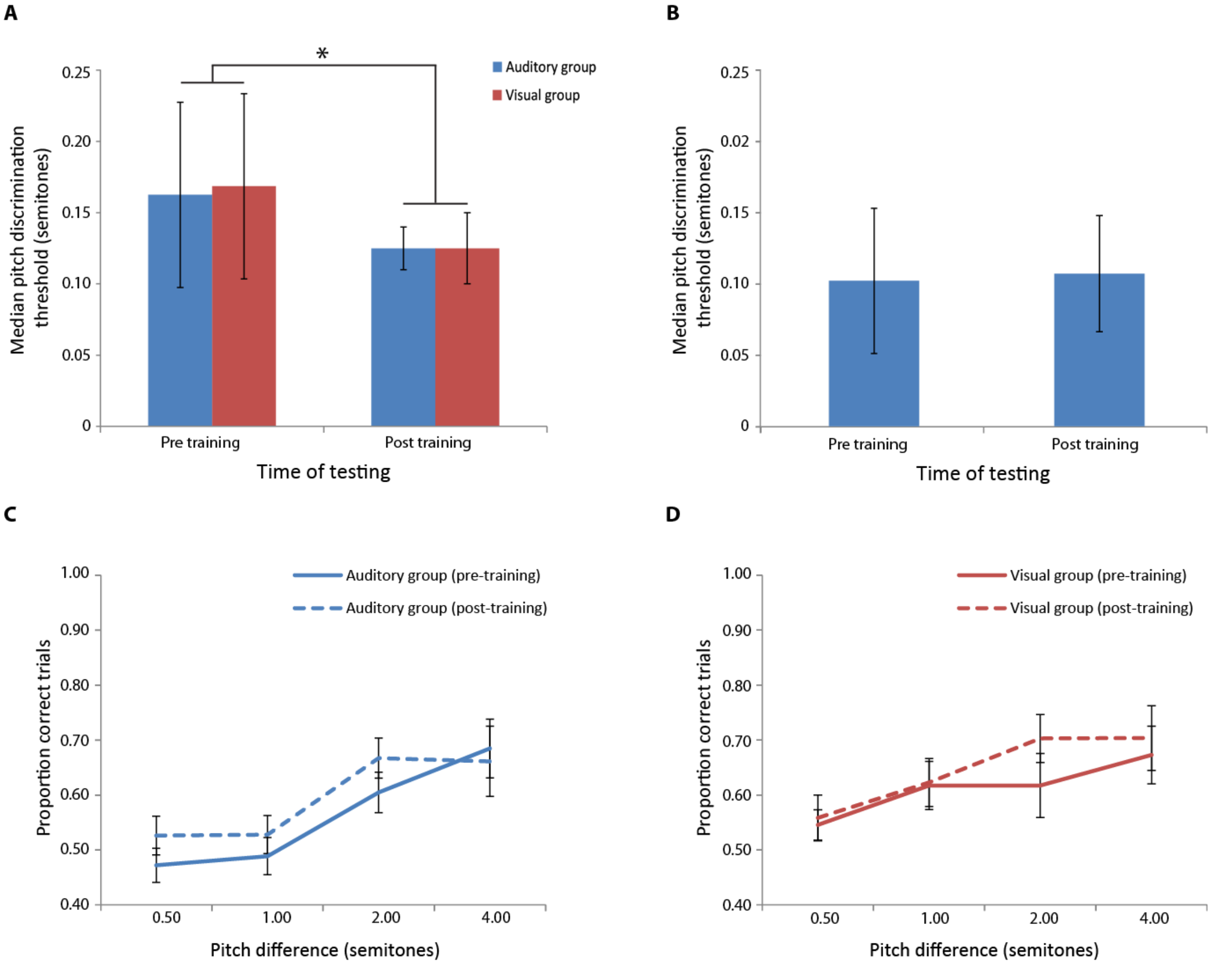

Figure 12. Performance on pitch discrimination and sensitivity to distress tasks before and after one training session (Experiment 2). A) Measures of pitch discrimination threshold for vocalisation clip 1, significant improvements were observed for both the auditory and visual training groups. B) There was no significant improvement in pitch discrimination on vocalisation clip 2 in the auditory training group. C-D) Performance on the ‘sensitivity to distress task’, showing the proportion of higher pitched stimuli chosen as sounding more distressed (proportion correct trials). There were no significant differences were observed after training in either the auditory (C) or visual (D) training groups. Error bars represent inter-quartile range (A-B) or mean +/- standard error (C-D). *denotes a significant difference ($p < .05$)

4.3.3 Results: Experiment 3

Effects of Training on Perceptual Discrimination Thresholds

As with Experiment 2, performance on pitch discrimination tasks was high, with evidence of a ceiling effect (78.9% scored within the top 10%, $M = .18$, $SD = .12$), necessitating the use of non-parametric tests. Wilcoxon signed rank tests were used to assess change in performance after one or two training sessions in the auditory and visual training groups.

Discrimination threshold of pitch in vocalisation clip 1 (all participants)

In the auditory training group, pitch discrimination thresholds were not significantly different after one training session (z = -1.43, p = .15, r = -.33), but did significantly improve after two training sessions (z = -2.22, p = .03, r = -.51). There was no significant main effect of training for session two alone (z = -1.74, p = .08, r = -.40) suggesting that the effects of training were cumulative over the two training sessions (see Figure 13, A). In the visual training group, there were no significant changes in pitch discrimination thresholds after either one (z = 1.05, p = .30, r = .24) or two training sessions (z = -.72, p = .47, r = -.16).

Discrimination threshold of pitch in vocalisation clip 2 (auditory training group)

Participants in the auditory group were additionally tested on their pitch discrimination threshold on a second, unfamiliar cry clip. Comparable to findings for vocalisation clip 1, pitch discrimination thresholds were not significantly different after one training session (z = -.54, p = .59, r = -.12), but did significantly improve after two training sessions (z = -2.77, p = .005, r = -.64). There was no significant effect of training session 2 alone (z = .39, p = .69, r = .09) indicative of a cumulative effect of training over the two sessions (see Figure 13, B).

Discrimination threshold of expression in infant faces (visual training group)

There were no significant effects of training on facial expression discrimination performance in the visual training group. Performance was similarly high before training (*Mdn* = 2.00, *IQR* = .50), after a single training session (*Mdn* = 2.25, *IQR* = 1.98; *z* = 1.23, *p* = .22, *r* = .28) and after two training sessions (*Mdn* = 2.00, *IQR* = .75; z = .26, *p* = .79, *r* = .06).

Effects of Training on Sensitivity to Distress in Infant Cries

As with Experiment 2, performance on the distress task was compared before and after training. Data were normally distributed, as confirmed by K-S tests (*p* > .20). Mean performance on this task was above chance (*M* = .68, *SD* = .11).

In order to investigate the effects of auditory and visual training sessions on the sensitivity to distress task, paired t-tests comparing performance on this task before and after training were conducted. . These analyses demonstrated a significant improvement in task performance in the auditory training group (*t*(18) = -2.15, *p* = .04, *r* = .45), but not in the visual training group (*t*(18) = -1.70, *p* = .11, *r* = .37). In the auditory training group, mean scores were significantly higher after training (percentage of correct trials: *M* = .74, *SD* = .14) compared with before training (*M* = .69, *SD* = .13), meaning that individuals were choosing more of the higher pitched stimuli as sounding more distressed. There was no significant change in performance in the visual training group after training (*M* = .70, *SD* = .13) compared with before training (*M* = .67, *SD* = .09). Within the auditory training group, paired samples t-tests were used to assess changes in performance at individual levels on the task. Results found significant increases in accuracy at the lower two levels on the task (0.5 and 1 semitone difference; *t*(18) = -2.10, *p* = .05, *r* = .44; *t*(18) = -2.86, *p* = .01, *r* = .56, respectively; see Figure 13, C-D).

The Effect of Gender on Perceptual Abilities Before and After Training

As with Experiment 2, the performance of men and women was compared using independent samples t-tests. No significant differences in pitch discrimination performance were observed: i) before training (*t*(36) = .59, *p* = .56, *r* = .10), ii) after a single training session (*t*(36) = -1.18, *p* = .25, *r* = .19) or iii) after two training sessions (*t*(36) = -1.34, *p* = .19, *r* = .22). There were no gender differences in performance on vocalisation clip 2: i) before training (*t*(18) = .27, *p* = .79, *r* = .06), ii) after a single training session (*t*(18) = .32, *p* = .76, *r* = .08), or iii) after two training sessions (*t*(18) = .17, *p* = .87, *r* = .04). Finally, there were also no gender differences in performance on the distress task before (*t*(36) = .74, *p* = .46, *r* = .12) or after (*t*(36) = .56, *p* = .58, *r* = .09) training.

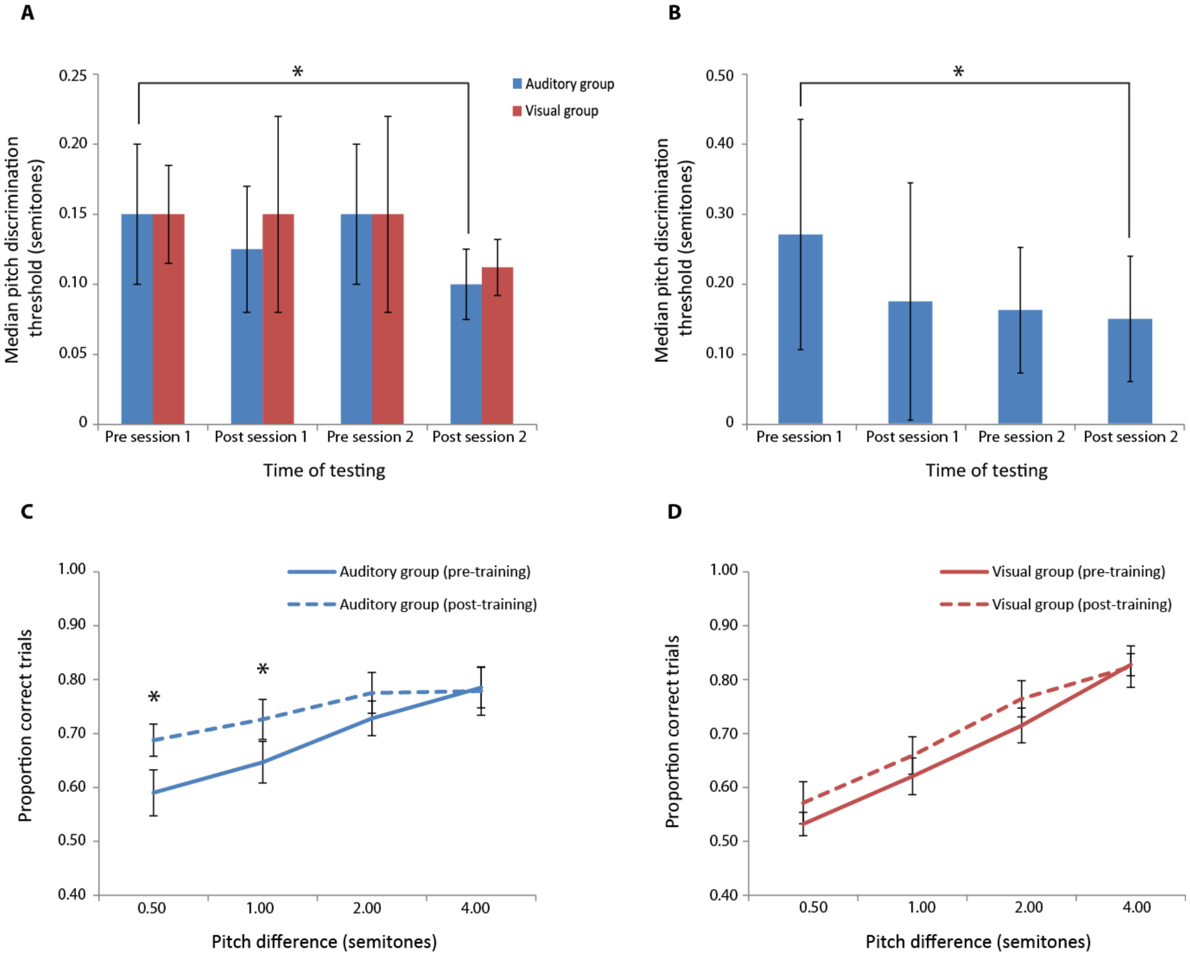

Figure 13. Performance on pitch discrimination and sensitivity to distress tasks before and after two training sessions (Experiment 3). A) Pitch discrimination performance for vocalisation clip 1 in the auditory and visual training groups. A significant improvement in pitch discrimination was observed after two training sessions in the auditory training group only. B) Pitch discrimination performance for vocalisation clip 2. A significant improvement in pitch discrimination was observed after two auditory training sessions. C-D) Change in performance on the distress task after two training sessions, graphs demonstrate the proportion of higher pitched stimuli selected as sounding more distressed (proportion correct trials). Significant improvements were observed in the auditory training group (C), on the two most difficult levels of the task but no change was observed in the visual training group (D). Error bars represent inter-quartile range (A-B) or mean +/- standard error (C-D). *denotes a significant difference ($p < .05$)

4.4 Discussion

4.4.1 Summary of Findings

Experiment 1 demonstrated a selective impairment in sensitivity to infant vocal distress in adults with depression and no previous musical training. Musicians with depression performed as well as healthy adults on this task, indicating that musical training may act as a potential protective factor against reduced sensitivity to infant vocal distress in depression. Findings from Experiment 2 demonstrated that a short perceptual training intervention significantly improved healthy adults' sensitivity to infant vocal distress. While a single training session resulted in specific improvements in discriminative ability of training stimuli, findings from Experiment 3 demonstrated that a second training session on a consecutive day resulted in generalised improvements to other stimuli and improved performance on a measure of sensitivity to infant vocal distress.

4.4.2 Discussion of Findings from Experiment 1

Consistent with previous studies and findings presented in Chapter 2, Experiment 1 demonstrated a strong association between infant cry pitch and perceived distress, with participants reliably interpreting higher pitched cry vocalisations as sounding more distressed (e.g., Schuetze & Zeskind, 2001). Adults with depression and no previous musical training demonstrated reduced sensitivity to infant vocal distress. This is comparable to previous findings demonstrating associations between maternal depression and reduced sensitivity to the pitch (Donovan et al., 1998) and emotional content (Lester et al., 1995; Schuetze & Zeskind, 2001) of infant cries. The impact of depression on appropriate responsiveness to infant vocalisations may be related to alterations to basic perceptual processing of infant cues. This possibility deserves further investigation as training interventions to improve basic perceptual processing may have a role for the improvement of caregiving sensitivity in depression.

Previous findings investigating the impact of depression on the perception of emotional cues have demonstrated negative interpretation biases. Adults with depression have been shown to respond more negatively when rating: adult vocal prosody (Emerson et al., 1999; Péron et al., 2011); adult facial expressions (Gollan et al., 2008; Leppänen et al., 2004); music (Naranjo et al., 2011; Punkanen, Eerola, & Erkkila, 2011); and even infant facial expressions (Arteche et al., 2011; Stein et al., 2010). Present findings demonstrating reduced sensitivity to infant vocal distress may be implicated in these negative biases. It is plausible that reductions in sensitivity to communicative cues introduce a level of ambiguity in the interpretation of these signals. Negative cognitive biases in depression may then result in overtly negative interpretations. This possibility should be considered further in studies exploring the relationship between sensitivity to communicative cues and biases in the appraisal of emotional content. In addition, more direct comparison of responses to infant vocalisations compared to other infant cues and emotional stimuli in adults with depression would allow investigation into these associations in depression more generally.

Adults with depression, who had small amounts of previous musical training, performed as well as healthy adults on this task. This suggests that musical training may act as a ‘protective factor’ against the processes in depression that disrupt sensitivity to infant vocal cues. This perceptual advantage of musicians is in fitting with previous literature demonstrating broad advantages in the auditory perceptual abilities of musicians (e.g., Lima & Castro, 2011; Thompson et al., 2004). Notably, the effect of musical training was apparent using a moderate definition of musicianship (four or more years of music lessons), relative to other studies (e.g., 'at least 10 years of lessons, starting before age five, with at least four hours of practice each week', Musacchia et al., 2007). The ameliorative effect of common levels of musical training on sensitivity to distress in infant vocalisations is of particular interest for the investigation of protective factors. Previous experience of musical training may help some individuals to maintain responsiveness to their infant's vocalisations, despite the presence of depressive symptoms.

The opposing influences of depression and musical training on sensitivity to infant vocal distress highlight the impact of experience on mechanisms supporting appropriate responsiveness to infant vocalisations. The current findings do not address how much or what type of training is required to improve sensitivity to infant vocalisations. Previous findings have suggested that focussed perceptual training interventions can enhance auditory skills to a comparable level to that observed in musicians (Micheyl et al., 2006). Experiments 2 and 3 sought to investigate whether similar perceptual training protocols could be used to enhance sensitivity to distress in infant cries.

4.4.3 Discussion of Findings from Experiments 2 and 3

Findings from Experiment 2 demonstrated preliminary evidence showing that it is possible to enhance adults' sensitivity to infant vocal distress through a short, targeted training intervention. A single 40-minute training session using a 3-AFC oddball paradigm led to specific improvements in pitch discrimination threshold of the stimulus used during training. Experiment 3 demonstrated that two training sessions on consecutive days resulted in more generalised improvements in pitch discrimination in infant vocalisations. Discrimination thresholds were improved on trained and untrained stimuli and sensitivity to infant vocal distress was enhanced. Together, these findings suggest that targeted perceptual training interventions might be able to improve sensitivity to infant vocalisations. These findings are in line with a large body of previous research, demonstrating a high degree of plasticity in human auditory and visual perceptual systems (Amitay, Irwin, & Moore, 2006; Karni & Sagi, 1993).

Throughout the literature on perceptual learning, there is a debate over whether learning reflects true modifications in perceptual abilities, or improvements in the response requirements of a task (known as procedural learning). Recent findings have demonstrated that early and robust effects of perceptual training programmes can be largely attributed to perceptual, rather than procedural processes (Hawkey et al., 2004). In the current experiment, measures were taken to minimise and assess the effects of procedural learning. First, a number of introductory trials were completed by all participants before the first testing session began, familiarising them with the response demands of the task. Secondly, a control group who completed a comparable visual training paradigm were included in the current design. These participants completed the same number of trials, with the same procedural demands, but responded to visual rather than auditory stimuli. While findings from Experiment 2 demonstrated an effect of visual training on pitch discrimination, these effects were not replicated in the larger group of participants tested in Experiment 3. The specific impact of auditory training in Experiment 3 suggests that at least some impact of training interventions is likely related to perceptual learning processes. This is further supported by observed improvements on the ‘sensitivity to distress’ measure, a task with different procedural demands.

Findings from these experiments also provide some insight into processes involved in the interpretation of distress in infant cry vocalisations. Participants performed well when discriminating infant cry stimuli by pitch alone. They demonstrated accurate responses to stimuli varying by small fractions of a semitone (pitch discrimination task). The same participants performed poorly when discriminating the same sounds on the basis of how distressed they seemed. They demonstrated inaccurate responses to stimuli varying by half a semitone (sensitivity to distress task). While the pitch threshold task could be described as ‘purely’ perceptual, the sensitivity to distress task additionally required appraisal of the affective content of stimuli. This appraisal therefore appears to be influenced by factors additional to the acoustic properties of infant cries. Experiments presented here suggest that these factors include current mood, previous musical experience and perceptual training.

4.4.4 Limitations of Experimental Procedures

The main limitation of Experiment 1 is that factors additional to the presence of musical training may have impacted on performance in this group. Factors such as socio-economic status and family support may be associated with greater likelihood of experiencing musical training. These factors may independently affect performance on an auditory task requiring attentive listening. As these factors were not measured in Experiment 1, their influence on results observed cannot be ruled out.

The main limitation of Experiments 2 and 3 was that measures of pitch discrimination threshold demonstrated ceiling effects. This may have masked the ability to measure the true impact of the training intervention, and may have prevented the detection of significant changes in some training conditions. Secondly, previous research has suggested a role for individual differences in initial perceptual abilities on the extent and success of perceptual learning (e.g., Amitay et al., 2005). These findings suggest that individuals with poorer initial listening abilities might be less able to learn through training paradigms. Individual differences in initial performance were not taken into account in the current study design. Future consideration of individual differences would be vital for the assessment of individuals who might benefit from this type of training.

Experiment 2 demonstrated that a short, focussed, training intervention can improve adults' pitch discrimination threshold of infant cries. Experiment 3 found that two auditory training sessions on consecutive days led to robust and generalisable improvements in pitch discrimination and sensitivity to distress in infant cries. The ability to improve sensitivity to distress in infant cries is of potential value for situations in which sensitivity in caregiver-infant interactions is disrupted.

4.4.5 Conclusions and Future Directions

The studies described in this chapter highlight a high degree of plasticity in mechanisms supporting appropriate responsiveness to infant vocalisations. These mechanisms can be affected by contextual and experiential factors including current mood, previous musical training and focussed perceptual training. Investigation of the neural mechanisms supporting these changes in responsiveness would be of much interest for understanding the nature of parental sensitivity to infant cues. It has been suggested that the specificity of improvements effected by perceptual training interventions is indicative of modifications to early stages in sensory processing systems (Irvine et al., 2000). From the field of visual perception, the role of expertise in face processing has been consistently linked with changes in functional activity in specialised areas of the cortical visual pathway (the fusiform gyrus; Gauthier, Tarr, Anderson, Skudlarski, & Gore, 1999). Studies assessing the neural correlates of heightened auditory perceptual abilities in musicians have demonstrated modulation in a number of areas of auditory cortical regions (e.g., Schneider et al., 2002). Recent evidence also indicates plasticity in the auditory brainstem, related to levels of musical expertise (Musacchia et al., 2007). Investigation of the impact of contextual and experiential factors affecting interpretation of infant vocalisations on underlying neural mechanisms could do much to increase understanding of the plasticity of the parental brain.

The experiments described in this chapter tested the perceptual abilities of non-parent participants, listening to the cries of unknown infants. The extent to which these findings can be applied to real-life caregiver-infant interactions is therefore uncertain. Previous research has suggested that parents have heightened sensitivities to infant cues in general and to cues from their own infants specifically (Green et al., 1987). It is possible that the learning processes involved in becoming attuned to cries from a specific infant are not directly comparable with the learning processes involved in general sensitivity to pitch in infant vocalisations. Future work investigating these questions with the current experimental design would be of much interest in this respect.

From a clinical perspective, these findings suggest that adults' sensitivity to infant vocal cues is malleable and can be rapidly enhanced using a short non-invasive intervention. As reviewed earlier, there is evidence to suggest that sensitivity to infant cues can be disrupted in depression. The investigation of the impact of auditory training on sensitivity to infant vocal distress in adults with depression would be of much clinical interest. However, the efficacy of such interventions in individuals with depression is as yet unclear. Effective training regimes of this kind would potentially have important implications for interventions aiming to improve sensitivity in caregiver-infant interactions.

CHAPTER 5: A SPECIFIC ROLE FOR THE PERIAQUEDUCTAL GRAY IN HUMAN CAREGIVING BEHAVIOUR

Infant vocalisations are biologically salient environmental sounds, subject to specialised processing in the brain (e.g., Lorberbaum et al., 2002). Neuroimaging studies of the human ‘parental brain’ have demonstrated a network of cortical and subcortical brain regions involved in the processing of infant cues (see section 1.5.1). Of these regions, the PAG of the midbrain has been ascribed a vital role in the production of maternal behaviour in rodents (Numan & Insel, 2003). Investigation of the role of the PAG in responding to infant cues in the human brain has previously been limited. This is in part due to methodological challenges of studying this small, deep subcortical structure with non-invasive methods. In this chapter, the role of the human PAG in the processing of infant vocalisations was investigated. This was achieved using invasive recordings made from electrodes implanted in this region in four patients who underwent deep brain stimulation (DBS) treatment for chronic pain.

5.1 Introduction

Across mammalian species, vocalisations from offspring serve to re-establish contact between a separated infant and their caregiver (Newman, 2007). The neural substrates underlying these evolutionarily conserved behaviours are likely to involve phylogenetically older brain regions, such as the ‘survival circuits’ of the brain (LeDoux, 2012). These circuits centre around regions of the midbrain and allow reflexive-type responses to environmental stimuli, while modulating wider brain processes to mobilise more elaborate responsive behaviours (LeDoux, 2012). One region in particular, the PAG, has been implicated in a range of survival-related processes. Recently conceptualised as an ‘interface for behavioural control’ (Benarroch, 2012), the PAG has previously been shown to be involved in specialised processing of emotionally salient stimuli and the expression of survival-related behavioural responses (Fanselow, 1994).

5.1.1 Anatomical connectivity of the PAG

Located in the midbrain, the PAG has an extensive network of anatomical connections with subcortical and cortical regions. Subcortically, it is closely interconnected with a number of midbrain and brainstem areas. Among these regions, the PAG receives inputs from auditory processing areas, such as the inferior colliculus (Dujardin & Jürgens, 2005). Primary subcortical outputs of the PAG include areas of the brainstem involved in regulating motor responses and physiological arousal (Holstege, Bandler, & Saper, 1996). Cortical connections of the PAG primarily consist of frontal regions, including the OFC (Cavada et al., 2000), known to be involved in emotional and reward-related processing (Kringelbach, 2005). These connections highlight the PAG as an interface between sensory and cognitive information processing, and as a key region in the control of autonomic functions and reflexive-type behaviour (Benarroch, 2012).

5.1.2 Current Understanding of the Function of the PAG

It is well-established in the animal literature that the PAG is implicated in the production of survival behaviours toward stressful environmental stimuli (such as fight, flight or freeze responses; Fanselow, 1994). Investigation of the functioning of the PAG in the human brain is limited by the challenges presented when attempting to study a small, deep subcortical structure with non-invasive neuroimaging techniques. The PAG is not accessible to MEG or EEG as both of these methods are primarily sensitive to cortical brain activity (Hillebrand & Barnes, 2002). Localisation of the PAG using fMRI and positron emission tomography (PET) is possible, although movement of the brainstem as a result of pulsatile cardiac blood flow presents additional challenges of motion artefacts (Poncelet, Wedeen, Weisskoff, & Cohen, 1992; Zhang et al., 2006). A number of studies have optimised fMRI parameters to improve the ability to detect activity in the PAG (e.g., Tracey et al., 2002).

A recent meta-analysis of the function of the PAG identified evidence from human neuroimaging studies suggesting involvement in the processing of emotionally salient stimuli (Linnman, Moulton, Barmettler, Becerra, & Borsook, 2012). The most well-established human neuroimaging findings, consistent with evidence from animal studies, have implicated the PAG in neural activity related to the experience of pain (Tracey et al., 2002; Wiech & Tracey, 2009). Other studies have demonstrated PAG activity in response to emotional visual or auditory stimuli. Studies of visual stimuli have demonstrated PAG activity in response to aversive images (Petrovic et al., 2005) and sad facial expressions (Kim et al., 2009). Similarly, studies of auditory stimuli found PAG activity both when listening to aversive sounds (e.g., nails on a blackboard; Zald & Pardo, 2002) and when experiencing pleasurable ‘chills’ while listening to music (Blood & Zatorre, 2001).

Specifically in relation to vocal stimuli, previous studies have demonstrated activity in the PAG in response to the sound of infant vocalisations (Laurent & Ablow, 2012b; Laurent et al., 2011) and in the production of voiced versus unvoiced speech (Schulz, Varga, Jeffires, Ludlow, & Braun, 2005). Direct stimulation of the PAG elicits vocalisations across a number of mammalian species, including rats, rhesus monkeys, and chimpanzees (e.g., Jürgens, 1994; Vanderhorst, Terasawa, Ralston Iii, & Holstege, 2000). Stimulation of the PAG in humans has also been shown to produce vocalisations (Sem-Jacobsen & Torkildsen, 1960), accompanied by experiences of strong emotions, including fear and distress (Tasker, 1982). It therefore seems that the human PAG may be involved in both the processing and production of vocalisations.

5.1.3 A Specific Role for the PAG in Caregiving Behaviour

Infant cues represent a privileged category of emotional cues which are high in biological salience, and much attention has therefore focused on PAG responses to such cues. The PAG is widely regarded as essential for the selection of maternal and other survival-related behaviours in response to salient stimuli in rodents (Numan & Woodside, 2010). Studies involving lesions to the rodent PAG have demonstrated disruption of the selection of appropriate behaviours towards pups in stressful conditions (Stern & Lonstein, 2001). Rodent maternal behaviour, such as pup-retrieval and active nursing, can also be disrupted by lesions to rostro-lateral PAG (Lonstein & Stern, 1997; Miranda-Paiva et al., 2007). Lesions to caudal PAG, on the other hand, resulted in increased maternal responses in the face of threatening stimuli, with a decrease in adaptive defensive responses (Sukikara et al., 2010). Together, these findings provide compelling evidence for the role of the PAG in the expression of maternal behaviour in rodents.

It has been strongly argued that the neural circuitry supporting evolutionarily-conserved behaviours is likely to be comparable across mammalian species (Corter & Fleming, 1990; LeDoux, 2012). In terms of caregiving behaviour, a number of neuroimaging studies have demonstrated activity in the human PAG in response to infant communicative cues. For instance the PAG was more active when adults heard sounds of infant cries, compared to frequency matched control sounds (Laurent et al., 2011) and when parents viewed images of their own infant's face, compared to images of other infants (Bartels & Zeki, 2004; Noriuchi et al., 2008). These studies highlight a role for the PAG in a wider neural network involved in responses to infant communicative cues (see section 1.5.1). Limitations in the capacities of non-invasive neuroimaging techniques have prevented focussed investigation regarding the timing of PAG involvement in response to infant cues. Understanding the temporal dynamics of neural activity may provide insight into the likely functioning of different regions.

Recent work has begun to demonstrate the temporal dynamics of cortical activity in response to infant cues, with early specialised processing of infant faces observed in the OFC using MEG (around 130ms after stimulus onset; Kringelbach et al., 2008; Parsons, Young, et al., 2013). This early activity in a reward-related cortical region is suggested to represent 'emotional tagging' of infant faces that may facilitate sensitive caregiving (Parsons, Stark, et al., 2013). Investigation of the timing of brain activity involved in the processing of infant vocalisations would allow greater understanding of the neural mechanisms supporting responsive caregiving behaviour to infant cues.

5.1.4 Deep Brain Stimulation

The study described in this chapter used an invasive technique to record neural activity directly from subcortical brain regions with high temporal resolution. Rare opportunities for direct recording from human brain structures are afforded by the increased utilisation of intracranial electrode implantation during DBS. DBS involves the delivery of an electrical pulse to specific subcortical structures through an intracranial electrode connected to a pulse generator (usually implanted under the skin of the chest). Therapeutic benefits have been observed with the use of DBS in a variety of otherwise treatment-resistant neurological conditions, including dystonia, essential tremor and chronic pain.

Procedures for DBS electrode implantation usually involve initial surgery to implant electrodes, leaving the connecting leads externalised to the head. In the week following surgery, the efficacy of different stimulation parameters is assessed and electrode leads are then internalised if a therapeutic effect is achieved. During this one week peri-operative period, there is a unique opportunity to directly record neural activity in targeted regions from implanted electrodes. In this manner, local field potentials (LFPs) can be recorded, providing a metric of localised neural activity. These rare opportunities to record neural activity directly from the human brain provide highly valuable information regarding the temporal dynamics of activity in human subcortical brain structures, inaccessible using non-invasive neuroimaging techniques. For example, previous studies using LFP recordings from DBS electrodes have demonstrated neural signatures for the experience of pain in the PAG (Green et al., 2009); the initiation of motor movements in the basal ganglia (Brown & Williams, 2005); and a potential marker of conscious reportability of stimuli in the ACC (Thomsen et al., 2011).

5.1.5 Aims and Hypotheses

In the current study, activity in the PAG and sensory thalamus was investigated in four patients who had DBS electrodes implanted in these regions for the treatment of chronic pain (Kringelbach, Pereira, Green, Owen, & Aziz, 2009; Young & Brechner, 1986). While performing an implicit listening task, patients heard clips of infant vocalisations varying in affect (cries, laughs and neutral ‘babbles’). LFP activity in response to these sounds was compared with responses to two categories of control sounds: constructed control sounds and ecological control sounds.

Constructed control sounds were artificially generated using physical parameters of the infant vocalisations, but with altered distribution of power in the frequency domain. These sounds were created to match the infant vocalisations in a number of physical parameters while lacking the emotional communicative properties of natural infant vocalisations. Ecological control sounds consisted of clips of adult female crying and domestic animal distress sounds (cats meowing and dogs whining). These stimuli were selected as natural sounds which convey emotion, but which do not necessarily elicit the same instinctive caregiving responses.

It was hypothesised that the PAG would demonstrate differential neural activity to infant vocalisations compared with either constructed control sounds or ecological control sounds. It was further hypothesised that no such differences would be observed in recordings made from the sensory thalamus.

5.2 Methods

5.2.1 Participants

Participants were four patients who underwent DBS surgery for chronic, intractable pain at the John Radcliffe Hospital, Oxford, UK. Three of the patients were male, aged 26-31 years and had no children. The male patients all had electrodes implanted in the lateral PAG and sensory thalamus. The fourth patient was female, aged 64 years and had two adult children. She had electrodes implanted in the medial PAG and sensory thalamus. Ethical permission for this study was obtained from the Oxfordshire Research Ethics Committee A (ref: 08/H0604/58). Participants provided written, informed consent.

5.2.2 Stimuli: Infant Vocalisations, Constructed and Ecological Control Sounds

Infant vocalisation stimuli were taken from the OxVoc Sounds Database (described in Chapter 2). Stimuli consisted of 1.5s sound clips of infant vocalisations varying in affect (15 infant cries, 15 infant laughs, 15 infant neutral vocalisations). Constructed control sounds were created to provide stimuli that were physically similar to the infant vocalisations, using the MIR toolbox for Matlab (Lartillot & Toiviainen, 2007). This procedure first involved a frame-by-frame analysis of the original infant stimuli (frame-length of 20ms, frame overlap of 10ms) to extract waveform amplitude values. These values were then multiplied by waveform amplitude values of a complex tone generated from sine waves with a F_0 = 500Hz and five harmonics (1000Hz, 1500Hz, 2000Hz, 2500Hz and 3000Hz). A Hamming window was also applied to each window to smooth frame edges. The resulting frames were then recombined to produce 1500ms control stimulus.

This resulted in electrostatic-like control sounds that match the original stimuli in RMS intensity, temporal pattern and overall frequency range. These stimuli sound clearly artificial and differ from the originals only in their exact frequency content (see Figure 14). Forty-five control stimuli were created, one to match each of the original 45 infant vocalisation stimuli. Ecological control sounds were also taken from the OxVoc Sounds Database. Forty-five stimuli were used, consisting of 15 adult cries, 15 cat meows and 15 dog whines. All stimuli used in this experiment were matched on duration and average RMS intensity and had 150ms linear rise and fall times applied to each clip.

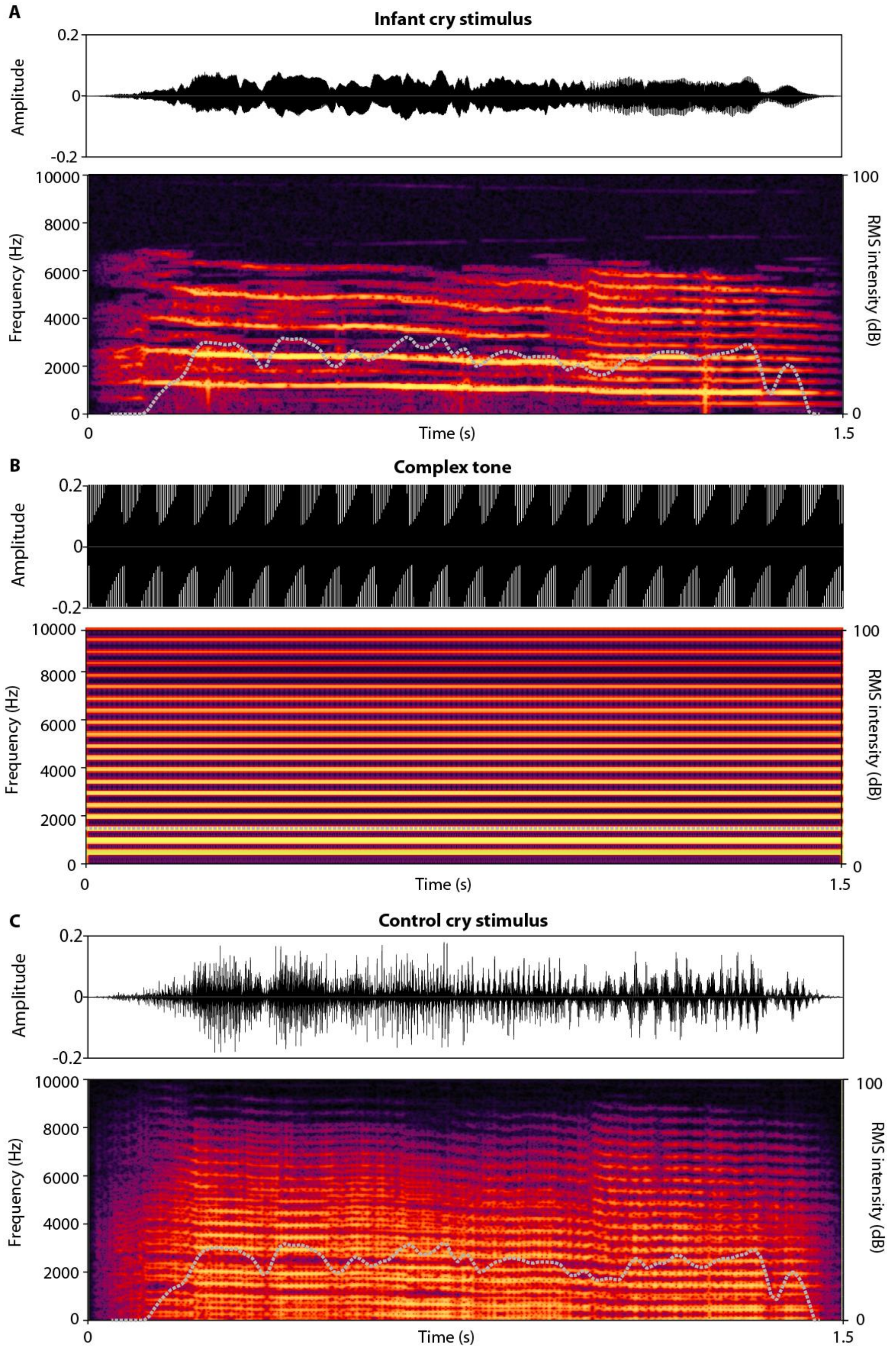

Figure 14. Waveforms and spectrograms of an example of an infant cry sound and its corresponding constructed control sound. A) The waveform amplitude (upper), spectrogram of frequency content (lower) and RMS intensity (superimposed white line) of a sample infant cry. B) Waveform amplitude and spectrogram of complex tone generated from sine waves (F_0 = 500Hz plus harmonics) C) The corresponding constructed control sound differs in frequency composition but retains the same temporal structure and amplitude envelope (RMS intensity over time).

5.2.3 Stimulus Presentation and Behavioural Task

Participants were presented with repeated blocks of stimuli from different sound categories. The length of the experimental session was shorter for patient one compared with the other patients, due to clinical time-constraints. The first patient tested (male, aged 26) listened to two categories of stimuli: infant vocalisations (15 cries, 15 neutral vocalisations and 15 laughs) and corresponding constructed control sounds (15 control cries, 15 control neutral vocalisations and 15 control laughs). Stimuli were presented in blocks contained one presentation of each of 45 unique stimuli from the same category. Patient 1 completed four repetitions of each stimulus block (4 repetitions of 2 categories of 45 stimuli = 360 trials in total). Each stimulus lasted 1.5s and an inter-stimulus-interval (ISI) was set at 2s, resulting in an experimental session lasting for 21mins. Patients two, three and four completed two additional blocks for each of these stimulus categories (in total, 6 repetitions of 2 categories of 45 stimuli = 540 trials). These patients also completed three blocks containing ecological control sounds (3 repetitions of 45 stimuli = 135 trials). In these patients, the ISI was shortened to 1s, resulting in a total experimental duration of 28mins.

During stimulus presentation, participants performed an implicit listening task in order to ensure maintenance of auditory attention. ‘Target’ and ‘distractor’ tones were pseudorandomly presented during the silent ISI (on average, one tone was presented every six stimuli, 50% were target tones, 50% distractor tones). These tones were 500ms pure tones at 400Hz and 500Hz and were randomly assigned as ‘target’ or ‘distracter’ tones. Participants were instructed to press a button each time they heard the target tone and to ignore the distracter tone. Stimuli were presented using Presentation® software (Version 14.4, Neurobehavioral Systems, Inc. www.neurobs.com) through Sony stereo headphones

(MDR-XD200). A parallel output cable transmitted stimulus triggers to the LFP recording system to allow millisecond precision in the logging of stimulus onset.

5.2.4 Electrode Placement and LFP Recordings

In line with standard procedures at Oxford Functional Neurosurgery, each patient was implanted with one electrode in the PAG and one in sensory thalamus (Kringelbach et al., 2009). Before surgery, anatomical MRIs were obtained from each patient. Patients were then fitted with a stereotactic frame (Cosman-Roberts-Wells base ring) and a computerised tomography (CT) scan was performed. Data from MRI and CT scans were combined (using localisation of the anterior and posterior commissures) and planning of coordinates for electrode placement was performed. Electrodes implanted were medtronic 3387 (Medtronic Neurological Division, Minneapolis, MN) electrodes, with four platinum-iridium cylindrical surfaces (1.27mm diameter, 2.0mm exposed tip and 1.5mm gap between two adjacent electrode contacts).

During surgery, electrodes were implanted in target locations and connected via extension leads to a pulse generator (IPG – Kinetra TM, Medtronic). Extension leads remained externalised for one week post-surgery and electrode settings were optimised for clinical outcomes (optimal stimulation is typically in the range of 20-40Hz, pulsewidth 90-300ms, 1-5V). If satisfactory pain relief was achieved, full implantation of a pulse generator was performed one week after initial DBS surgery. Immediately after surgery, CT scans were performed to confirm target location. Electrode points were first identified on post-operative CT scans (1x1x1mm). The FLIRT tool from the FMRIB Software Library (FSL version 4.1, www.fmrib.ox.ac.uk/fsl, University of Oxford) was used to perform a 12 parameter linear transform. This transformed individuals patients' post-operative CT scans into the coordinates of their pre-operative MRI and then into standard MNI (Montreal Neurological Institute) space. This set of transformations was then performed on individual electrode points (using nearest neighbour interpolation), which were then plotted in standard MNI space (see Figure 15 and Table 7).

LFP recordings from PAG electrodes were performed in all patients and recordings from thalamus electrodes were performed in patients 2-4. Electrode extension leads were disconnected from pulse generators and connected to a data acquisition unit (CED 1401 Mark II, Cambridge Electronic Design Ltd., Cambridge, UK). LFPs were recorded using a bipolar configuration from four circumferential contacts on each electrode. This method resulted in three channels of data reflecting the relative voltage difference between two electrode contact points. Data were filtered online (0.5-500Hz) and amplified by a factor of 10,000 using an electrically isolated amplifier (CED 1902, Cambridge Electronic Design Ltd., Cambridge, UK). Data were recorded with a measurement range of +/-500µV, resolution of 0.24µV and sampling rate of 2480Hz. LFPs were visualised online and saved using Spike II software® (CED version 5.0). Due to equipment failure, data from two channels (contacts 2-3 in patient 2 and patient 4, for both PAG and sensory thalamus electrodes) were not recorded. Analysis was performed using the Fieldtrip toolbox for MATLAB (http://fieldtip.fcdonders.nl, Oostenveld et al., 2011; MathWorks Inc., Natick, MA).

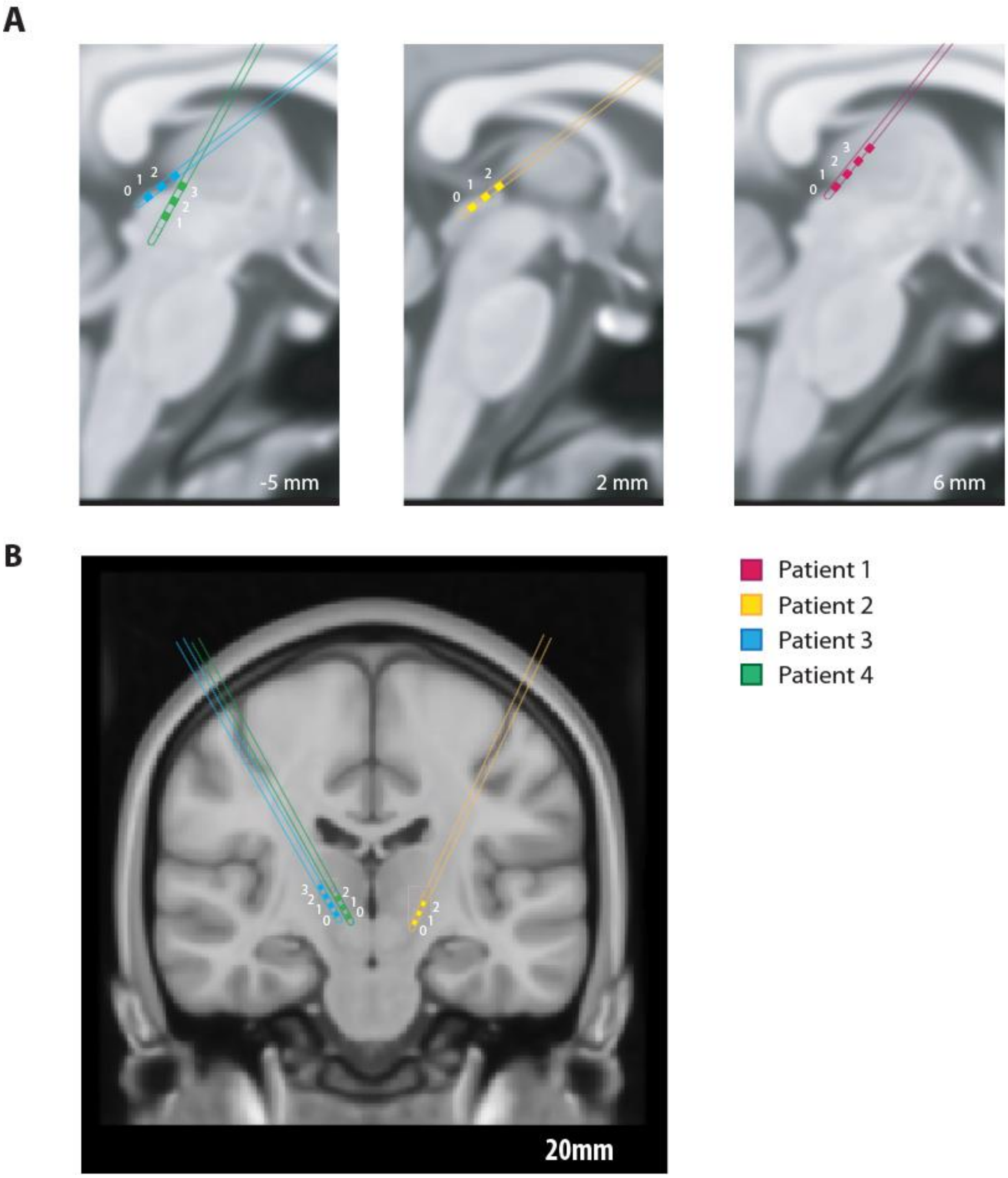

Figure 15. Deep brain stimulation (DBS) electrode placement in patients. A) Placement of PAG electrodes is represented on sagittal slices of standard MNI images. B) Placement of thalamic electrodes is represented on a coronal plane MNI image.

Table 7. *Standard MNI coordinates of each electrode contact point*

Electrodes	Patient 1			Patient 2			Patient 3			Patient 4		
	x	*y*	*z*	*x*	*y*	*z*	*x*	*y*	*z*	*x*	*y*	*z*
PAG												
Contact 3	-6	-21	8	*-4*	*-25*	*2*	7	-24	5	*6*	*-26*	*0*
Contact 2	-5	-24	5	-2	-27	0	6	-27	3	5	-28	-4
Contact 1	-5	-26	3	-1	-30	-3	4	-31	-1	4	-30	-6
Contact 0	-5	-28	0	-1	-33	-4	4	-33	-2	3	-31	-9
Thalamus												
Contact 3	-	-	-	*-17*	*-15*	*3*	14	-17	7	*16*	*-18*	*5*
Contact 2	-	-	-	-16	-18	0	11	-20	1	15	-20	2
Contact 1	-	-	-	-15	-20	-3	11	-21	0	13	-22	0
Contact 0	-	-	-	-14	-22	-5	10	-21	-2	12	-23	-2

Note. Coordinates in italics indicate contact points with no recorded data, due to a technical problem with the recording equipment. Dashes indicate no recorded data.

5.2.5 Data Analysis: Time-locked LFPs and Time-frequency Analysis

Two types of analysis were performed to investigate differential LFP activity across stimulus categories from macroelectrodes implanted in the PAG and sensory thalamus (see Figure 16 for overview). First, time-locked analyses investigated differences in the amplitude of averaged LFPs across stimulus categories at high temporal resolution. Secondly, time-frequency analyses were performed to assess differences in evoked power across alpha, beta and gamma frequency bands between stimulus categories.

Pre-processing

Data pre-processing was performed using the *ft_preprocessing* tools from the Fieldtrip toolbox. Data was first epoched from 200ms before stimulus onset to 500ms after stimulus onset. Line noise (50Hz and harmonics) was removed using 2nd order Butterworth notch filters (sub-function, *dftfreq*). This function performed discrete Fourier transforms to reduce the impact of ‘line noise’ artefacts, the frequency of oscillations of alternating current in mains electricity circuits. High-pass (2Hz) and low-pass (80Hz) filters were then applied to limit data to the frequency range of interest (sub-functions, *hpfilter* and *lpfilter*). A baseline correction (-200 to 0ms before stimulus onset) was also applied to data from each trial in order to remove any initial offsets in LFP amplitude (sub-function *demean*).

Trial-wise data was then averaged across channels within each individual patient and converted to standardised values in order to allow equal weighting of data from each patient in further stages of analysis (MATLAB functions *mean* and *zscore*).Variance across trials within each patient was then assessed and trials with a mean amplitude (averaged across the whole duration of the trial) greater than +/-3SD of the mean (averaged across all trials) were rejected. On average, between three and six trials were removed for each patient. Data were then averaged across patients, resulting in mean LFP time-courses for each trial. These data were then used for event-related potential and time-frequency statistical analyses.

Time-locked LFPs: Event-Related Potential (ERP) Analysis

Sample-by-sample one-way ANOVAs with stimulus category as a main factor were then performed (similar to methods described in Blair & Karniski, 1993; Guthrie & Buchwald, 1991; Kraskov, Quiroga, Reddy, Fried, & Koch, 2007). These analyses produced time-resolved significance levels for the difference in averaged voltage across stimulus categories.

Time-Frequency Analysis

Time-frequency analyses were performed using the multitaper method implemented in FieldTrip (*ft_freqanalysis*). A frequency range of 2-80Hz was analysed, using a sliding window with a width of 10ms from 200ms pre-stimulus to 500ms post-stimulus onset. Power spectra were plotted using a relative baseline correction and time-windows of interest identified for statistical analysis. One-way ANOVAs and post-hoc Scheffe tests were used to assess power differences in the alpha (8-12Hz), beta (13-20Hz) and gamma (30-80Hz) frequency bands.

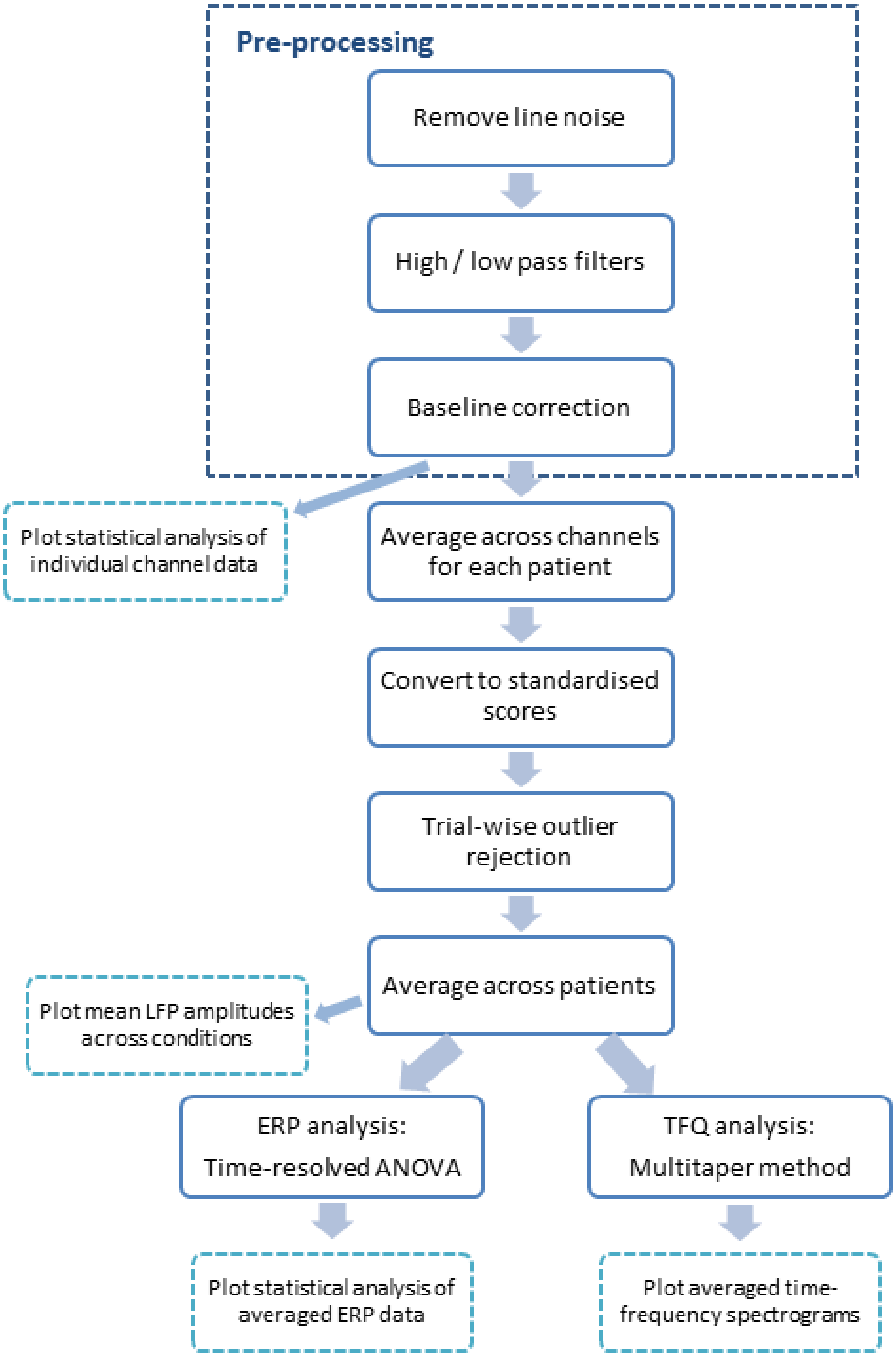

Figure 16. LFP analysis pipeline demonstrating stages of data analysis. Event-related potential (ERP) analysis allowed comparison of time-locked fluctuations in voltage of neural activity. These analyses are demonstrated in plots of mean LFP amplitudes for each stimulus condition (Figures 17A-C, 18A-B), statistical plots of comparisons across stimulus categories (Figures 17D-F, 18C-D) and individual channel statistical plots (Appendix D). Time-frequency analyses allowed investigation of time-locked changes in oscillatory neural power across categories. These analyses are demonstrated in spectrogram plots (Figures 19-20).

5.3 Results

5.3.1 Behavioural Task

Participants all performed at ceiling on the target tone detection task, with more than 95% correct button presses. A one-way ANOVA with stimulus category (infant vocalisations, constructed control sounds and ecological control sounds) as a within-subject factor demonstrated no significant differences in reaction times to the target tones during presentation of different stimulus categories ($F(2, 240) = .76$, $p = .47$, $r = .06$).

5.3.2 Event-related Potentials

Findings from the PAG

Sample-by-sample one-way ANOVAs were performed with stimulus category (infant vocalisation, constructed control sound, ecological control sound) as a main factor. A significant difference was observed in PAG activity in response to infant vocalisations compared with constructed control sounds beginning at 49ms post-stimulus onset ($F(1, 523) = 6.74$, $p < .01$, $r = 0.11$; see Figure 17). A similar analysis comparing responses to infant vocalisations with responses to ecological control sounds found a significant difference later in time, beginning at 85ms ($F(1, 401) = 6.92$, $p < .01$, $r = .13$, see Figure 17).

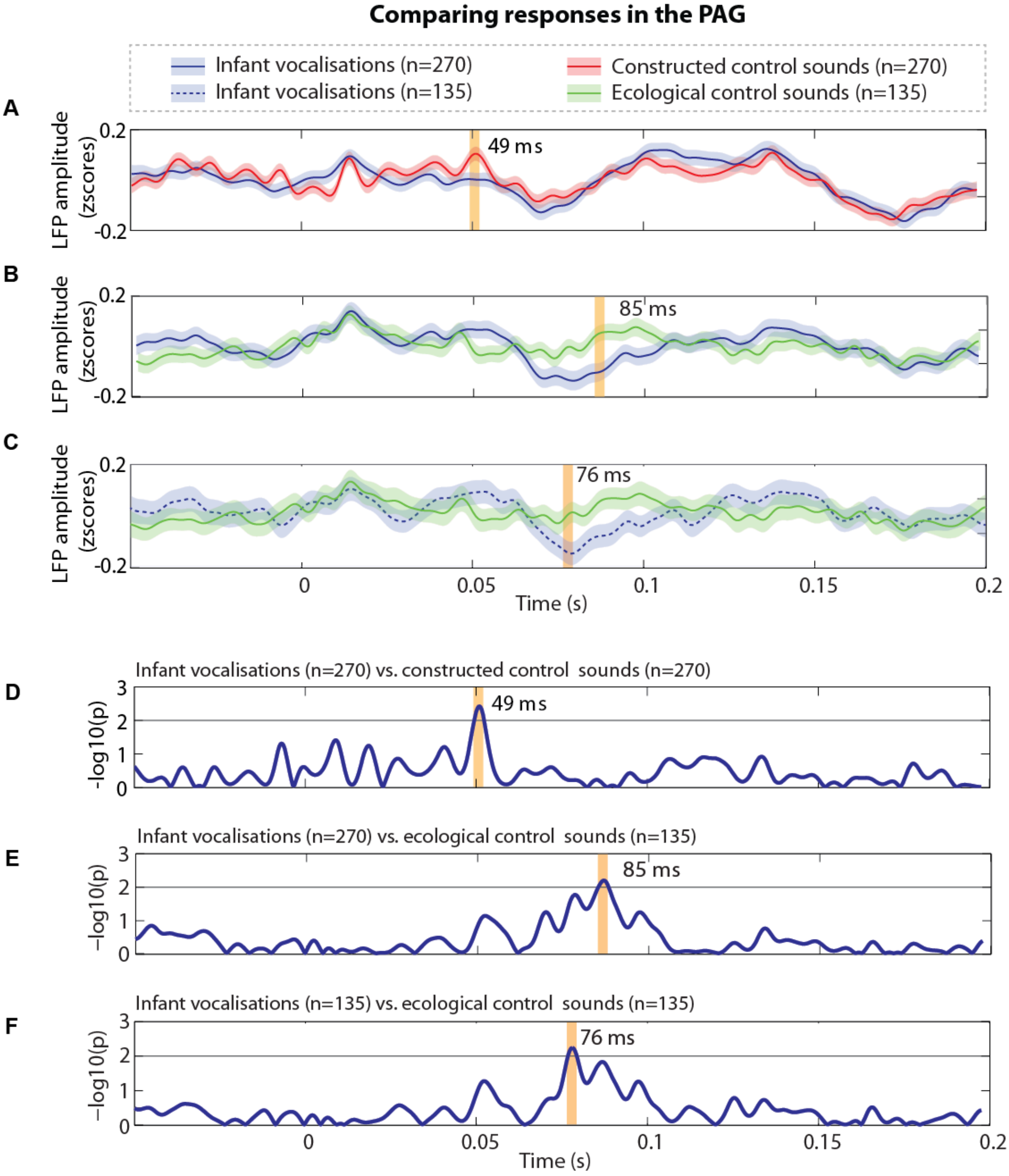

Figure 17. Comparison of event-related LFP activity in the PAG between stimulus categories. A-C) Mean time courses of LFP amplitudes across stimulus categories (shading indicates standard error). D-F) Statistical differences in LFPs recorded from the PAG showing normalized p-values obtained from ANOVA tests. A&D) Significant differences in averaged LFPs at 49ms in response to infant vocalisations compared with constructed control sounds. B&E) Significant differences in averaged LFPs at 86ms in response to infant vocalisations compared with ecological control sounds. C&F) An additional comparison using equal numbers of trials for infant vocalisations and ecological control sounds demonstrated slightly earlier significant differences, at 76ms. Horizontal lines indicate an alpha threshold of 0.01, orange highlighting indicates samples that cross the alpha threshold.

A number of additional analyses were performed to investigate other comparisons of interest. First, in the analysis of infant vocalisations compared to ecological control sounds described above, the number of trials of infant vocalisations ($n = 266$) was greater than the number of trials of ecological control sounds ($n = 135$). An additional analysis was performed with equal numbers of trials, using a random subset of 135 infant vocalisations. Sample-by-sample one-way ANOVAs demonstrated a significant difference beginning at 77ms ($F(1, 270) = 6.87$, $p < .01$, $r = .16$; see Figure 17). This is similar to the findings described above (beginning at 86ms) observed with the larger number of trials.

Comparisons of infant vocalisations varying in affect (e.g. infant cries compared with infant laughs) were not possible due to the unstable nature of averaged data over relatively small numbers of trials (90 trials per participant for each individual category). Preliminary analyses did not manage to achieve a stable signal in baseline periods (i.e. there were significant differences in LFP amplitude between stimulus categories in the pre-stimulus period). Analyses including significant differences in the pre-stimulus period would diminish confidence that any significant difference in the post-stimulus period was related to stimulus presentation rather than some external factor. Further analyses based on these comparisons were therefore not performed.

Analyses of PAG data on an individual channel basis were also carried out to explore the impact of individual patient data on the findings observed. These analyses revealed significant differences between infant vocalisations and constructed control sounds in all patients, beginning 17-80ms after stimulus onset. Differences between infant vocalisations and ecological control sounds were also found in all patients and generally occurred later in time, ranging from 44-184ms (see Appendix D).

Findings from the Sensory Thalamus

The analyses described above were repeated with data recorded from the sensory thalamus in three patients. No significant differences were observed in comparisons of either: i) infant vocalisations versus constructed control sounds or ii) infant vocalisations versus ecological control sounds (all $p > .01$, see Figure 18). Given the lack of a main effect of stimulus category, analyses of sub-categories of infant vocalisations were not performed.

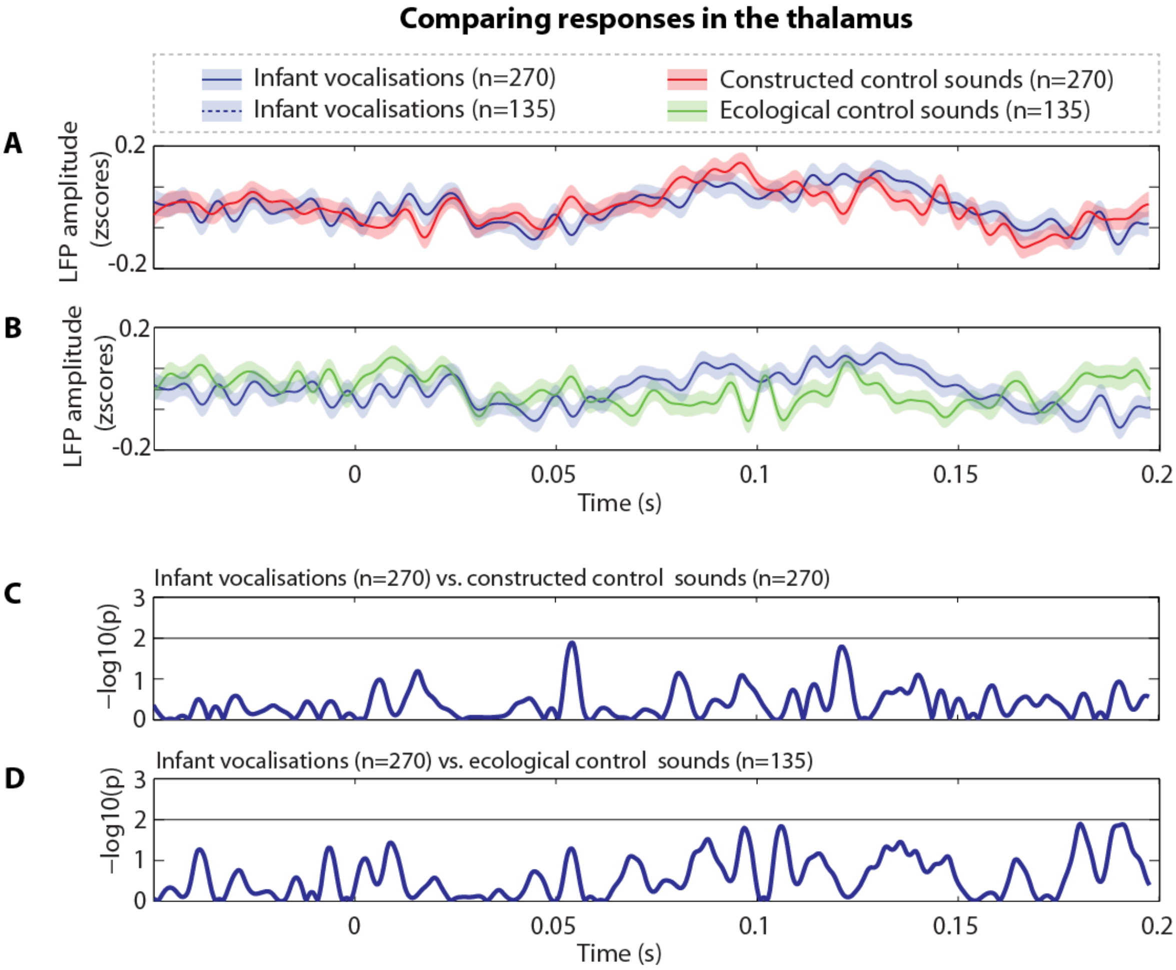

Figure 18. Comparisons of averaged event-related LFP activity in the thalamus. A-B) averaged time courses of LFP amplitudes (shading indicates standard error), C-D) statistical plots of differences between categories. No significant differences in time-locked activity were observed between: A&C) infant vocalisations and constructed control sounds or B&D) infant vocalisations and ecological control sounds. Plots represent normalized p-values from ANOVA tests, horizontal lines indicate an alpha threshold of 0.01.

5.3.3 Time-frequency Analyses

Findings from the PAG

Differences in patterns of oscillatory activity were investigated using time-frequency analyses. All stimulus conditions induced stronger activity in the alpha (8-12Hz) and beta (13-30Hz) frequency bands at around 100-200ms post-stimulus onset, relative to the baseline (pre-stimulus) period (see Figure 19).

Visual inspection of time-frequency plots identified 100-200ms post-stimulus onset as a time-window of interest. Differences in power during this time-window across stimulus categories were assessed using one-way ANOVAs. No significant differences in induced power were found between all infant vocalisations compared with a random subset of 135 infant vocalisations in the alpha ($F(1, 269) = 2.15$, $p = .14$, $r = .19$), beta ($F(1, 269) = 2.83$, $p = .09$, $r = .10$) or gamma frequency bands ($F(1,269) = 3.51$, $p = .06$, $r = .11$). Consequently, data from all infant vocalisations were used in all analyses.

One-way ANOVAs with stimulus category (infant vocalisations, constructed control sounds and ecological control sounds) as a factor demonstrated significant differences in the alpha ($F(2, 674) = 10.07$, $p < .001$, $r = .12$) and beta frequency bands ($F(2, 674) = 4.45$, $p = .01$, $r = .08$). No significant differences were found in the gamma frequency band ($F(2, 674) = 1.32$, $p = .27$, $r = .04$).

For the alpha band activity (8-12Hz), post-hoc Scheffe tests revealed significant differences between the ecological control sounds and infant vocalisations ($p < .0001$), and between the ecological control sounds and constructed control sounds ($p < .0001$). No differences were observed between the infant vocalisations and constructed control sounds (*p* = .83). In the beta band (13-30Hz), post-hoc Scheffe tests demonstrated a significant difference in infant vocalisations compared with ecological control sounds (*p* < .01). There were no significant differences between infant vocalisations and constructed control sounds (*p* = .71) or between constructed control sounds and ecological control sounds (*p* = .07).

Findings from the Sensory Thalamus

Similar analyses were conducted for data from the thalamic electrodes. One-way ANOVAs with stimulus category as a factor showed no significant differences in power from 100-200ms in alpha (*F*(2, 269) = 1.67, *p* = .19, *r* = .08), beta (*F*(2,269) = .87, *p* = .42, *r* = .06) or gamma (*F*(2, 269) = .87, *p* = .42, *r* = .06) frequency bands (see Figure 20).

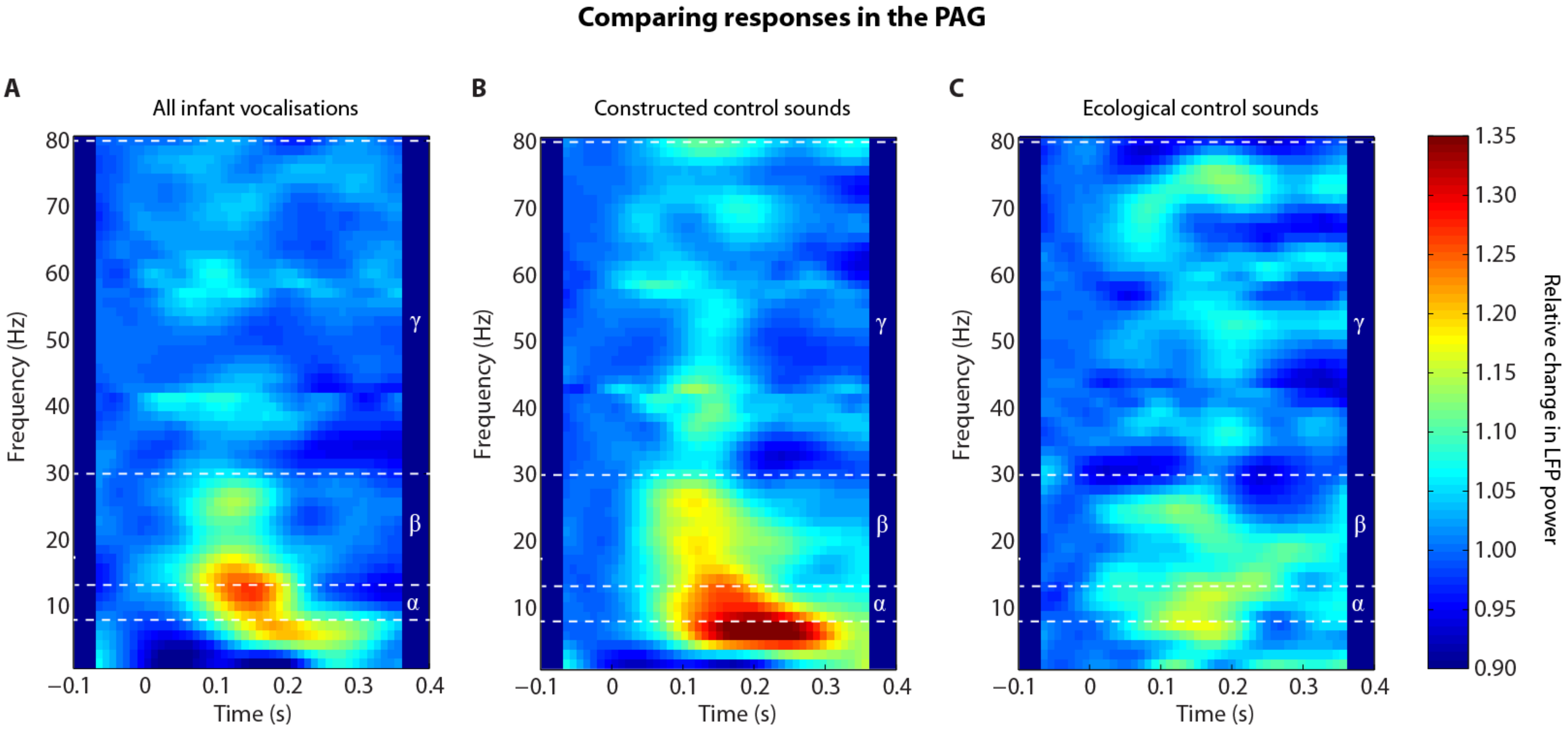

Figure 19. Time-frequency plots in the PAG representing post-stimulus power in different stimulus conditions: A) Infant vocalisations, B) constructed control sounds, C) ecological control sounds. Between 100-200ms post-stimulus onset, significant differences in power in the alpha frequency band (8-12Hz) were observed between infant vocalisations (A) and ecological control sounds (C); and between constructed control sounds (B) and ecological control sounds (C). Significant differences in power in the beta frequency band (13-30Hz) were observed between infant vocalisations (A) and ecological control sounds (C). No significant differences in the gamma frequency band (30-80Hz) were observed. White dashed lines indicate boundaries of frequency bands.

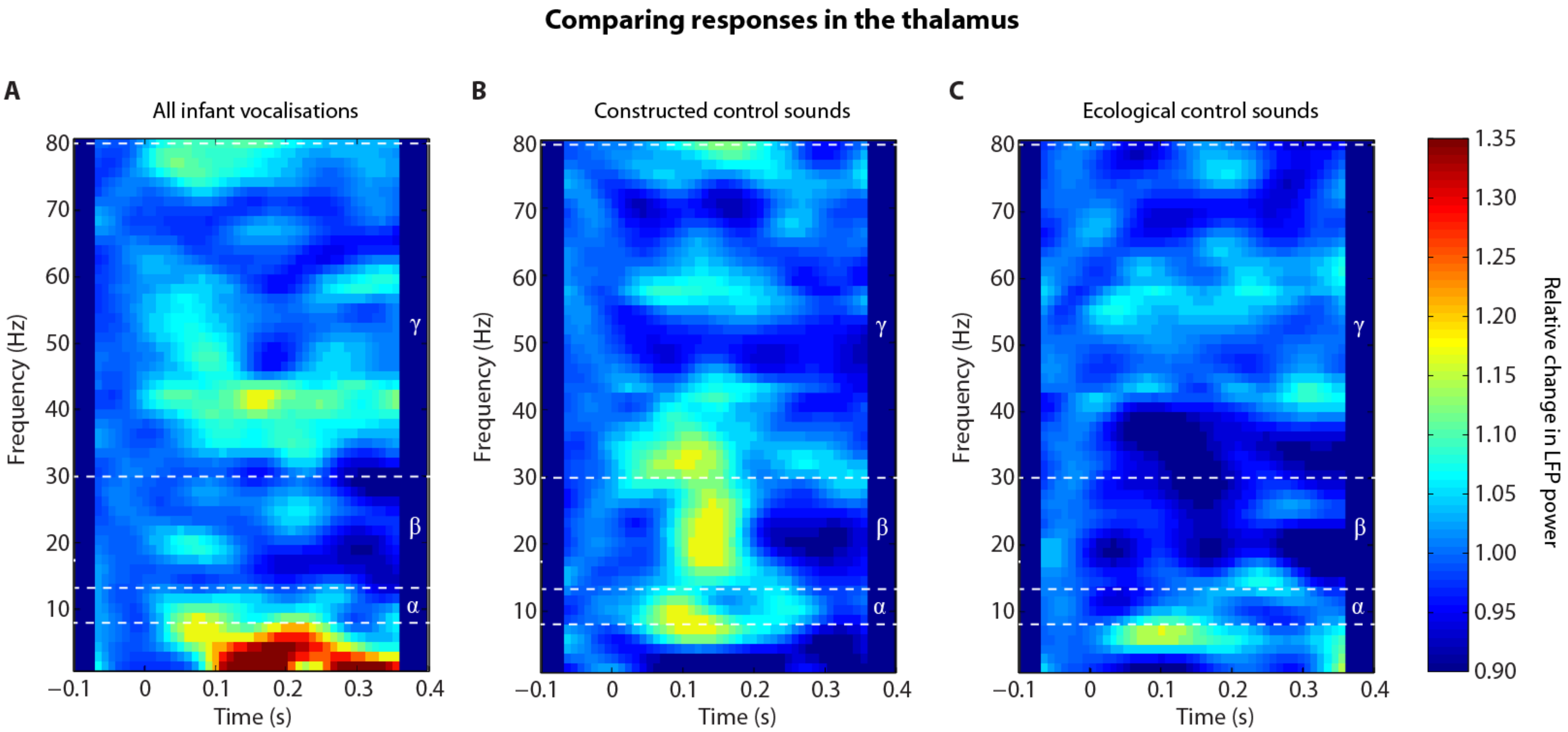

Figure 20. Results of time-frequency analyses performed on averaged LFP recordings from the thalamus in response to different stimulus categories: A) infant vocalisations, B) constructed control sounds, C) ecological control sounds. There were no significant differences in power of activity between 100-200ms in the alpha (8-12Hz), beta (13-30Hz) or gamma (30-80Hz) frequency bands. White dashed lines indicate boundaries of frequency bands.

5.4 Discussion

Findings from this study provide evidence suggesting that infant vocalisations may be differentiated from other sounds early in the auditory pathway. Specifically, significant early differences were observed in event-related activity recorded from the PAG of the midbrain. These differences were observed between responses to infant vocalisations and either physically similar sounds or emotional vocalisations from adults and animals. In addition, there were significant differences in power of oscillatory activity in the alpha and beta frequency bands between responses to infant vocalisations and ecological control sounds recorded from the same electrodes. No such differences were observed in LFP activity recorded from electrodes implanted in the sensory thalamus.

5.4.1 ERP Analyses

Results of ERP analyses of data recorded from electrodes implanted in the PAG revealed significant differences in response to infant vocalisations and constructed control sounds at 50ms after stimulus onset. Activity differentiating infant vocalisations from ecological control sounds was observed slightly later in time, at around 80ms. No significant differences in event-related activity between categories were observed in LFP recordings from electrodes implanted in the sensory thalamus. These findings demonstrate that activity early in the auditory pathway discriminates physically and emotionally similar auditory stimuli in less than 100ms. It is tentatively suggested that discrimination of sounds from the same functional category (i.e. emotional vocalisations) may be more time-consuming than discrimination of natural from unnatural sounds. However, the functional significance of this difference in timing is unclear.

5.4.2 Time-frequency Analyses

Results from time-frequency analyses demonstrated that low frequency activity in recordings from the PAG differentiated infant vocalisations and control sounds from ecological control sounds within 100-200ms after stimulus onset. In the alpha band, significant differences were observed between i) infant vocalisations and ecological control sounds, and ii) constructed and ecological control sounds. In the beta band, significant differences were observed between infant vocalisations and ecological control sounds. Time-frequency analyses of LFP recordings from sensory thalamus demonstrated no such differences. These findings suggest that induced power changes in low frequency activity in the midbrain are sensitive to physical differences between different types of vocalisations.

The functional significance of low frequency oscillations in the brain is a topic of much debate. Originally thought to be involved in ‘idling’ or ‘inhibition’, more recent studies suggest that synchronisation of alpha activity with beta and gamma oscillations is essential for perceptual and memory processes (Palva & Palva, 2007). It has been suggested that the function of oscillatory activity in different brain regions may vary (e.g., Meltzer et al., 2008), but this has not been extensively investigated in human subcortical brain regions.

5.4.3 Implications of Findings

In terms of caregiving behaviour, the early specialised neural activity in the midbrain observed here may be involved in the identification of infant cues and the initiation of rapid responses. While LFP recordings were obtained from electrodes implanted in the PAG of four patients, it is not possible to conclusively state that differences in neural activity observed occurred as a result of local processing within the PAG as the spatial resolution of LFPs in subcortical areas is currently debated. Some studies, employing microelectrode arrays, estimate the spread of neural activity to be around 0.25mm in subcortical brain areas, while estimates from cortical regions range from 20-40µm up to 6mm from the source (Berens, Keliris, Ecker, Logothetis, & Tolias, 2008; Kajikawa & Schroeder, 2011; Katzner et al., 2009; Kreiman et al., 2006; Nauhaus, Busse, Carandini, & Ringach, 2009; Wang, Ulbert, Schomer, Marinkovic, & Halgren, 2005; Xing, Yeh, & Shapley, 2009). Other studies have demonstrated much wider spreading of neural activity. For example, auditory brainstem responses are evoked potentials detectable using electrodes located on the scalp, many centimetres from their recognised source in subcortical areas (e.g., Skoe & Kraus, 2010).

This process, known as volume conduction, is likely to depend on many factors, including the strength of the electrical potential. It is plausible that activity reported here is affected, to a greater or lesser extent, by activity from other subcortical or cortical areas (Brown et al., 2001). The most likely candidate for this type of modulatory activity is the inferior colliculus of the midbrain (particularly the brachium), which is located close to the PAG and is a major relay of the primary auditory pathway. Further clarification of this possibility would require simultaneous recording of electrical potentials from multiple midbrain nuclei, including the PAG and inferior colliculus.

Recent empirical work addressing this issue in different human subcortical regions (the subthalamic nucleus and substantia nigra of the basal ganglia) demonstrated that recorded LFPs were likely to be a combination of locally generated potentials and volume conduction of distally generated potentials (Alavi, Dostrovsky, Hodaie, Lozano, & Hutchison, 2013). Such studies using multiple microelectrode recordings in human patients are necessarily rare. Future work using animal models to investigate LFP activity within the PAG and other subcortical brain regions in response to emotional vocalisations would be of much use to investigate these findings further.

Findings of early differential activity in subcortical brain regions, perhaps localised to the PAG, are plausible in the context of existing literature. Previous studies investigating the human PAG have demonstrated selective activity in response to emotional or biologically salient stimuli, including infant communicative cues (Laurent et al., 2011; Mobbs et al., 2007). Animal studies have demonstrated a vital role for the PAG in the expression of maternal behaviour in rodents (Lonstein & Stern, 1997). The current study would suggest that activity in the PAG may support early specialised processing of infant vocalisations. The established connectivity of the PAG with brainstem and frontal cortical areas would further suggest this region may be uniquely placed to initiate rapid caregiving responses to infant vocalisations.

In terms of auditory processing, findings presented here suggest that auditory stimuli varying in salience may be differentiated at an earlier stage of the auditory pathway than previously examined. This is in fitting with other findings in auditory neuroscience demonstrating complex analysis processes in early stages of the auditory pathway (Kraus & Nicol, 2005; Schnupp, 2008). More broadly, it has been suggested that subcortical auditory regions should not be considered as simple relay stations, but instead as nodes in interconnected sensory, cognitive and motor networks involved in synthesising robust auditory percepts (Gruters & Groh, 2012). While the PAG is not a part of the primary the auditory pathway, it is highly interconnected with the inferior colliculus (Holmstrom, Roberts, & Portfors, 2007). As discussed above, while activity was recorded from electrodes implanted in the PAG, it is not possible to localise activity specifically to this region. Nonetheless, evidence presented here suggests functionally important classes of auditory stimuli can be differentiated in subcortical brainstem regions.

5.4.4 Limitations of Experimental Procedures

Unexpected levels of noise in the data prevented reliable analysis of individual subcategories of infant vocalisations or species-specific vocalisations. Both of these questions would be interesting to address in future studies given findings of differential cortical activity to the valence of vocalisations (e.g., Grandjean et al., 2005) and the role of subcortical brain regions in the processing species-specific vocalisations (e.g., Šuta, Kvašňák, Popelář, & Syka, 2003; Šuta, Popelář, Kvašňák, & Syka, 2007). The reasons for the high and variable levels of noise in the data remain unclear, but may be related to limitations of recording techniques, or variations in electrode placement across individuals.

The signal-to-noise ratio of data in this study was further limited by the high degree of variability in stimulus identity used. A good signal-to-noise ratio is often achieved using many repetitions of few stimuli. In the current design, there were many repetitions of stimuli from an individual category (270 trials containing infant vocalisations for example) but within each category, there was limited repetition of each stimulus (each individual stimulus was presented only six times). The aim of this design was enhance the representativeness of recorded responses to infant vocalisations in general, rather than a few specific exemplars. This type of design, however, introduces a level of variance into the data, limiting the ability to detect more subtle differences in responses. Future studies of this type should aim to include a high degree of repetition to overcome these limitations.

5.4.5 Conclusion and Future Directions

This study demonstrated early differential activity in response to infant vocalisations, compared to constructed and ecological control sounds, in LFP activity recorded from the PAG of the midbrain. This differential processing may reflect changes in responses to stimuli varying in biological salience. This notion is in line with current thinking around the functioning of the PAG, namely that it is involved in the selection of appropriate reflexive-type responses to salient environmental stimuli. Further work is required to investigate how localised these differential responses are and to assess the contributions of other nearby auditory brainstem nuclei, including the inferior colliculus. From a behavioural perspective, early differentiation of infant vocalisations from other sounds might be critically involved in the initiation of caregiving responses

CHAPTER 6: DISCUSSION

The findings from this thesis support the notion that infant vocalisations undergo differential processing in the adult listener. This is evidenced by demonstrations of sensitive responsiveness to these sounds, even in adults who are not parents. Two core features of responsiveness to infant vocalisations were assessed: the ability to respond ‘promptly’, as measured by motor performance, and the ability to respond ‘appropriately’, as measured by sensitivity to functionally important variations of these sounds. Evidence of differential processing of infant vocalisations at a neural level was found within recordings from subcortical midbrain regions.

Prompt parental responsiveness is defined in behavioural literature as rapid reactions to infant communicative cues (Bell & Ainsworth, 1972). Work in this thesis demonstrated that listening to infant cry vocalisations enhanced adults’ ability to make rapid, effortful motor movements, compared to other sounds. Compared with healthy adults, adults with depression displayed similarly enhanced motor responses after hearing infant cries, although overall performance in this group was reduced. Previous studies have demonstrated substantial impairment in psychomotor functioning in depression, including general slowing of responses to environmental stimuli (e.g., Schrijvers, Hulstijn, & Sabbe, 2008). As a result of current findings, it is hypothesised that differential processing of infant vocalisations supports prompt responsiveness through enhancing the speed and coordination of motor movements.

Appropriate responsiveness is defined as caregiving behaviour that is highly attuned to the infant’s needs (Bell & Ainsworth, 1972). This thesis showed that adults are highly sensitive to the physical parameters of infant vocalisations and interpret changes in these parameters as reflecting varying emotional states. Sensitivity to infant vocalisations was shown to vary across individuals, related to experiences of depression and musical training. Among adults with depression, those with no musical training demonstrated reduced sensitivity compared to those who did have musical training. No such differences were observed among healthy adults. Sensitivity to subtle features of infant vocalisations is therefore suggested to rely on flexible systems that are subject to modification by environmental factors. Further evidence supporting this hypothesis showed that sensitivity to infant vocalisations can be improved following a targeted training intervention.

The neural substrates supporting responsiveness to infant cues were investigated using intracranial recordings of neural activity from human subcortical brain regions. Findings provided evidence supporting the idea of early differential processing of infant vocalisations compared with physically or environmentally similar sounds. While recordings were made from intracranial DBS electrodes implanted in the PAG, the localisation of neural activity recorded to this region specifically remains untested. It is suggested that this differential early activity in a subcortical brain area may support prompt behavioural responses to infant cues.

6.1 Current Findings in the Context of Existing Theories

The findings presented here are generally consistent with the current major theories of emotional processing (Bar, 2007; LeDoux, 2000). The ability to respond or move rapidly after hearing a distressed infant is in line with the notion of fast, initial processing of environmental stimuli, especially those with survival value. Adults' sensitivity to fine-grained features of infant vocalisations is in line with the notion of a subsequent, detailed analysis of salient stimuli that ultimately helps refine behaviour.

This work can also be considered in terms of existing theoretical models of adults' responses to infant crying. In brief, models consider infant crying to be either: an aversive stimulus that should be 'terminated'; a stimulus inducing empathic responses in the listener; or a general 'motivational entity', promoting responses which are further refined by additional factors (for overview, see Murray, 1985). Evidence to date favours the motivational entity model as a comprehensive account of responsive behaviour, capturing a range of behavioural responses and allowing scope for the influence of various factors at different levels of processing.

The motivational entity model suggests that there are relatively 'hard-wired' responses to infant crying that include attentiveness and a state of arousal. Beyond this initial reaction, the nature of specific caregiving behaviours is thought to be influenced by other factors such as: acoustic features of cries, current caregiving context, listeners' motivational state and cognitive appraisal. There are a number of parallels between this model and general neuroscientific theories of emotion which comprise rapid, approximate processing followed by slower, more detailed analysis (e.g., LeDoux, 2000; Rolls, 2000).

Data presented in this thesis broadly support this type of model, demonstrating evidence for specialised processing of infant cries, the presence of a generalised motivational state and the influence of additional factors on appraisal processes. Findings from Chapter 5 demonstrated evidence for early differential neural processing of infant cries from other emotive or acoustically similar sounds. Together with previous neuroimaging work demonstrating patterns of neural activity in response to infant cries, this lends credence to the suggestion that infant cries are subject to a certain degree of specialised neural processing. Work in this thesis provided the first human evidence for early differential processing of infant vocalisations in a subcortical brain region. While further work is required to investigate the nature of this differential processing, early subcortical 'tagging' of infant vocalisations may be vital for the initiation of a heightened motivational state.

In Chapter 3, evidence was presented showing that exposure to infant cries, relative to other emotive or non-emotive sounds, can result in enhanced motor performance. These findings support the notion of a 'motivational entity', with exposure to infant cries associated with the induction of a behavioural state in which adults are capable of making more rapid, effortful and coordinated motor movements. Importantly, these findings were apparent on a simple motor task, highlighting the general nature of the effects of exposure to infant cries. This general state may promote heightened responsiveness in caregiving behaviour, although the specific content of this behaviour is likely to be affected by additional factors.

Previous work has begun to demonstrate the effects of additional factors on caregiving behaviours. Factors previously described include acoustic cues (e.g., Zeskind & Collins, 1987), contextual cues (e.g., Wood & Gustafson, 2001), parental experience (e.g., Green et al., 1987) and listener's depressive state (e.g., Stein et al., 1991). Findings described in Chapter 4 presented evidence for the interactive nature of some of these factors. In one study, previous experience of musical training was shown to enhance sensitivity to distress in infant cry vocalisations and act as a 'protective factor' against impairments to this ability associated with depression. In a second study, training focussed on enhancing sensitivity to acoustic features of infant cries was associated with improvements in individuals' sensitivity to distress in these sounds. These findings suggest that responsiveness to infant cries is a dynamic process affected by ongoing experiences. This highlights the need for further investigation into processes that govern changes in responsiveness as parents gain caregiving experience, and how these processes can become perturbed, for example in depression.

While findings presented here are informative for current models of adults' responses to infant cries, further investigation is required for a comprehensive understanding of the mechanisms guiding behavioural responses to these emotive stimuli. One under-investigated area concerns individuals' motivations for behavioural responses. Evolutionary theory suggests that caring for offspring is a primitive desire, ensuring the survival of one's own infant. More broadly, 'altruistic' theories suggest that the recognition of distress in another individual results in behavioural strategies to reduce this distress (Moss & Robson, 1968). Alternatively, it has been suggested that infant cries are simply aversive stimuli which provoke behavioural responses resulting in their termination (Hoffman, 1975).

These opposing motivations can be conceptualised as an example of 'approach-avoidance conflict' (see Aupperle & Paulus, 2010 for review). This term describes a behaviour or event that has both rewarding and punishing outcomes. Infant crying may induce both an empathic desire to approach the infant in distress as well as the aversive desire to avoid further exposure to the unpleasant sound. Measures of approach/avoidance behaviour (e.g., Aupperle, Sullivan, Melrose, Paulus, & Stein, 2011) would allow assessment of the contribution of each of these factors to behavioural responses to infant cries. Investigating these features in individuals with varying parental experience and with different types of infant vocalisations would allow assessment of the contribution of motivational factors to caregiving behaviours. Greater understanding of these mechanisms could result in development of interventions encouraging approach behaviour in parent-infant dyads with reduced levels of interaction (e.g., in depression Stein et al., 1991).

There is also little work to date systematically investigating the influence of cognitive factors on caregiving behaviour. One area that may be of particular relevance is that of attribution theory. It is plausible that the perceived cause of infant cries may impact on the provision of caregiving behaviour. One recent study demonstrated that labelling infant cries as either ‘sick’ or ‘bored’ affected neural activity in brain regions involved in processing these sounds, including the inferior frontal gyrus and amygdala (Riem, Voorthuis, Bakermans-Kranenburg, & Van Ijzendoorn, 2013). Attribution theory states that the perceived locus of control, stability and controllability of behaviours can influence the perceived cause of behaviour (Weiner, 1985). The role that these factors play in the initiation and selection of caregiving responses should be assessed within future models of caregiving behaviour.

6.2 Prompt Responsiveness to Infant Vocalisations

Infant cry vocalisations elicited enhanced motor performance in adults, compared with other environmental sounds. Performance on an objective behavioural measure of rapid and coordinated motor movements was significantly higher after listening to infant cry vocalisations, compared with listening to adult cries or bird sounds. These findings suggest that listening to infant cries induces a type of ‘high-alert’ state, improving adults’ abilities to make rapid and effortful movements. Previous studies have demonstrated that hearing infant cries results in a general desire to respond (Del Vecchio et al., 2009). Findings presented here refine this notion, suggesting specific enhancement of the core capacity of ‘prompt’ responsiveness to infant cues.

Adults with depression also demonstrated enhanced motor performance after listening to infant cries compared to other sounds. This suggests that the differential processing of infant cries resulting in enhanced responsiveness is robust to the impact of specific psychopathology. Overall, adults with depression demonstrated a reduction in motor performance compared with healthy adults. Comparable with previous findings (White et al., 1997), altered motor performance in depression was indicative of general psychomotor disturbance. Observational studies of parent-infant interactions in postnatal depression have demonstrated a reduction in the speed of parental responses to infant cues (e.g., Bettes, 1988). Findings presented here suggest that psychomotor disturbance in depression impacts on the ability to make rapid behavioural responses. .

6.3 Appropriate Responsiveness to Infant Vocalisations

Adults demonstrated sensitivity to physical parameters of infant vocalisations, evidenced by their ability to interpret changes in pitch of infant cries as indicating varying levels of distress. Previous findings have demonstrated a strong association between the pitch of infant cries and the level of perceived distress (e.g., Zeskind & Marshall, 1988). Specific impairments in sensitivity to pitch in infant vocalisations were observed in adults with depression, unless they had previous experience of musical training. Musically trained individuals with depression retained sensitivity to infant vocalisations, performing at similar levels to healthy adults. This suggests that musical training (or perhaps an associated variable) may act as a protective factor against disrupted sensitivity to infant vocalisations in depression. Previous work has suggested that mothers with depression are less sensitive to features of infant vocalisations compared with healthy mothers (Donovan et al., 1998). Other studies have shown that musicians have enhanced sensitivity to auditory stimuli, including emotional speech (e.g., Lima & Castro, 2011; Thompson et al., 2004). Work presented in this thesis demonstrated for the first time that the enhanced auditory perceptual skills of musicians may protect against the reduced sensitivity to infant vocalisations observed in depression.

The flexibility of mechanisms that may be involved in supporting appropriate responsiveness to infant vocalisations was further highlighted by the effects of a training intervention in healthy adults. An auditory perceptual training intervention was found to improve adults' sensitivity to infant vocalisations. After two 40-minute auditory training sessions on consecutive days, adults' sensitivity to distress in infant cry vocalisations was significantly improved. This effect was observed exclusively in adults who were randomly allocated to an auditory training intervention, with no effect in adults allocated to a visual training intervention. This suggests that improved sensitivity to infant vocalisations was related to the training of auditory perceptual skills. A large body of literature has previously shown the impact of short training interventions on the ability to discriminate pitch in pure tones (e.g., Amitay, Irwin, & Moore, 2006). Findings in this thesis demonstrate that established principles of auditory training can be used to enhance sensitivity to infant vocalisations. It has not yet been established whether similar interventions might be effective in adults with depression. In sum, findings indicate that appropriate responsiveness is supported by flexible systems which can be detrimentally affected by disease, but positively affected by intervention.

6.4 A Role for the Periaqueductal Gray in Adults' Responses to Infant Vocalisations

Early, differential neural activity in response to infant vocalisations was observed in recordings made from electrodes implanted in the PAG of the human midbrain in four patients. These recordings from intracranial electrodes showed evidence for differences in activity in response to infant vocalisations compared with either constructed control sounds or ecological control sounds. Recordings made from the sensory thalamus in three of these patients demonstrated no significant differences. Differential responses to infant vocalisations and constructed control sounds were observed at 50ms, while differences with ecological control sounds were observed slightly later, at around 80ms. In addition, variation in oscillatory activity between different types of sounds was observed for up to 200ms after stimulus onset in data recorded from electrodes implanted in the PAG. The specificity of these findings to the PAG compared with other midbrain structures cannot be established from the current data. However, these findings do suggest that infant vocalisations can be differentiated from other sounds at an earlier stage of auditory processing than previously examined.

It is well-established that the PAG is involved in the expression of caregiving behaviour (Numan & Woodside, 2010). This was previously evidenced by disruptions to rodent maternal behaviour by lesions of the PAG (Lonstein & Stern, 1998). Human fMRI studies have also shown that the PAG is active in response to infant vocalisations and facial expressions (e.g., Bartels & Zeki, 2004; Laurent et al., 2011). Findings presented in this thesis are the first demonstration of the temporal dynamics of PAG activity in response to infant cues. Findings of early differential activity to infant vocalisations in the PAG may be of particular importance for the rapid identification of infant vocalisations. Furthermore, the unique anatomical connectivity of the PAG as an interface for behavioural control

(Benarroch, 2012) suggests that this region may be involved in the initiation of rapid responsive caregiving behaviour.

6.5 Discussion of Methods

6.5.1 Strengths of Experimental Design

There were a number of strengths of the experimental design employed in this thesis. First, work in this thesis used a definition of responsiveness that could be readily operationalized into quantifiable measures. For instance, prompt responsiveness was operationalized as accuracy and force of motor movements and measured using a motor task. Previous studies have used a general definition of responsiveness and as a consequence measures were loosely defined (e.g., asking participants to describe how they would respond to an infant cry; Out, Pieper, Bakermans-Kranenburg, Zeskind, et al., 2010). The use of quantifiable measures with a high degree of objectivity minimises the level of subjective error during data collection. This is an advantage when compared with observational studies of parental behaviour which rely on subjective ratings by trained experimenters and typically achieve inter-rater reliability levels of a maximum of 0.7 (e.g., Stein et al., 1991).

In this work, the first large-scale database of authentic, natural vocalisations from infants, adults and domestic animals was developed. Most previous experimental work of responses to infant vocalisations used a single infant cry (e.g., Out, Pieper, Bakermans-Kranenburg, Zeskind, et al., 2010; Schuetze & Zeskind, 2001) or cries from a single infant (e.g., Frodi et al., 1978). Findings reported in this thesis are based on studies including multiple exemplars of infant vocalisations from different infants. This confers a greater degree of confidence that findings observed are related to infant vocalisations as a category of salient environmental sounds. The inclusion of authentic adult and domestic animal vocalisations in this database allowed investigation of the specificity of responses to infant vocalisations in the context of comparable environmental sounds. This allowed work here to build upon previous studies investigating infant vocalisations in isolation (e.g., Zeskind

& Marshall, 1988) or comparing responses to infant vocalisations with responses to physically matched control sounds (e.g., Lorberbaum et al., 2002). Stimuli in the database were also standardised on a number of physical parameters (duration, intensity and onset/offset amplitude envelope) making them suitable for use in both behavioural and neural studies investigating responses to affective vocalisations.

The development of a perceptual training intervention to enhance sensitivity to infant vocalisations (described in section 4.3) was grounded in a wealth of fundamental auditory science. Experimental design capitalised on previous investigation of the flexibility of auditory perceptual capacities which has specified a number of optimal parameters for perceptual learning (e.g., Hawkey et al., 2004; Wright & Sabin, 2007). These parameters have previously been shown to be effective in applied settings, for instance in targeted interventions aimed at improving children's language abilities (Bradley & Bryant, 1983; Moore et al., 2005). In the context of this large body of previous fundamental and applied research, the effects demonstrated in this thesis suggest that broad theories of perceptual learning can be applied to emotional vocalisations.

Finally, the use of intracranial recordings provided temporally sensitive data from highly localised regions of the human brain. While the opportunity to perform intracranial recordings in humans is necessarily rare, the use of this technique provides direct evidence for the functioning of a specified brain area. This provides useful additional information to that obtained using fMRI which allows investigation of networks of brain regions involved in a particular process. As a measure sensitive to changes in blood flow, fMRI is a technique that is limited in temporal resolution and the correlational nature of the association between blood flow and neuronal activity. Findings presented here using intracranial electrodes to directly record localised neural activity provides compelling evidence for the early differentiation of infant vocalisations from other sounds in the PAG.

6.5.2 Limitations of Studies

There were a number of limitations of the techniques used in this thesis. Possibly the most significant of these concerns the fact that the translation between experimental measures of responsiveness and actual responsiveness in face-to-face interactions with an infant is tentative. While it is likely that the core capacities of prompt and appropriate responsiveness demonstrated in this work are related to parental responsiveness during interactions, the extent of this link has not been tested. There are many factors that impact on parental behaviour during face-to-face interactions, including enhanced sensitivity to own infant cues (Wiesenfeld et al., 1981), the broader context of caregiving (Wood & Gustafson, 2001) and previous parenting experience (Donate-Bartfield & Passman, 1985). The effects of these factors on prompt and appropriate aspects of responsive caregiving behaviour remain a priority for future work.

The vocalisation stimuli used in these studies were obtained from naturalistic social interactions. How representative these stimuli are of the range of natural human emotional vocalisations, however, remains untested. In addition, there are some obvious omissions to the stimulus database described. It is currently limited to emotional vocalisations from six- to eight-month old infants, cries from female adults and distress sounds from domestic animals. Useful future additions would include cries from male adults as well as positive and neutral vocalisations from both male and female adults. Further to this, inclusion of comparable stimuli from individuals of different ages, developmental stages and cultures would be of much benefit to increase the versatility of this database.

Related to this point, studies included in this thesis focussed on responses to infant cry vocalisations. The reason for this specific focus was to compare findings from novel experimental procedures to the large body of previous literature concentrating on reactivity to cry stimuli. These cues are of specific interest as they often lead to the initiation of caregiving behaviour. More comprehensive understanding of adults' responses to infant vocalisations would require investigation of responses to positively and neutrally valenced stimuli. In addition, comparison of responses to artificial control sounds, such as those described in Chapter 5, would help to isolate which acoustic features of infant cries might be specifically associated with responsive behaviour.

There were also some specific limitations related to individual studies in this work. First, previous work has demonstrated changes in adults' heart rate when listening to infant cries. Although the direction of these changes in men and women, parents and non-parents remain unclear (see section 3.1.1), the majority of studies do report some change in heart rate after exposure to infant cries. In this thesis, the experiment investigating the impact of infant cries on motor performance demonstrated no change in heart rate while listening to infant cries (see section 3.3.3). This was likely related to a lack of sensitivity of methods employed to measure heart rate. Observed differences would have allowed investigation of how heart rate reactivity might be related to enhanced motor abilities following exposure to infant cries.

Measures of individual factors that might affect responsiveness to infant vocalisations were limited to current mood (depression and anxiety), previous musical training and gender. Participants with depression were recruited through community advertising processes. While diagnostic procedures ensured that participants met minimum criteria for major depressive disorder, it is likely that individuals who respond to such adverts are those who are experiencing relatively mild symptoms. Recruitment of participants more representative of the full range of symptomatology might be obtained by recruiting in a clinical setting. In addition, it is possible that other factors, such as level of social functioning or adults' own attachment experiences, might impact on responsive behaviour to infant vocalisations (Ward & Carlson, 1995). Additional self-report measures of social behaviour and childhood experiences could be readily included as variables of interest in future work.

In Chapter 4 (section 4.3), an intervention was described that improved adults' sensitivity to infant vocalisations through perceptual training. The wider impact of this training intervention was untested. Previous work has demonstrated that musical training confers generalised auditory advantages, including sensitivity to emotion in vocalisations (Thompson et al., 2004). It is plausible that the effects of the intervention used here might generalise to enhanced sensitivity to adult vocalisations, or other auditory stimuli.

Intracranial recordings allow investigation of localised neural activity. However, the extent to which activity reflects processing within a circumscribed neural area is not well-established (Kajikawa & Schroeder, 2011). Evidence presented in this thesis showed differential neural activity in response to infant vocalisations in macroelectrode recordings. While these recordings were obtained from electrodes implanted in the PAG, interpretation of findings is restricted to the functioning of human midbrain regions more generally. Furthermore, as nodes in an interconnected network, the interpretation of findings obtained from investigation of isolated brain regions is naturally limited.

6.6 Future Directions

Future research should aim to assess the extent to which experimental measures of responsiveness to infant vocalisations relate to behaviour in natural parent-infant interactions. A first stage in this investigation might assess whether and how the experience of parenting affects responsiveness to infant cues. It has previously been demonstrated that parents are highly sensitive to cues from their own infants (e.g., Cismaresco & Montagner, 1990). However, it is unclear whether parenthood is also associated with general changes in sensitivity to ‘infancy’. This could be investigated by comparing patterns of responsiveness to own infant and other infant cues in parents, alongside general responsiveness to infant cues in non-parents. In addition, future studies of parental responsiveness to infant cues should aim to include observational measures of parental behaviour during interactions with an infant. This would allow comparison of performance on an experimental task with patterns of caregiving behaviour in the same individuals. Longitudinal studies assessing how responsiveness to infant cues and parenting behaviour develops with experience would be especially informative in this regard.

A second area for future work should aim to relate disrupted performance of adults with depression on experimental measures of responsiveness to actual changes in responding to an infant. In this thesis, previous musical training was highlighted as a potential protective factor against disrupted sensitivity to infant vocalisations in adults with depression. Further investigation should assess this association in a naturalistic setting. In the context of postnatal depression, the impact of musical training on parental sensitivity during interactions with an infant should be assessed. Evidence in favour of a protective role of musical training should be followed by investigation of potential interventions. One candidate intervention strategy would use perceptual training to enhance sensitivity to infant vocalisations. The efficacy of this type of training in healthy adults was demonstrated in this thesis. Further investigation of this intervention strategy is warranted and may be of particular relevance for treatment of antenatal groups identified as at risk of postnatal depression.

Finally, future neuroimaging work should aim to investigate the temporal dynamics of activity in neural networks of the ‘parental brain’. Findings from this thesis point to a role for midbrain regions in the early differentiation of infant vocalisations. It is unknown how this early midbrain activity is related to functioning of wider neural networks supporting caregiving behaviour. Future investigation of the role of subcortical midbrain structures within a broader network of brain regions involved in sensitive responses to infant vocalisations is warranted. While it would not be possible to accurately identify regions from specific midbrain structures using EEG or MEG alone, methods for simultaneous recording from intracranial electrodes and EEG or MEG have recently been developed (Lalo et al., 2008; Litvak et al., 2010). This approach may allow investigation of the timing of neural activity in subcortical structures in human patients with implanted electrodes, as well as wider cortical activity from EEG or MEG sensors.

Future studies in humans, however, will necessarily be limited by the availability of patients with appropriately located electrodes. Comparative work in animal models would allow more comprehensive investigation of the underlying neural mechanisms. For example, recent work using single-unit recordings from auditory brainstem regions in the guinea-pig has demonstrated evidence for subcortical mechanisms of auditory scene analysis (Pressnitzer, Sayles, Micheyl, & Winter, 2008). Studies using similar methodologies would provide highly detailed information regarding the role of subcortical regions in the early processing of infant vocalisations.

With current neuroimaging methods, there is some scope to further investigate spatio-temporal patterns of neural activity in human cortical regions. For example, using MEG, recent work has identified a role for the OFC in early emotional ‘tagging’ of infant faces (Kringelbach et al., 2008; Parsons, Young, et al., 2013). Similar investigations should aim to identify the temporal dynamics of cortical activity in response to infant vocalisations. An important goal for this type of neuroimaging research should be to link patterns of neural activity more directly with behaviour during parent-infant interactions. This could be achieved in part with the use of behavioural tasks during scans, allowing correlation of neural activity with behavioural performance. Another approach would be to correlate patterns of neural activity with observational measures of sensitivity of caregiving behaviour.

6.7 Conclusion

This thesis has provided evidence for infant vocalisations as emotive, communicative cues which elicit fast, but also attuned, responses in adults. These responses are present in adults responding to unfamiliar infants, reflecting a potentially ‘universal’ responsiveness to the young of our species. Listening to infant cries was found to enhance the ability to make rapid, effortful movements. Sensitivity to functionally important parameters of infant vocalisations was found to be affected by environmental factors, including depression, previous musical experience and perceptual training. Focussed investigation of the neural basis of responsiveness to infant vocalisations highlighted a potential role for the PAG of the midbrain in the early identification of these sounds. Future translational work relating findings from clinical, experimental and neural studies would help to improve understanding of how these parental capacities develop, how they can be altered by life experiences and how they might be improved by intervention.

REFERENCES

Adolphs, R. (2003). Cognitive neuroscience of human social behaviour. *Nat Rev Neurosci, 4*(3), 165-178.

Ainsworth, M. D. (1969). Object relations, dependency, and attachment: a theoretical review of the infant-mother relationship. *Child development, 40*(4), 969-1025.

Ainsworth, M. D., Bell, S. M., & Stayton, D. J. (Eds.). (1974). *Infant-mother attachment and social development: Socialization as a product of reciprocal responsiveness to signals*. London: Cambridge University Press.

Ainsworth, M. D., Blehar, M. C., Waters, E., & Wall, S. (1978). Patterns of attachment: A psychological study of the strange situation. *Patterns of Attachment: A Psychological Study of the Strange Situation*.

Alavi, M., Dostrovsky, J., Hodaie, M., Lozano, A., & Hutchison, W. (2013). Spatial extent of β oscillatory activity in and between the subthalamic nucleus and substantia nigra pars reticulata of Parkinson's disease patients. *Experimental neurology, 245*, 60-71. doi: 10.1016/j.expneurol.2012.09.021

Alley, T. R. (1981). Head shape and the perception of cuteness. *Developmental Psychology, 17*(5), 650-654.

Amitay, S., Halliday, L., Taylor, J., Sohoglu, E., & Moore, D. R. (2010). Motivation and intelligence drive auditory perceptual learning. *PLoS One, 5*(3).

Amitay, S., Hawkey, D. J. C., & Moore, D. R. (2005). Auditory frequency discrimination learning is affected by stimulus variability. *Perception and Psychophysics, 67*(4), 691-698.

Amitay, S., Irwin, A., Hawkey, D. J. C., Cowan, J. A., & Moore, D. R. (2006). A comparison of adaptive procedures for rapid and reliable threshold assessment and training in naive listeners. *Journal of the Acoustical Society of America, 119*(3), 1616-1625.

Amitay, S., Irwin, A., & Moore, D. R. (2006). Discrimination learning induced by training with identical stimuli. *Nature Neuroscience, 9*(11), 1446-1448.

Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nat Rev Neurosci, 7*(4), 268-277.

Amoss, R. T., Martin, N. B., & Owren, M. J. (2011, 04/00/). *Physiological arousal and laughter acoustics.* Paper presented at the J Acoust Soc Am.

Arteche, A., Joormann, J., Harvey, A., Craske, M., Gotlib, I. H., Lehtonen, A., . . . Stein, A. (2011). The effects of postnatal maternal depression and anxiety on the processing of infant faces. *Journal of Affective Disorders, 133*(1-2), 197-203.

Atkinson, L., Paglia, A., Coolbear, J., Niccols, A., Parker, K. C., & Guger, S. (2000). Attachment security: a meta-analysis of maternal mental health correlates. *Clin Psychol Rev, 20*(8), 1019-1040. doi: S0272-7358(99)00023-9 [pii]

Audibert, N., Aubergé, V., & Rilliard, A. (2010). *Prosodic correlates of acted vs. spontaneous discrimination of expressive speech: a pilot study*. Paper presented at the 5th International Conference on Speech, Chicago.

Aupperle, R., & Paulus, M. (2010). Neural systems underlying approach and avoidance in anxiety disorders. *Dialogues in clinical neuroscience, 12*(4), 517-531.

Aupperle, R., Sullivan, S., Melrose, A., Paulus, M., & Stein, M. (2011). A reverse translational approach to quantify approach-avoidance conflict in humans. *Behavioural Brain Research, 225*(2), 455-463. doi: 10.1016/j.bbr.2011.08.003

Aziz-Zadeh, L., Sheng, T., & Gheytanchi, A. (2010). Common premotor regions for the perception and production of prosody and correlations with empathy and prosodic ability. *PLoS One, 5*(1). doi: 10.1371/journal.pone.0008759

Bader, A. P., & Phillips, R. D. (1999). Fathers' proficiency at recognizing their newborns by tactile cues. *Infant Behavior and Development, 22*(3), 405-409.

Bakermans-Kranenburg, M. J., Van Ijzendoorn, M. H., & Juffer, F. (2003). Less is More: Meta-Analyses of Sensitivity and Attachment Interventions in Early Childhood. *Psychological Bulletin, 129*(2), 195-215.

Bakermans-Kranenburg, M. J., van Ijzendoorn, M. H., Riem, M. M. E., Tops, M., & Alink, L. R. A. (2012). Oxytocin decreases handgrip force in reaction to infant crying in females without harsh parenting experiences. *Social Cognitive and Affective Neuroscience, 7*(8), 951-957.

Banse, R., & Scherer, K. R. (1996). Acoustic Profiles in Vocal Emotion Expression. *Journal of Personality and Social Psychology, 70*(3), 614-636.

Bänziger, T., Mortillaro, M., & Scherer, K. R. (2012). Introducing the Geneva Multimodal expression corpus for experimental research on emotion perception. *Emotion, 12*(5), 1161-1179.

Bänziger, T., & Scherer, K. R. (2005). The role of intonation in emotional expressions. *Speech Communication, 46*(3-4), 252-267.

Bar, M. (2007). The proactive brain: using analogies and associations to generate predictions. *Trends in Cognitive Sciences, 11*(7), 280-289.

Barker, P. L. (2013). *Recognition of spontaneous and acted vocalizations of emotion. (Masters dissertation).* Retrieved from Universiteit van Amsterdam, Digital Academic Repository.

Barkhuysen, P., Krahmer, E., & Swerts, M. (2007). *Cross-modal perception of emotional speech.* Paper presented at the International Congress of Phonetic Sciences, Saarbrücken.

Barrett, J., Wonch, K. E., Gonzalez, A., Ali, N., Steiner, M., Hall, G. B., & Fleming, A. S. (2012). Maternal affect and quality of parenting experiences are related to amygdala response to infant faces. *Social Neuroscience, 7*(3), 252-268.

Barrett, L. F., & Bar, M. (2009). See it with feeling: Affective predictions during object perception. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1521), 1325-1334.

Bartels, A., & Zeki, S. (2004). The neural correlates of maternal and romantic love. *NeuroImage, 21*(3), 1155-1166.

Belin, P., Fillion-Bilodeau, S., & Gosselin, F. (2008). The Montreal Affective Voices: A validated set of nonverbal affect bursts for research on auditory affective processing. *Behavior Research Methods, 40*(2), 531-539.

Belin, P., Zatorre, R. J., & Ahad, P. (2002). Human temporal-lobe response to vocal sounds. *Cognitive Brain Research, 13*(1), 17-26.

Bell, S. M., & Ainsworth, M. D. (1972). Infant crying and maternal responsiveness. *Child Dev, 43*(4), 1171-1190.

Benarroch, E. E. (2012). Periaqueductal gray: An interface for behavioral control. *Neurology, 78*(3), 210-217.

Berens, P., Keliris, G. A., Ecker, A. S., Logothetis, N. L., & Tolias, A. S. (2008). Feature Selectivity of the Gamma-Band of the Local Field Potential in Primate Primary Visual Cortex. *Frontiers in Neuroscience, 2*(2), 199-207.

Berntson, G., Boysen, S., & Cacioppo, J. (1993). Neurobehavioral organization and the cardinal principle of evaluative bivalence. *Annals of the New York Academy of Sciences, 702*, 75-102. doi: 10.1111/j.1749-6632.1993.tb17243.x

Bettes, B. A. (1988). Maternal depression and motherese: temporal and intonational features. *Child development, 59*(4), 1089-1096.

Bigelow, A. E., MacLean, K., Proctor, J., Myatt, T., Gillis, R., & Power, M. (2010). Maternal sensitivity throughout infancy: Continuity and relation to attachment security. *Infant Behavior and Development, 33*(1), 50-60.

Blair, R. C., & Karniski, W. (1993). An alternative method for significance testing of waveform difference potentials. *Psychophysiology, 30*(5), 518-524.

Bleichfeld, B., & Moely, B. E. (1984). Psychophysiological responses to an infant cry: Comparison of groups of women in different phases of the maternal cycle. *Developmental Psychology, 20*(6), 1082-1091.

Blood, A. J., & Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences, 98*(20), 11818-11823.

Bos, P., Hermans, E., Montoya, E., Ramsey, N., & van Honk, J. (2010). Testosterone administration modulates neural responses to crying infants in young females. *Psychoneuroendocrinology, 35*(1), 114-121. doi: 10.1016/j.psyneuen.2009.09.013

Bowlby, J. (1969). *Attachment and Loss, Vol 1: Attachment*. London: Hogarth Press.

Bowlby, J. (1982). Attachment and loss: Retrospect and prospect. *American Journal of Orthopsychiatry, 52*(4), 664-678.

Bradley, L., & Bryant, P. E. (1983). Categorizing sounds and learning to read - a causal connection. *Nature, 301*(5899), 419-421.

Brazelton, T. B., Koslowski, B., & Main, M. (1974). The origins of reciprocity: The early mother-infant interaction. In M. Lewis & L. Rosenblum, A. (Eds.), *The effect of the infant on its caregiver*. New York: John Wiley.

Brewster, A., Nelson, J., McCanne, T., & Lucas…, D. (1998). Gender differences in physiological reactivity to infant cries and smiles in military families. *Child abuse & Neglect, 22*(8), 775-788.

Brosch, T., Sander, D., & Scherer, K. R. (2007). That Baby Caught My Eye... Attention Capture by Infant Faces. *Emotion, 7*(3), 685-689.

Brown, P., Oliviero, A., Mazzone, P., Insola, A., Tonali, P., & Di Lazzaro, V. (2001). Dopamine dependency of oscillations between subthalamic nucleus and pallidum in Parkinson's disease. *Journal of Neuroscience, 21*(3), 1033-1038.

Brown, P., & Williams, D. (2005). Basal ganglia local field potential activity: Character and functional significance in the human. *Clinical Neurophysiology, 116*(11), 2510-2519. doi: http://dx.doi.org/10.1016/j.clinph.2005.05.009

Bruckert, L., Bestelmeyer, P., Latinus, M., Rouger, J., Charest, I., Rousselet, G. A., . . . Belin, P. (2010). Vocal Attractiveness Increases by Averaging. *Current biology, 20*(2), 116-120.

Bruner, J. S. (1975). The Ontogenesis of Speech Acts. *Journal of Child Language, 2*(1), 1-19.

Buist, A., Morse, C. A., & Durkin, S. (2003). Men's adjustment to fatherhood: implications for obstetric health care. *Journal of obstetric, gynecologic, and neonatal nursing : JOGNN / NAACOG, 32*(2), 172-180.

Burkhardt, F., Paeschke, A., Rolfes, M., Sendlmeier, W., & Weiss, B. (2005). *A database of German emotional speech.* Paper presented at the 9th European Conference on Speech Communication and Technology, Lisbon, Portugal.

Caclin, A., Brattico, E., Tervaniemi, M., Näätänen, R., Morlet, D., Giard, M. H., & McAdams, S. (2006). Separate neural processing of timbre dimensions in auditory sensory memory. *J Cogn Neurosci, 18*(12), 1959-1972.

Caria, A., de Falco, S., Venuti, P., Lee, S., Esposito, G., Rigo, P., . . . Bornstein, M. H. (2012). Species-specific response to human infant faces in the premotor cortex. *NeuroImage, 60*(2), 884-893.

Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development, 63*(4).

Cavada, C., Compañy, T., Tejedor, J., Cruz-Rizzolo, R. J., & Reinoso-Suárez, F. (2000). The anatomical connections of the macaque monkey orbitofrontal cortex. A review. *Cerebral Cortex, 10*(3), 220-242.

Chang, R. S., & Thompson, N. S. (2011). Whines, cries, and motherese: Their relative power to distract. *Journal of Social, Evolutionary, and Cultural Psychology, 5*(2), 131-141.

Chaumon, M., Kveraga, K., Barrett, L., & Bar, M. (2013). Visual Predictions in the Orbitofrontal Cortex Rely on Associative Content. *Cerebral Cortex*. doi: 10.1093/cercor/bht146

Cismaresco, A., & Montagner, H. (1990). Mothers' discrimination of their neonates' cry in relation to cry acoustics: The first week of life. *Early Child Development and Care, 65*, 3-11.

Cohn, J. F., & Tronick, E. Z. (1983). Three-month-old infants' reaction to simulated maternal depression. *Child development, 54*(1), 185-193.

Cooper, P. J., Tomlinson, M., Swartz, L., Landman, M., Molteno, C., Stein, A., . . . Murray, L. (2009). Improving quality of mother-infant relationship and infant attachment in socioeconomically deprived community in South Africa: randomised controlled trial. *BMJ (Clinical research ed.), 338*.

Cornell, D. G., Suarez, R., & Berent, S. (1984). Psychomotor retardation in melancholic and nonmelancholic depression: cognitive and motor components. *J Abnorm Psychol, 93*(2), 150-157.

Corter, C. M., & Fleming, A. S. (1990). Maternal responsiveness in humans: Emotional, cognitive and biological factors *Advances in the Study of Behavior* (Vol. 19, pp. 83-136).

Cox, J. L., Chapman, G., Murray, D., & Jones, P. (1996). Validation of the Edinburgh postnatal depression scale (EPDS) in non-postnatal women. *Journal of Affective Disorders, 39*(3), 185-189.

Cox, J. L., Holden, J. M., & Sagovsky, R. (1987). Detection of Postnatal Depression: Development of the 10-item Edinburgh Postnatal Depression scale. *British Journal of Psychiatry, 150*(JUNE), 782-786.

Craske, M., Kircanski, K., Zelikowsky, M., Mystkowski, J., Chowdhury, N., & Baker, A. (2008). Optimizing inhibitory learning during exposure therapy. *Behaviour research and therapy, 46*(1), 5-27. doi: 10.1016/j.brat.2007.10.003

Cunningham, M. R., Barbee, A. P., & Pike, C. L. (1990). What Do Women Want? Facialmetric Assessment of Multiple Motives in the Perception of Male Facial Physical Attractiveness. *Journal of Personality and Social Psychology, 59*(1), 61-72.

Darwin, C. (1872). *The Expression of the Emotions in Man and Animals*: John Murray, London.

De Pisapia, N., Bornstein, M. H., Rigo, P., Esposito, G., De Falco, S., & Venuti, P. (2013). Sex differences in directional brain responses to infant hunger cries. *NeuroReport, 24*(3), 142-146.

De Wolff, M. S., & Van Ijzendoorn, M. H. (1997). Sensitivity and Attachment: A Meta-Analysis on Parental Antecedents of Infant Attachment. *Child development, 68*(4), 571-591.

Deater-Deckard, K., Pickering, K., Dunn, J. F., & Golding, J. (1998). Family structure and depressive symptoms in men preceding and following the birth of a child. The Avon Longitudinal Study of Pregnancy and Childhood Study Team. *Am J Psychiatry, 155*(6), 818-823.

Del Vecchio, T., Walter, A., & O'Leary, S. G. (2009). Affective and physiological factors predicting maternal response to infant crying. *Infant Behavior and Development, 32*(1), 117-122.

Delhommeau, K., Micheyl, C., Jouvent, R., & Collet, L. (2002). Transfer of learning across durations and ears in auditory frequency discrimination. *Perception and Psychophysics, 64*(3), 426-436.

Demany, L. (1985). Perceptual learning in frequency discrimination. *Journal of the Acoustical Society of America, 78*(3), 1118-1120.

Denham, S. A., Blair, K. A., DeMulder, E., Levitas, J., Sawyer, K., Auerbach–Major, S., & Queenan, P. (2003). Preschool Emotional Competence: Pathway to Social Competence? *Child development, 74*(1), 238-256. doi: 10.1111/1467-8624.00533

Dessureau, B. K., Kurowski, C. O., & Thompson, N. S. (1998). A reassessment of the role of pitch and duration in adults' responses to infant crying. *Infant Behavior & Development, 21*(2), 367-371.

Dietrich, S., Hertrich, I., Alter, K., Ischebeck, A., & Ackermann, H. (2007). Semiotic aspects of human nonverbal vocalizations: A functional imaging study. *NeuroReport, 18*(18), 1891-1894.

Donate-Bartfield, E., & Passman, R. H. (1985). Attentiveness of mothers and fathers to their baby's cries. *Infant Behavior and Development, 8*(4), 385-393.

Donovan, W. L., Leavitt, L. A., & Walsh, R. O. (1998). Conflict and depression predict maternal sensitivity to infant cries. *Infant Behavior and Development, 21*(3), 505-517.

Drolet, M., Schubotz, R. I., & Fischer, J. (2012). Authenticity affects the recognition of emotions in speech: Behavioral and fMRI evidence. *Cognitive, Affective and Behavioral Neuroscience, 12*(1), 140-150.

Dujardin, E., & Jürgens, U. (2005). Afferents of vocalization-controlling periaqueductal regions in the squirrel monkey. *Brain Research, 1034*(1-2), 114-131.

Eisenberg, N., Zhou, Q., Losoya, S. H., Fabes, R. A., Shepard, S. A., Murphy, B. C., . . . Cumberland, A. (2003). The Relations of Parenting, Effortful Control, and Ego Control to Children's Emotional Expressivity. *Child development, 74*(3), 875-895. doi: 10.1111/1467-8624.00573

Ekman, P. (1993). Facial expression and emotion. *American Psychologist, 48*(4), 384-392.

Emerson, C. S., Harrison, D. W., & Everhart, D. E. (1999). Investigation of receptive affective prosodic ability in school-aged boys with and without depression. *Neuropsychiatry, Neuropsychology and Behavioral Neurology, 12*(2), 102-109.

Engberg, I. S., & Hansen, A. V. (1996). Documentation of the danish emotional speech database DES. *Internal AAU report, Center for Person Kommunikation, Denmark.*

Ethofer, T., Van De Ville, D., Scherer, K. R., & Vuilleumier, P. (2009). Decoding of Emotional Information in Voice-Sensitive Cortices. *Current biology, 19*(12), 1028-1033.

Fanselow, M. S. (1994). Neural organization of the defensive behavior system responsible for fear. *Psychonomic Bulletin & Review, 1*(4), 429-438. doi: 10.3758/bf03210947

Fecteau, S., Armony, J. L., Joanette, Y., & Belin, P. (2005). Judgment of emotional nonlinguistic vocalizations: Age-related differences. *Applied Neuropsychology, 12*(1), 40-48.

Fecteau, S., Belin, P., Joanette, Y., & Armony, J. L. (2007). Amygdala responses to nonlinguistic emotional vocalizations. *NeuroImage, 36*(2), 480-487.

Feinberg, D. R., Jones, B. C., Law Smith, M. J., Moore, F. R., DeBruine, L. M., Cornwell, R. E., . . . Perrett, D. I. (2006). Menstrual cycle, trait estrogen level, and masculinity preferences in the human voice. *Hormones and Behavior, 49*(2), 215-222.

Field, T. (1995). Infants of depressed mothers. *Infant Behavior and Development, 18*(1), 1-13.

Field, T., Morrow, C., & Adlestein, D. (1993). Depressed mothers' perceptions of infant behavior. *Infant Behavior and Development, 16*(1), 99-108.

Field, T., Woodson, R., Greenberg, R., & Cohen, D. (1982). Discrimination and imitation of facial expressions by neonates. *Science, 218*(4568), 179-181.

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (2002). Structured clinical interview for DSM-IV-TR Axis i Disorders, research version, non-patient edition.

(SCID-I/NP). New York: Biometrics Research, New York State Psychiatric Institute.

Frith, C. D. (2007). The social brain? *Philosophical Transactions of the Royal Society B: Biological Sciences, 362*(1480), 671-678.

Frith, C. D., & Frith, U. (2007). Social Cognition in Humans. *Current biology, 17*(16).

Frith, U., & Frith, C. (2010). The social brain: Allowing humans to boldly go where no other species has been. *Philosophical Transactions of the Royal Society B: Biological Sciences, 365*(1537), 165-175.

Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society B: Biological Sciences, 358*(1431), 459-473.

Frodi, A. M., & Lamb, M. E. (1980). Child abusers' responses to infant smiles and cries. *Child development, 51*(1), 238-241.

Frodi, A. M., Lamb, M. E., Leavitt, L. A., & Donovan, W. L. (1978). Fathers' and mothers' responses to infant smiles and cries. *Infant Behavior and Development, 1*(2), 187-198.

Fullard, W., & Reiling, A. M. (1976). An Investigation of Lorenz's "Babyness". *Child development, 47*(4), 1191-1193. doi: 10.2307/1128462

Furedy, J. J., Fleming, A. S., Ruble, D., Scher, H., Daly, J., Day, D., & Loewen, R. (1989). Sex differences in small-magnitude heart-rate responses to sexual and infant-related stimuli: A psychophysiological approach. *Physiology and Behavior, 46*(5), 903-905.

Gallagher, H. L., & Frith, C. D. (2003). Functional imaging of 'theory of mind'. *Trends Cogn Sci, 7*(2), 77-83. doi: S1364661302000256 [pii]

Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (1999). Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects. *Nat Neurosci, 2*(6), 568-573. doi: 10.1038/9224

Gavin, N. I., Gaynes, B. N., Lohr, K. N., Meltzer-Brody, S., Gartlehner, G., & Swinson, T. (2005). Perinatal depression: A systematic review of prevalence and incidence. *Obstetrics and Gynecology, 106*(5 I), 1071-1083.

Glocker, M. L., Langleben, D. D., Ruparel, K., Loughead, J. W., Gur, R. C., & Sachser, N. (2009). Baby schema in infant faces induces cuteness perception and motivation for caretaking in adults. *Ethology, 115*(3), 257-263.

Gollan, J. K., Pane, H. T., McCloskey, M. S., & Coccaro, E. F. (2008). Identifying differences in biased affective information processing in major depression. *Psychiatry Research, 159*(1-2), 18-24.

Golub, H. L., & Corwin, M. J. (1985). A physioacoustic model of the infant cry. In B. M. Lester & C. F. Boukydis (Eds.), *Infant crying: Theoretical and Research Perspectives* (pp. 59-82). New York: Plenum Press.

Gotlib, I. H., Krasnoperova, E., Yue, D. N., & Joormann, J. (2004). Attentional Biases for Negative Interpersonal Stimuli in Clinical Depression. *Journal of Abnormal Psychology, 113*(1), 127-135.

Grandjean, D., Sander, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., & Vuilleumier, P. (2005). The voices of wrath: Brain responses to angry prosody in meaningless speech. *Nature Neuroscience, 8*(2), 145-146.

Green, A., Wang, S., Stein, J., Pereira, E., Kringelbach, M. L., Liu, X., . . . Aziz, T. (2009). Neural signatures in patients with neuropathic pain. *Neurology, 72*(6), 569-571.

Green, J. A., & Gustafson, G. E. (1983). Individual recognition of human infants on the basis of cries alone. *Developmental Psychobiology, 16*(6), 485-504.

Green, J. A., Jones, L. E., & Gustafson, G. E. (1987). Perception of Cries by Parents and Nonparents: Relation to Cry Acoustics. *Developmental Psychology, 23*(3), 370-382.

Grimault, N., Micheyl, C., Carlyon, R. P., & Collet, L. (2002). Evidence for two pitch encoding mechanisms using a selective auditory training paradigm. *Perception and Psychophysics, 64*(2), 189-197.

Gruters, K. G., & Groh, J. M. (2012). Sounds and beyond: multisensory and other non-auditory signals in the inferior colliculus. *Frontiers in Neural Circuits, 6*. doi: 10.3389/fncir.2012.00096

Gur, R. C., Erwin, R. J., Gur, R. E., Zwil, A. S., Heimberg, C., & Kraemer, H. C. (1992). Facial emotion discrimination: II. Behavioral findings in depression. *Psychiatry Research, 42*(3), 241-251.

Gustafson, G. E., & Green, J. A. (1989). On the importance of fundamental frequency and other acoustic features in cry perception and infant development. *Child development, 60*(4), 772-780.

Gustafson, G. E., Wood, R. M., & Green, J. A. (2000). Can we hear the causes of infants' crying? In R. G. Barr, B. Hopkins & J. A. Green (Eds.), *Crying as a Sign, a Symptom, and a Signal.* (pp. 8-22). Cambridge: Cambridge University Press.

Gustafsson, E., Levréro, F., Reby, D., & Mathevon, N. (2013). Fathers are just as good as mothers at recognizing the cries of their baby. *Nature Communications, 4*, 1698. doi: 10.1038/ncomms2713

Guthrie, D., & Buchwald, J. S. (1991). Significance testing of difference potentials. *Psychophysiology, 28*(2), 240-244.

Hawkey, D. J. C., Amitay, S., & Moore, D. R. (2004). Early and rapid perceptual learning. *Nature Neuroscience, 7*(10), 1055-1056.

Hildebrandt, K. A., & Fitzgerald, H. E. (1978). Adults' responses to infants varying in perceived cuteness. *Behavioural Processes, 3*(2), 159-172.

Hillebrand, A., & Barnes, G. R. (2002). A Quantitative Assessment of the Sensitivity of Whole-Head MEG to Activity in the Adult Human Cortex. *NeuroImage, 16*(3, Part A), 638-650. doi: 10.1006/nimg.2002.1102

Hoffman, M. L. (1975). Developmental synthesis of affect and cognition and its implications for altruistic motivation. *Developmental Psychology, 11*(5), 607.

Holmstrom, L., Roberts, P. D., & Portfors, C. V. (2007). Responses to social vocalizations in the inferior colliculus of the mustached bat are influenced by secondary tuning curves. *Journal of Neurophysiology, 98*(6), 3461-3472.

Holstege, G., Bandler, R., & Saper, C. (1996). The emotional motor system. *Progress in brain research, 107*, 3-6.

Huttar, G. L. (1968). Relations between prosodic variables and emotions in normal American English utterances. *Journal of Speech and Hearing Research, 11*(3), 481-487.

Irvine, D. R. F., Martin, R. L., Klimkeit, E., & Smith, R. (2000). Specificity of perceptual learning in a frequencey discrimination task. *Journal of the Acoustical Society of America, 108*(6), 2964-2968.

Izard, C. E. (1971). *The face of emotion*. New York: Appleton-Century-Crofts.

Joosen, K. J., Mesman, J., Bakermans-Kranenburg, M. J., Pieper, S., Zeskind, P. S., & Van Ijzendoorn, M. H. (2012). Physiological Reactivity to Infant Crying and Observed Maternal Sensitivity. *Infancy, 18*. doi: 10.1111/j.1532-7078.2012.00122.x

Juffer, F., Bakermans-Kranenburg, M. J., & van Ijzendoorn, M. H. (2007). *Promoting positive parenting: An attachment-based intervention*. New York: Taylor Francis Group.

Jürgens, R., Hammerschmidt, K., & Fischer, J. (2011). Authentic and play-acted vocal emotion expressions reveal acoustic differences. *Frontiers in Psychology, 2*(JUL).

Jürgens, U. (1994). The role of the periaqueductal grey in vocal behaviour. *Behavioural Brain Research, 62*(2), 107-117.

Juslin, P. N., & Laukka, P. (2001). Impact of Intended Emotion Intensity on Cue Utilization and Decoding Accuracy in Vocal Expression of Emotion. *Emotion, 1*(4), 381-412.

Juslin, P. N., & Laukka, P. (2003). Communication of Emotions in Vocal Expression and Music Performance: Different Channels, Same Code? *Psychological Bulletin, 129*(5), 770-814.

Kaitz, M., Lapidot, P., Bronner, R., & Eidelman, A. (1992). Parturient women can recognize their infants by touch. *Developmental Psychology, 28*(1), 35-39.

Kaitz, M., Shiri, S., Danziger, S., Hershko, Z., & Eidelman, A. (1994). Fathers can also recognize their newborns by touch. *Infant Behavior and Development, 17*, 205-207.

Kajikawa, Y., & Schroeder, C. (2011). How local is the local field potential? *Neuron, 72*(5), 847-858.

Karni, A., & Sagi, D. (1993). The time course of learning a visual skill. *Nature, 365*(6443), 250-252.

Katzner, S., Nauhaus, I., Benucci, A., Bonin, V., Ringach, D. L., & Carandini, M. (2009). Local Origin of Field Potentials in Visual Cortex. *Neuron, 61*(1), 35-41.

Kim, J.-W., Kim, S.-E., Kim, J.-J., Jeong, B., Park, C.-H., Son, A. R., . . . Ki, S. W. (2009). Compassionate attitude towards others' suffering activates the mesolimbic neural system. *Neuropsychologia, 47*(10), 2073-2081.

Kim, P., Feldman, R., Mayes, L. C., Eicher, V., Thompson, N., Leckman, J. F., & Swain, J. E. (2011). Breastfeeding, brain activation to own infant cry, and maternal sensitivity. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 52*(8), 907-915.

Konner, M. (2010). *The Evolution of Childhood: Relationships, Emotion, Mind*: Harvard University Press.

Kraskov, A., Quiroga, R. Q., Reddy, L., Fried, I., & Koch, C. (2007). Local field potentials and spikes in the human medial temporal lobe are selective to image category. *J Cogn Neurosci, 19*(3), 479-492.

Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews Neuroscience, 11*(8), 599-605.

Kraus, N., & Nicol, T. (2005). Brainstem origins for cortical 'what' and 'where' pathways in the auditory system. *Trends Neurosci, 28*(4), 176-181. doi: 10.1016/j.tins.2005.02.003

Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: A review. *Biological Psychology, 84*(3), 394-421.

Kreiman, G., Hung, C. P., Kraskov, A., Quiroga, R. Q., Poggio, T., & DiCarlo, J. J. (2006). Object selectivity of local field potentials and spikes in the macaque inferior temporal cortex. *Neuron, 49*(3), 433-445.

Kringelbach, M. L. (2005). The human orbitofrontal cortex: Linking reward to hedonic experience. *Nature Reviews Neuroscience, 6*(9), 691-702.

Kringelbach, M. L., Lehtonen, A., Squire, S., Harvey, A. G., Craske, M. G., Holliday, I. E., . . . Stein, A. (2008). A specific and rapid neural signature for parental instinct. *PLoS One, 3*(2), e1664. doi: 10.1371/journal.pone.0001664

Kringelbach, M. L., Pereira, E. A. C., Green, A. L., Owen, S. L. F., & Aziz, T. Z. (2009). Deep brain stimulation for chronic pain. *Journal of Pain Management, 3*, 301-314.

Kringelbach, M. L., & Rolls, E. T. (2003). Neural correlates of rapid reversal learning in a simple model of human social interaction. *NeuroImage, 20*(2), 1371-1383.

Kruger, A. C., & Konner, M. (2010). Who Responds to Crying?: Maternal Care and Allocare among the !Kung. *Human Nature, 21*(3), 309-329.

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience, 5*(11), 831-843.

Kuhl, P. K., & Meltzoff, A. N. (1982). The bimodal perception of speech in infancy. *Science, 218*(4577), 1138-1140.

Lalo, E., Thobois, S., Sharott, A., Polo, G., Mertens, P., Pogosyan, A., & Brown, P. (2008). Patterns of bidirectional communication between cortex and basal ganglia during movement in patients with Parkinson disease. *J Neurosci, 28*(12), 3008-3016. doi: 10.1523/jneurosci.5295-07.2008

Lamb, M. E. (1975). Fathers: forgotten contributors to child development. *Human Development, 18*(4), 245-266.

Lamb, M. E. (1977). Father-infant and mother-infant interaction in the first year of life. *Child development*, 167-181.

Lamb, M. E. (2012). Mothers, Fathers, Families, and Circumstances: Factors Affecting Children's Adjustment. *Applied Developmental Science, 16*(2), 98-111.

Lamers, F., Van Oppen, P., Comijs, H. C., Smit, J. H., Spinhoven, P., Van Balkom, A. J. L. M., . . . Penninx, B. W. J. H. (2011). Comorbidity patterns of anxiety and depressive disorders in a large cohort study: The Netherlands Study of Depression and Anxiety (NESDA). *Journal of Clinical Psychiatry, 72*(3), 342-348.

Lartillot, O., & Toiviainen, P. (2007). *MIR in Matlab (II): a toolbox for musical feature extraction from audio.* Paper presented at the International Conference on Music Information Retrieval, Vienna.

Laukka, P., Neiberg, D., Forsell, M., Karlsson, I., & Elenius, K. (2011). Expression of affect in spontaneous speech: Acoustic correlates and automatic detection of irritation and resignation. *Computer Speech and Language, 25*(1), 84-104.

Laurent, H. K., & Ablow, J. C. (2012a). A cry in the dark: Depressed mothers show reduced neural activation to their own infant's cry. *Social Cognitive and Affective Neuroscience, 7*(2), 125-134.

Laurent, H. K., & Ablow, J. C. (2012b). The missing link: Mothers' neural response to infant cry related to infant attachment behaviors. *Infant Behavior and Development, 35*(4), 761-772.

Laurent, H. K., Stevens, A., & Ablow, J. C. (2011). Neural correlates of hypothalamic-pituitary-adrenal regulation of mothers with their infants. *Biological Psychiatry, 70*(9), 826-832.

Lawrie, S. M., MacHale, S. M., Cavanagh, J. T., O'Carroll, R. E., & Goodwin, G. M. (2000). The difference in patterns of motor and cognitive function in chronic fatigue syndrome and severe depressive illness. *Psychol Med, 30*(2), 433-442.

Lederman, D. (2010). Estimation of Infants' Cry Fundamental Frequency using a Modified SIFT algorithm. *arXiv preprint arXiv:1009.2796*.

LeDoux, J. E. (2000). Emotion Circuits in the Brain. *Annual Review of Neuroscience, 23*(1), 155-184. doi: 10.1146/annurev.neuro.23.1.155

LeDoux, J. E. (2012). Evolution of human emotion. A view through fear. *Progress in brain research, 195*, 431-442.

Leger, D. W., Thompson, R. A., Merritt, J. A., & Benz, J. J. (1996). Adult Perception of Emotion Intensity in Human Infant Cries: Effects of Infant Age and Cry Acoustics. *Child development, 67*(6), 3238-3249.

Leppänen, J. M., Milders, M., Bell, J. S., Terriere, E., & Hietanen, J. K. (2004). Depression biases the recognition of emotionally neutral faces. *Psychiatry Research, 128*(2), 123-133.

Lester, B. M., Boukydis, C. F. Z., Garcia-Coll, C. T., Peucker, M., McGrath, M. M., Vohr, B. R., . . . Oh, W. (1995). Developmental outcome as a function of the goodness of fit between the infant's cry characteristics and the mother's perception of her infant's cry. *Pediatrics, 95*(4), 516-521.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America, 49*(2).

Lima, C. F., & Castro, S. L. (2011). Speaking to the trained ear: musical expertise enhances the recognition of emotions in speech prosody. *Emotion, 11*(5), 1021-1031.

Lima, C. F., Castro, S. L., & Scott, S. K. (2013). When voices get emotional: A corpus of nonverbal vocalizations for research on emotion processing. *Behav Res Methods*. doi: 10.3758/s13428-013-0324-3

Linnman, C., Moulton, E. A., Barmettler, G., Becerra, L., & Borsook, D. (2012). Neuroimaging of the periaqueductal gray: State of the field. *NeuroImage, 60*(1), 505-522.

Litvak, V., Eusebio, A., Jha, A., Oostenveld, R., Barnes, G. R., Penny, W. D., . . . Friston, K. J. (2010). Optimized beamforming for simultaneous MEG and intracranial local field potential recordings in deep brain stimulation patients. *NeuroImage, 50*(4), 1578-1588.

Lobmaier, J. S., Sprengelmeyer, R., Wiffen, B., & Perrett, D. I. (2010). Female and male responses to cuteness, age and emotion in infant faces. *Evolution and Human Behavior, 31*(1), 16-21.

Lonstein, J. S., & Stern, J. M. (1997). Role of the midbrain periaqueductal gray in maternal nurturance and aggression: C-fos and electrolytic lesion studies in lactating rats. *Journal of Neuroscience, 17*(9), 3364-3378.

Lonstein, J. S., & Stern, J. M. (1998). Site and behavioral specificity of periaqueductal gray lesions on postpartum sexual, maternal, and aggressive behaviors in rats. *Brain Research, 804*(1), 21-35.

Lorberbaum, J. P., Newman, J. D., Horwitz, A. R., Dubno, J. R., Lydiard, R. B., Hamner, M. B., . . . George, M. S. (2002). A potential role for thalamocingulate circuitry in human maternal behavior. *Biological Psychiatry, 51*(6), 431-445.

Lorenz, K. (1943). Die angeborenen Formen Möglicher Erfahrung. [Innate forms of potential experience]. *Zeitschrift für Tierpsychologie, 5*, 235-519.

Lorenz, K. (1950). The comparative method in studying innate behaviour patterns. *Symp. Soc. Exp. Biol., 4*, 221-268.

Magne, C., Schon, D., & Besson, M. (2006). Musician children detect pitch violations in both music and language better than nonmusician children: behavioral and electrophysiological approaches. *J Cogn Neurosci, 18*(2), 199-211. doi: 10.1162/089892906775783660

Mampe, B., Friederici, A. D., Christophe, A., & Wermke, K. (2009). Newborns' cry melody is shaped by their native language. *Current biology, 19*(23), 1994-1997.

Matthey, S., Barnett, B., Kavanagh, D. J., & Howie, P. (2001). Validation of the Edinburgh Postnatal Depression Scale for men, and comparison of item endorsement with their partners. *Journal of Affective Disorders, 64*(2-3), 175-184.

McCanne, T. R., & Hagstrom, A. H. (1996). Physiological hyperreactivity to stressors in physical child abusers and individuals at risk for being physically abusive. *Aggression and Violent Behavior, 1*(4), 345-358. doi: http://dx.doi.org/10.1016/S1359-1789(96)00004-3

McComb, K., Taylor, A. M., Wilson, C., & Charlton, B. D. (2009). The cry embedded within the purr. *Current biology, 19*(13), R507-R508.

Meltzer, J. A., Zaveri, H. P., Goncharova, I. I., Distasio, M. M., Papademetris, X., Spencer, S. S., . . . Constable, R. T. (2008). Effects of working memory load on oscillatory power in human intracranial EEG. *Cerebral Cortex, 18*(8), 1843-1855.

Meltzoff, A. N., & Moore, M. K. (1977). Imitation of Facial and Manual Gestures by Human Neonates. *Science, 198*(4312), 75-78. doi: 198/4312/75 [pii] 10.1126/science.198.4312.75

Mesman, J., van Ijzendoorn, M. H., & Bakermans-Kranenburg, M. J. (2009). The many faces of the Still-Face Paradigm: A review and meta-analysis. *Developmental Review, 29*(2), 120-162.

Messinger, D. S. (2002). Positive and negative: Infant facial expressions and emotions. *Current Directions in Psychological Science, 11*(1), 1-6.

Meyer, M., Baumann, S., Wildgruber, D., & Alter, K. (2007). How the brain laughs. Comparative evidence from behavioral, electrophysiological and neuroimaging studies in human and monkey. *Behavioural Brain Research, 182*(2), 245-260.

Micheyl, C., Delhommeau, K., Perrot, X., & Oxenham, A. J. (2006). Influence of musical and psychoacoustical training on pitch discrimination. *Hearing Research, 219*(1-2), 36-47.

Milgrom, J., Westley, D. T., & Gemmill, A. W. (2004). The mediating role of maternal responsiveness in some longer term effects of postnatal depression on infant development. *Infant Behavior and Development, 27*(4), 443-454. doi: 10.1016/j.infbeh.2004.03.003

Miranda-Paiva, C. M., Canteras, N. S., Sukikara, M. H., Nasello, A. G., Mackowiak, II, & Felicio, L. F. (2007). Periaqueductal gray cholecystokinin infusions block morphine-induced disruption of maternal behavior. *Peptides, 28*(3), 657-662. doi: 10.1016/j.peptides.2006.11.005

Mitchell, R. L. C., Elliott, R., Barry, M., Cruttenden, A., & Woodruff, P. W. R. (2003). The neural response to emotional prosody, as revealed by functional magnetic resonance imaging. *Neuropsychologia, 41*(10), 1410-1421.

Miyazono, H., Glasberg, B. R., & Moore, B. C. J. (2010). Perceptual learning of fundamental frequency discrimination: Effects of fundamental frequency, harmonic number, and component phase. *Journal of the Acoustical Society of America, 128*(6), 3649-3657.

Mobbs, D., Petrovic, P., Marchant, J. L., Hassabis, D., Weiskopf, N., Seymour, B., . . . Frith, C. D. (2007). When fear is near: threat imminence elicits prefrontal-periaqueductal gray shifts in humans. *Science, 317*(5841), 1079-1083. doi: 10.1126/science.1144298

Montoya, J. L., Landi, N., Kober, H., Worhunsky, P. D., Rutherford, H. J. V., Mencl, W. E., . . . Potenza, M. N. (2012). Regional brain responses in nulliparous women to emotional infant stimuli. *PLoS One, 7*(5).

Moore, C., & Corkum, V. (1994). Social Understanding at the End of the First Year of Life. *Developmental Review, 14*(4), 349-372.

Moore, D. R., Rosenberg, J. F., & Coleman, J. S. (2005). Discrimination training of phonemic contrasts enhances phonological processing in mainstream school children. *Brain and Language, 94*(1), 72-85.

Morris, J. S., Scott, S. K., & Dolan, R. J. (1999). Saying it with feeling: Neural responses to emotional vocalizations. *Neuropsychologia, 37*(10), 1155-1163.

Morsbach, G., McCulloch, M., & Clark, A. (1986). Infant crying as a potential stressor concerning mothers' concentration ability. *Psychologia: An International Journal of Psychology in the Orient, 29*(1), 18-20.

Moss, J., & Robson, K. (1968). *The role of protest behavior in the development of mother-infant attachment.* Paper presented at the American Psychological Association, San Fransisco.

Muller, E., Hollien, H., & Murry, T. (1974). Perceptual responses to infant crying: identification of cry types. *Journal of Child Language, 1*(1), 89-95.

Murray, A. D. (1979). Infant crying as an elicitor of parental behavior: An examination of two models. *Psychological Bulletin, 86*(1), 191-215.

Murray, A. D. (1985). Aversiveness is in the mind of the beholder: Perception of infant crying by adults. In B. M. Lester & C. F. Boukydis (Eds.), *Infant Crying: Theoretical and Research Perspectives* (pp. 217-239). New York: Plenum Press.

Murray, I. R., & Arnott, J. L. (1993). Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. *Journal of the Acoustical Society of America, 93*(2), 1097-1108.

Murray, L., Cooper, P. J., Wilson, A., & Romaniuk, H. (2003). Controlled trial of the short- and long-term effect of psychological treatment of post-partum depression. 2. Impact on the mother-child relationship and child outcome. *British Journal of Psychiatry, 182*(MAY), 420-427.

Murray, L., Fiori-Cowley, A., Hooper, R., & Cooper, P. (1996). The impact of postnatal depression and associated adversity on early mother-infant interactions and later infant outcome. *Child Dev, 67*(5), 2512-2526.

Murray, L., Hipwell, A., Hooper, R., Stein, A., & Cooper, P. (1996). The cognitive development of 5-year-old children of postnatally depressed mothers. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 37*(8), 927-935.

Murray, L., Kempton, C., Woolgar, M., & Hooper, R. (1993). Depressed mothers' speech to their infants and its relation to infant gender and cognitive development. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 34*(7), 1083-1101.

Murray, L., Sinclair, D., Cooper, P., Ducournau, P., Turner, P., & Stein, A. (1999). The socioemotional development of 5-year-old children of postnatally depressed mothers. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 40*(8), 1259-1271.

Murray, L., & Trevarthen, C. (1985). Emotional regulation of interactions between two-month-olds and their mothers. In T. M. Field & N. A. Fox (Eds.), *Social perception in infants* (pp. 177-197). Norwood, NJ: Ablex.

Musacchia, G., Sams, M., Skoe, E., & Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proceedings of the National Academy of Sciences of the United States of America, 104*(40), 15894-15898.

Musser, E. D., Kaiser-Laurent, H., & Ablow, J. C. (2012). The neural correlates of maternal sensitivity: An fMRI study. *Developmental Cognitive Neuroscience, 2*(4), 428-436.

Naranjo, C., Kornreich, C., Campanella, S., Noël, X., Vandriette, Y., Gillain, B., . . . Constant, E. (2011). Major depression is associated with impaired processing of emotion in music as well as in facial and vocal stimuli. *Journal of Affective Disorders, 128*(3), 243-251.

Nauhaus, I., Busse, L., Carandini, M., & Ringach, D. L. (2009). Stimulus contrast modulates functional connectivity in visual cortex. *Nature Neuroscience, 12*(1), 70-76.

Newman, J. D. (2007). Neural circuits underlying crying and cry responding in mammals. *Behavioural Brain Research, 182*(2), 155-165.

Newman, M. O., Zuellig, A. R., Kachin, K. E., Constantino, M. J., Przeworski, A., Erickson, T., & Cashman-McGrath, L. (2002). Preliminary reliability and validity of the generalized anxiety disorder questionnaire-IV: A revised self-report diagnostic measure of generalized anxiety disorder. *Behavior Therapy, 33*(2), 215-233.

Nitschke, J. B., Nelson, E. E., Rusch, B. D., Fox, A. S., Oakes, T. R., & Davidson, R. J. (2004). Orbitofrontal cortex tracks positive mood in mothers viewing pictures of their newborn infants. *NeuroImage, 21*(2), 583-592.

Noriuchi, M., Kikuchi, Y., & Senoo, A. (2008). The Functional Neuroanatomy of Maternal Love: Mother's Response to Infant's Attachment Behaviors. *Biol Psychiatry, 63*(4), 415-423.

Numan, M. J., & Insel, T. R. (2003). *The neurobiology of parental behavior*. New York: Springer.

Numan, M. J., & Woodside, B. (2010). Maternity: Neural Mechanisms, Motivational Processes, and Physiological Adaptations. *Behavioral Neuroscience, 124*(6), 715-741.

Nwokah, E. E., Hsu, H. C., Dobrowolska, O., & Fogel, A. (1994). The development of laughter in mother-infant communication: Timing parameters and temporal sequences. *Infant Behavior and Development, 17*(1), 23-35.

Nylen, K. J., Moran, T. E., Franklin, C. L., & O'Hara, M. W. (2006). Maternal depression: A review of relevant treatment approaches for mothers and infants. *Infant Mental Health Journal, 27*(4), 327-343.

O'Hara, M. W., & Mc Cabe, J. E. (2013) Postpartum depression: Current status and future directions. *Vol. 9* (pp. 379-407).

O'Hara, M. W., & Swain, A. M. (1996). Rates and risk of postpartum depression - a meta-analysis. *International Review of Psychiatry, 8*, 37-54.

Oller, D. K., Buder, E. H., Ramsdell, H. L., Warlaumont, A. S., Chorna, L., & Bakeman, R. (2013). Functional flexibility of infant vocalization and the emergence of language. *Proceedings of the National Academy of Sciences of the United States of America, 110*(16), 6318-6323.

Oller, D. K., & Eilers, R. E. (1988). The role of audition in infant babbling. *Child development, 59*(2), 441-449.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational intelligence and neuroscience, 2011*, 156869. doi: 10.1155/2011/156869

Ostwald, P. F., & Murray, T. (1985). The communicative and diagnostic significance of infant sounds. In B. M. Lester & C. F. Boukydis (Eds.), *Infant Crying: Theoretical and Research Perspectives* (pp. 139-158). New York: Plenum Press.

Out, D., Pieper, S., Bakermans-Kranenburg, M. J., & Van Ijzendoorn, M. H. (2010). Physiological reactivity to infant crying: A behavioral genetic study. *Genes, Brain and Behavior, 9*(8), 868-876.

Out, D., Pieper, S., Bakermans-Kranenburg, M. J., Zeskind, P. S., & van Ijzendoorn, M. H. (2010). Intended sensitive and harsh caregiving responses to infant crying: The role of cry pitch and perceived urgency in an adult twin sample. *Child Abuse and Neglect, 34*(11), 863-873.

Palva, S., & Palva, J. M. (2007). New vistas for α-frequency band oscillations. *Trends in Neurosciences, 30*(4), 150-158.

Papousek, H., & Papousek, M. (1975). Cognitive aspects of preverbal social interaction between human infants and adults. *Ciba Foundation symposium*(33), 241-269.

Papousek, H., & Papousek, M. (1983). Biological basis of social interactions: Implications of research for an understanding of behavioural deviance. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 24*(1), 117-129.

Papousek, H., & Papousek, M. (1987). Intuitive Parenting: A Dialectic Counterpart to the Infant's Integrative Competence. In J. D. Osofsky (Ed.), *Handbook of Infant Development* (2nd ed., pp. 669-720). New York: Wiley.

Papousek, M. (2007). Communication in early infancy: An arena of intersubjective learning. *Infant Behavior and Development, 30*(2), 258-266.

Papousek, M., & Papousek, H. (1989). Forms and Functions of Vocal Matching in Interactions between Mothers and Their Precanonical Infants. *First Language, 9*(2), 137-157.

Parsons, C. E., Stark, E. A., Young, K. S., Stein, A., & Kringelbach, M. L. (2013). Understanding the human parental brain: A critical role of the orbitofrontal cortex. *Social Neuroscience, 8*(6), 525-543.

Parsons, C. E., Young, K. S., Bhandari, R., Van Ijzendoorn, M. H., Bakermans-Kranenburg, M. J., Stein, A., & Kringelbach, M. L. (in press). The Bonnie baby: experimentally manipulated temperament affects perceived cuteness and motivation to view infant faces. *Developmental Science*.

Parsons, C. E., Young, K. S., Kumari, N., Stein, A., & Kringelbach, M. L. (2011). The motivational salience of infant faces is similar for men and women. *PLoS One, 6*(5), e20632.

Parsons, C. E., Young, K. S., Mohseni, H., Woolrich, M. W., Thomsen, K. R., Joensson, M., . . . Kringelbach, M. L. (2013). Minor structural abnormalities in the infant face disrupt neural processing: A unique window into early caregiving responses. *Social Neuroscience, 8*(4), 268-274.

Parsons, C. E., Young, K. S., Murray, L., Stein, A., & Kringelbach, M. L. (2010). The functional neuroanatomy of the evolving parent-infant relationship. *Progress in Neurobiology, 91*, 220-241.

Parsons, C. E., Young, K. S., Parsons, E., Dean, A., Murray, L., Goodacre, T., . . . Kringelbach, M. L. (2011). The effect of cleft lip on adults' responses to faces: cross-species findings. *PLoS One, 6*(10), e25897.

Paulson, J. F., & Bazemore, S. D. (2010). Prenatal and postpartum depression in fathers and its association with maternal depression: A meta-analysis. *JAMA - Journal of the American Medical Association, 303*(19), 1961-1969.

Pearson, R. M., Cooper, R. M., Penton-Voak, I. S., Lightman, S. L., & Evans, J. (2010). Depressive symptoms in early pregnancy disrupt attentional processing of infant emotion. *Psychological Medicine, 40*(4), 621-631.

Pearson, R. M., Lightman, S. L., & Evans, J. (2011). Attentional processing of infant emotion during late pregnancy and mother-infant relations after birth. *Archives of Women's Mental Health, 14*(1), 23-31.

Péron, J., El Tamer, S., Grandjean, D., Leray, E., Travers, D., Drapier, D., . . . Millet, B. (2011). Major depressive disorder skews the recognition of emotional prosody. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 35*(4), 987-996.

Petitto, L. A., & Marentette, P. F. (1991). Babbling in the manual mode: Evidence for the ontogeny of language. *Science, 251*(5000), 1493-1496.

Petrovic, P., Dietrich, T., Fransson, P., Andersson, J., Carlsson, K., & Ingvar, M. (2005). Placebo in emotional processing—induced expectations of anxiety relief activate a generalized modulatory network. *Neuron, 46*(6), 957-969.

Poncelet, B. P., Wedeen, V. J., Weisskoff, R. M., & Cohen, M. S. (1992). Brain parenchyma motion: Measurement with cine echo-planar MR imaging. *Radiology, 185*(3), 645-651.

Pongrácz, P., Molnár, C., Miklósi, Á., & Csányi, V. (2005). Human listeners are able to classify dog (Canis familiaris) barks recorded in different situations. *Journal of Comparative Psychology, 119*(2), 136-144.

Poobalan, A. S., Aucott, L. S., Ross, L., Smith, W. C. S., Helms, P. J., & Williams, J. H. G. (2007). Effects of treating postnatal depression on mother-infant interaction and child development: Systematic review. *British Journal of Psychiatry, 191*(NOV.), 378-386.

Porges, S. W. (1995). Cardiac vagal tone: A physiological index of stress. *Neuroscience and Biobehavioral Reviews, 19*(2), 225-233.

Porter, F. L., Miller, R. H., & Marshall, R. E. (1986). Neonatal pain cries: effect of circumcision on acoustic features and perceived urgency. *Child development, 57*(3), 790-802.

Porter, F. L., Porges, S. W., & Marshall, R. E. (1988). Newborn pain cries and vagal tone: parallel changes in response to circumcision. *Child development, 59*(2), 495-505.

Porter, R. H., Cernoch, J. M., & McLaughlin, F. J. (1983). Maternal recognition of neonates through olfactory cues. *Physiol Behav, 30*(1), 151-154. doi: 0031-9384(83)90051-3 [pii]

Powers, M., & Emmelkamp, P. (2008). Virtual reality exposure therapy for anxiety disorders: A meta-analysis. *Journal of anxiety disorders, 22*(3), 561-569. doi: 10.1016/j.janxdis.2007.04.006

Pressnitzer, D., Sayles, M., Micheyl, C., & Winter, I. (2008). Perceptual organization of sound begins in the auditory periphery. *Current biology, 18*(15), 1124-1128. doi: 10.1016/j.cub.2008.06.053

Protopapas, A., & Lieberman, P. (1997). Fundamental frequency of phonation and perceived emotional stress. *Journal of the Acoustical Society of America, 101*(4), 2267-2277.

Punkanen, M., Eerola, T., & Erkkila, J. (2011). Biased emotional recognition in depression: perception of emotions in music by depressed patients. *J Affect Disord, 130*(1-2), 118-126.

Quadflieg, S., Mohr, A., Mentzel, H. J., Miltner, W. H. R., & Straube, T. (2008). Modulation of the neural network involved in the processing of anger prosody: The role of task-relevance and social phobia. *Biological Psychology, 78*(2), 129-137.

Ramchandani, P., Stein, A., Evans, J., & O'Connor, T. G. (2005). Paternal depression in the postnatal period and child development: A prospective population study. *Lancet, 365*(9478), 2201-2205.

Ramchandani, P. G., Stein, A., O'Connor, T. G., Heron, J., Murray, L., & Evans, J. (2008). Depression in men in the postnatal period and later child psychopathology: A population cohort study. *Journal of the American Academy of Child and Adolescent Psychiatry, 47*(4), 390-398.

Ranote, S., Elliott, R., Abel, K. M., Mitchell, R., Deakin, J. F. W., & Appleby, L. (2004). The neural basis of maternal responsiveness to infants: An fMRI study. *NeuroReport, 15*(11), 1825-1829.

Riem, M. M. E., Bakermans-Kranenburg, M. J., van Ijzendoorn, M. H., Out, D., & Rombouts, S. A. R. B. (2012). Attachment in the brain: Adult attachment representations predict amygdala and behavioral responses to infant crying. *Attachment and Human Development, 14*(6), 533-551.

Riem, M. M. E., Pieper, S., Out, D., Bakermans-Kranenburg, M. J., & van Ijzendoorn, M. H. (2011). Oxytocin receptor gene and depressive symptoms associated with physiological reactivity to infant crying. *Social Cognitive and Affective Neuroscience, 6*(3), 294-300.

Riem, M. M. E., Van Ijzendoorn, M. H., Tops, M., Boksem, M. A. S., Rombouts, S. A. R. B., & Bakermans-Kranenburg, M. J. (2012). No laughing matter: Intranasal oxytocin administration changes functional brain connectivity during exposure to infant laughter. *Neuropsychopharmacology, 37*(5), 1257-1266.

Riem, M. M. E., Voorthuis, A., Bakermans-Kranenburg, M. J., & Van Ijzendoorn, M. H. (2013). Pity or peanuts? Oxytocin induces different neural responses to the same infant crying labeled as sick or bored. *Developmental Science*. doi: 10.1111/desc.12103

Rolls, E. T. (2000). Précis of The brain and emotion. *Behavioral and Brain Sciences, 23*(02), 177-191. doi: 10.1017/S0140525X00002429

Roth, D. A. E., Kishon-Rabin, L., Hildesheimer, M., & Karni, A. (2005). A latent consolidation phase in auditory identification learning: Time in the awake state is sufficient. *Learning and Memory, 12*(2), 159-164.

Rothgänger, H. (2003). Analysis of the sounds of the child in the first year of age and a comparison to the language. *Early Human Development, 75*(1-2), 55-69.

Ruddy, M. G., & Bornstein, M. H. (1982). Cognitive correlates of infant attention and maternal stimulation over the first year of life. *Child development, 53*(1), 183-188.

Rutter, M., & Quinton, D. (1984). Parental psychiatric disorder: Effects on children. *Psychological Medicine, 14*(4), 853-880.

Sabbe, B., Hulstijn, W., van Hoof, J., Tuynman-Qua, H. G., & Zitman, F. (1999). Retardation in depression: assessment by means of simple motor tasks. *J Affect Disord, 55*(1), 39-44. doi: S0165-0327(98)00087-1 [pii]

Salekin, R. T., & Lochman, J. E. (2008). Child and adolescent psychopathy: The search for protective factors. *Criminal Justice and Behavior, 35*(2), 159-172.

Salomon, K., Clift, A., Karlsdóttir, M., & Rottenberg, J. (2009). Major Depressive Disorder Is Associated With Attenuated Cardiovascular Reactivity and Impaired Recovery Among Those Free of Cardiovascular Disease. *Health Psychology, 28*(2), 157-165.

Sander, K., Frome, Y., & Scheich, H. (2007). FMRI activations of amygdala, cingulate cortex, and auditory cortex by infant laughing and crying. *Human Brain Mapping, 28*(10), 1007-1022.

Sander, K., & Scheich, H. (2001). Auditory perception of laughing and crying activates human amygdala regardless of attentional state. *Cognitive Brain Research, 12*(2), 181-198.

Sander, K., & Scheich, H. (2005). Left auditory cortex and amygdala, but right insula dominance for human laughing and crying. *J Cogn Neurosci, 17*(10), 1519-1531.

Sauter, D. A., Eisner, F., Calder, A. J., & Scott, S. K. (2010). Perceptual cues in nonverbal vocal expressions of emotion. *Quarterly Journal of Experimental Psychology, 63*(11), 2251-2272.

Scherer, K. R. (1986). Vocal Affect Expression. A Review and a Model for Future Research. *Psychological Bulletin, 99*(2), 143-165.

Scherer, K. R. (1994). Affect bursts. In S. H. M. Van Goozen, N. E. Van de Poll & J. A. Sergeant (Eds.), *Emotions: Essays on Emotion Theory* (pp. 161-193). Hillsdale, NJ: Erlbaum.

Scherer, K. R. (1995). Expression of emotion in voice and music. *Journal of voice, 9*(3), 235-248.

Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication, 40*(1-2), 227-256.

Scherer, K. R., Banse, R., & Wallbott, H. G. (2001). Emotion inferences from vocal expression correlate across languages and cultures. *Journal of Cross-Cultural Psychology, 32*(1), 76-92.

Scherer, K. R., Ladd, D. R., & Silverman, K. E. A. (1984). Vocal cues to speaker affect: Testing two models. *Journal of the Acoustical Society of America, 76*(5), 1346-1356.

Schirmer, A., & Kotz, S. A. (2006). Beyond the right hemisphere: Brain mechanisms mediating vocal emotional processing. *Trends in Cognitive Sciences, 10*(1), 24-30.

Schneider, P., Scherg, M., Dosch, H., Specht, H., Gutschalk, & Rupp, A. (2002). Morphology of Heschl's Gyrus Reflects Enhanced Activation in the Auditory Cortex of Musicians. *Nature Neuroscience, 5*(7), 688-694. doi: citeulike-article-id:153667

Schnupp, J. W. H. (2008). Auditory Neuroscience: Sound Segregation in the Brainstem? *Current biology, 18*(16), R705-R706.

Schrijvers, D., Hulstijn, W., & Sabbe, B. (2008). Psychomotor symptoms in depression: a diagnostic, pathophysiological and therapeutic tool. *Journal of Affective Disorders, 109*(1-2), 1-20. doi: 10.1016/j.jad.2007.10.019

Schröder, M. (2003). Experimental study of affect bursts. *Speech Communication, 40*(1-2), 99-116.

Schuetze, P., & Zeskind, P. S. (2001). Relations between Women's Depressive Symptoms and Perceptions of Infant Distress Signals Varying in Pitch. *Infancy, 2*(4), 483-499.

Schulz, G., Varga, M., Jeffires, K., Ludlow, C., & Braun, A. (2005). Functional neuroanatomy of human vocalization: an H215O PET study. *Cerebral Cortex, 15*(12), 1835-1847.

Scott, S. K., Sauter, D., & McGettigan, C. (2010). Brain Mechanism for Processing Perceived Emotional Vocalization in Humans. In S. M. Brudzynski (Ed.), *Handbook of mammalian vocalization: An integrative neuroscience approach* (pp. 187-197). Amsterdam: Elsevier.

Seifritz, E., Esposito, F., Neuhoff, J. G., Lüthi, A., Mustovic, H., Dammann, G., . . . Di Salle, F. (2003). Differential sex-independent amygdala response to infant crying and laughing in parents versus nonparents. *Biological Psychiatry, 54*(12), 1367-1375.

Sem-Jacobsen, C. W., & Torkildsen, A. (1960). Depth recording and electrical stimulation in the human brain. In E. R. Ramey & D. S. O'Doherty (Eds.), *Electrical Studies on the Unanesthetized Brain* (pp. 275-287). New York: Paul B. Hoeber, Inc.

Sillery, E., Bittar, R. G., Robson, M. D., Behrens, T. E., Stein, J., Aziz, T. Z., & Johansen-Berg, H. (2005). Connectivity of the human periventricular-periaqueductal gray region. *J Neurosurg, 103*(6), 1030-1034. doi: 10.3171/jns.2005.103.6.1030

Skoe, E., & Kraus, N. (2010). Auditory brain stem response to complex sounds: a tutorial. *Ear and hearing, 31*(3), 302-324. doi: 10.1097/AUD.0b013e3181cdb272

Smith, D. R., Patterson, R. D., Turner, R., Kawahara, H., & Irino, T. (2005). The processing and perception of size information in speech sounds. *J Acoust Soc Am, 117*(1), 305-318.

Sohr-Preston, S. L., & Scaramella, L. V. (2006). Implications of timing of maternal depressive symptoms for early cognitive and language development. *Clinical Child and Family Psychology Review, 9*(1), 65-83.

Soltis, J. (2004). The signal functions of early infant crying. *Behavioral and Brain Sciences, 27*(4), 443-458.

Sprengelmeyer, R., Lewis, J., Hahn, A., & Perrett, D. I. (2013). Aesthetic and Incentive Salience of Cute Infant Faces: Studies of Observer Sex, Oral Contraception and Menstrual Cycle. *PLoS One, 8*(5).

Sprengelmeyer, R., Perrett, D. I., Fagan, E. C., Cornwell, R. E., Lobmaier, J. S., Sprengelmeyer, A., . . . Young, A. W. (2009). The cutest little baby face: A hormonal link to sensitivity to cuteness in infant faces. *Psychological Science, 20*(2), 149-154.

Squire, S., & Stein, A. (2003). Functional MRI and parental responsiveness: a new avenue into parental psychopathology and early parent-child interactions? *The British journal of psychiatry, 183*(6), 481-483.

Sroufe, L. A. (2005). Attachment and development: A prospective, longitudinal study from birth to adulthood. *Attachment and Human Development, 7*(4), 349-367.

Sroufe, L. A., & Wunsch, J. P. (1972). The development of laughter in the first year of life. *Child development, 43*(4), 1326-1344.

Stein, A., Arteche, A., Lehtonen, A., Craske, M., Harvey, A., Counsell, N., & Murray, L. (2010). Interpretation of infant facial expression in the context of maternal postnatal depression. *Infant Behavior and Development, 33*(3), 273-278.

Stein, A., Gath, D. H., Bucher, J., Bond, A., Day, A., & Cooper, P. J. (1991). The relationship between post-natal depression and mother-child interaction. *British Journal of Psychiatry, 158*(JAN.), 46-52.

Stein, A., Lehtonen, A., Harvey, A. G., Nicol-Harper, R., & Craske, M. (2009). The influence of postnatal psychiatric disorder on child development: Is maternal preoccupation one of the key underlying processes? *Psychopathology, 42*(1), 11-21.

Stein, A., Malmberg, L. E., Leach, P., Barnes, J., Sylva, K., & Team, F. (2013). The influence of different forms of early childcare on children's emotional and behavioural development at school entry. *Child: Care, Health and Development, 39*(5), 676-687. doi: 10.1111/j.1365-2214.2012.01421.x

Stein, A., Malmberg, L. E., Sylva, K., Barnes, J., Leach, P., & the FCCC Team. (2008). The influence of maternal depression, caregiving, and socioeconomic status in the post-natal year on children's language development. *Child: Care, Health and Development, 34*(5), 603-612.

Stern, J. M., & Lonstein, J. S. (2001). Neural mediation of nursing and related maternal behaviors. *Progress in brain research, 133*, 263-278.

Strathearn, L., Li, J., Fonagy, P., & Montague, P. R. (2008). What's in a smile? Maternal brain responses to infant facial cues. *Pediatrics, 122*(1), 40-51.

Sukikara, M. H., Mota-Ortiz, S. R., Baldo, M. V., Felicio, L. F., & Canteras, N. S. (2010). The periaqueductal gray and its potential role in maternal behavior inhibition in response to predatory threats. *Behavioural Brain Research, 209*(2), 226-233.

Sullivan, M. W., & Lewis, M. (2003). Emotional expressions of young infants and children: A practitioner's primer. *Infants & Young Children, 16*(2), 120-142.

Šuta, D., Kvašňák, E., Popelář, J., & Syka, J. (2003). Representation of Species-Specific Vocalizations in the Inferior Colliculus of the Guinea Pig. *Journal of Neurophysiology, 90*(6), 3794-3808.

Šuta, D., Popelář, J., Kvašňák, E., & Syka, J. (2007). Representation of species-specific vocalizations in the medial geniculate body of the guinea pig. *Experimental Brain Research, 183*(3), 377-388.

Swain, J. E., Leckman, J. F., Mayes, L. C., Feldman, R., & Schultz, R. T. (2005). Early human parent-infant bond development: fMRI, thoughts and behaviors. *Biol Psychiatry, 57*.

Swain, J. E., Tasgin, E., Mayes, L. C., Feldman, R., Todd Constable, R., & Leckman, J. F. (2008). Maternal brain response to own baby-cry is affected by cesarean section delivery. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 49*(10), 1042-1052.

Tallal, P., Miller, S. L., Bedi, G., Byma, G., Wang, X., Nagarajan, S. S., . . . Merzenich, M. M. (1996). Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science, 271*(5245), 81-84.

Tasker, R. R. (1982). Identification of pain processing systems by electrical stimulation of the brain. *Human Neurobiology, 1*(4), 261-272.

Taylor, B. P., Bruder, G. E., Stewart, J. W., McGrath, P. J., Halperin, J., Ehrlichman, H., & Quitkin, F. M. (2006). Psychomotor slowing as a predictor of fluoxetine nonresponse in depressed outpatients. *Am J Psychiatry, 163*(1), 73-78. doi: 163/1/73 [pii] 10.1176/appi.ajp.163.1.73

Tervaniemi, M., Just, V., Koelsch, S., Widmann, A., & Schroger, E. (2005). Pitch discrimination accuracy in musicians vs nonmusicians: an event-related potential and behavioral study. *Exp Brain Res, 161*(1), 1-10. doi: 10.1007/s00221-004-2044-5

Thompson, R., Bruzek, J., & Cotnoir-Bichelman, N. (2011). The role of negative reinforcement in infant caregiving: an experimental simulation. *Journal of applied behavior analysis, 44*(2), 295-304. doi: 10.1901/jaba.2011.44-295

Thompson, W. F., Schellenberg, E. G., & Husain, G. (2004). Decoding Speech Prosody: Do Music Lessons Help? *Emotion, 4*(1), 46-64.

Thomsen, K. R., Lou, H. C., Joensson, M., Hyam, J. A., Holland, P., Parsons, C. E., . . . Green, A. L. (2011). Impact of emotion on consciousness: positive stimuli enhance conscious reportability. *PLoS One, 6*(4), e18686.

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., . . . Nelson, C. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research, 168*(3), 242-249.

Tracey, I., Ploghaus, A., Gati, J. S., Clare, S., Smith, S., Menon, R. S., & Matthews, P. M. (2002). Imaging Attentional Modulation of Pain in the Periaqueductal Gray in Humans. *Journal of Neuroscience, 22*(7), 2748-2752.

Trevarthen, C. (1977). Descriptive Analyses of Infant Communicative Behaviour. In H. R. Schaffer (Ed.), *Studies in Mother-Infant Interaction*. London: Academic Press.

Trevarthen, C., Murray, L., & Hubley, P. (1981). Psychology of infants. *Scientific Foundations of Clinical Paediatrics*, 211-274.

Tronick, E., Als, H., & Adamson, L. (1978). The infant's response to entrapment between contradictory messages in face-to-face interaction. *Journal of the American Academy of Child Psychiatry, 17*(1), 1-13.

Tronick, E., & Gianino, A. F. (1986). The transmission of maternal disturbance to the infant. In E. Tronick & T. Field (Eds.), *Maternal Depression and Infant Disturbance* (pp. 5-11). San Fransisco: Jossey-Bass.

Truong, K. P., & van Leeuwen, D. A. (2007). Automatic discrimination between laughter and speech. *Speech Communication, 49*(2), 144-158. doi: http://dx.doi.org/10.1016/j.specom.2007.01.001

Uekermann, J., Abdel-Hamid, M., Lehmkämper, C., Vollmoeller, W., & Daum, I. (2008). Perception of affective prosody in major depression: A link to executive functions? *Journal of the International Neuropsychological Society, 14*(4), 552-561.

Vanderhorst, V. G. J. M., Terasawa, E., Ralston Iii, H. J., & Holstege, G. (2000). Monosynaptic projections from the lateral periaqueductal gray to the nucleus retroambiguus in the rhesus monkey: Implications for vocalization and reproductive behavior. *Journal of Comparative Neurology, 424*(2), 251-268.

Verhage, M. L., Oosterman, M., & Schuengel, C. (2013). Parenting Self-Efficacy is Associated with Cry Perception, Not Autonomic Responses, during a Cry Response Task. *Parenting, 13*(4), 253-265.

Vygotsky, L. S. (1978). *Mind in Society: the Development of Higher Psychological Processes*. Cambridge, MA: Harvard University Press.

Wang, C., Ulbert, I., Schomer, D. L., Marinkovic, K., & Halgren, E. (2005). Responses of human anterior cingulate cortex microdomains to error detection, conflict monitoring, stimulus-response mapping, familiarity, and orienting. *Journal of Neuroscience, 25*(3), 604-613.

Ward, M. J., & Carlson, E. A. (1995). Associations among Adult Attachment Representations, Maternal Sensitivity, and Infant-Mother Attachment in a Sample of Adolescent Mothers. *Child development, 66*(1), 69-79.

Wasz-Höckert, O., Michelsson, K., & Lind, J. (1985). Twenty-five years of Scandinavian cry research. In B. M. Lester & C. F. Boukydis (Eds.), *Infant Crying: Theoretical and Research Perspectives* (pp. 83-104). New York: Plenum Press.

Wasz-Höckert, O., Partanen, T. J., Vuorenkoski, V., Michelsson, K., & Valanne, E. (1964). The identification of some specific meanings in infant vocalization. *Experientia, 20*(3), 154.

Watson, D., Wiese, D., Vaidya, J., & Tellegen, A. (1999). The two general activation systems of affect: Structural findings, evolutionary considerations, and psychobiological evidence. *Journal of Personality and Social Psychology, 76*(5), 820.

Weinberg, M. K., & Tronick, E. Z. (1996). Infant Affective Reactions to the Resumption of Maternal Interaction after the Still-Face. *Child development, 67*(3), 905-914.

Weiner, B. (1985). An Attributional Theory of Achievement Motivation and Emotion. *Psychological Review, 92*(4), 548-573.

Wendland-Carro, J., Piccinini, C. A., & Millar, W. S. (1999). The role of an early intervention on enhancing the quality of mother-infant interaction. *Child development, 70*(3), 713-721.

White, D. A., Myerson, J., & Hale, S. (1997). How cognitive is psychomotor slowing in depression? Evidence from a meta-analysis. *Aging, Neuropsychology, and Cognition, 4*(3), 166-174.

Wiech, K., & Tracey, I. (2009). The influence of negative emotions on pain: Behavioral effects and neural mechanisms. *NeuroImage, 47*(3), 987-994.

Wiesenfeld, A. R., Malatesta, C. Z., & DeLoach, L. L. (1981). Differential parental response to familial and unfamiliar infant distress signals. *Infant Behavior and Development, 4*(3), 281-295.

Wildgruber, D., Ackermann, H., Kreifelts, B., & Ethofer, T. (2006). Cerebral processing of linguistic and emotional prosody: fMRI studies. *Progress in brain research, 156*, 249-268.

Wood, R. M., & Gustafson, G. E. (2001). Infant crying and adults' anticipated caregiving responses: Acoustic and contextual influences. *Child development, 72*(5), 1287-1300.

Wright, B. A., & Sabin, A. T. (2007). Perceptual learning: How much daily training is enough? *Experimental Brain Research, 180*(4), 727-736.

Xing, D., Yeh, C. I., & Shapley, R. M. (2009). Spatial spread of the local field potential and its laminar variation in visual cortex. *Journal of Neuroscience, 29*(37), 11540-11549.

Yamamoto, R., Ariely, D., Chi, W., Langleben, D. D., & Elman, I. (2009). Gender differences in the motivational processing of babies are determined by their facial attractiveness. *PLoS One, 4*(6), e6042. doi: 10.1371/journal.pone.0006042

Young, R. F., & Brechner, T. (1986). Electrical stimulation of the brain for relief of intractable pain due to cancer. *Cancer, 57*(6), 1266-1272.

Zald, D. H., & Pardo, J. V. (2002). The neural correlates of aversive auditory stimulation. *NeuroImage, 16*(3 Pt 1), 746-753.

Zeskind, P. S., & Collins, V. (1987). Pitch of infant crying and caregiver responses in a natural setting. *Infant Behavior and Development, 10*(4), 501-504.

Zeskind, P. S., & Marshall, T. R. (1988). The Relation between Variations in Pitch and Maternal Perceptions of Infant Crying. *Child development, 59*(1), 193-196. doi: 10.2307/1130401

Zhang, W. T., Mainero, C., Kumar, A., Wiggins, C. J., Benner, T., Purdon, P. L., . . . Sorensen, A. G. (2006). Strategies for improving the detection of fMRI activation in trigeminal pathways with cardiac gating. *NeuroImage, 31*(4), 1506-1512.

Appendix A. Example waveforms and spectrograms for different vocalisation categories

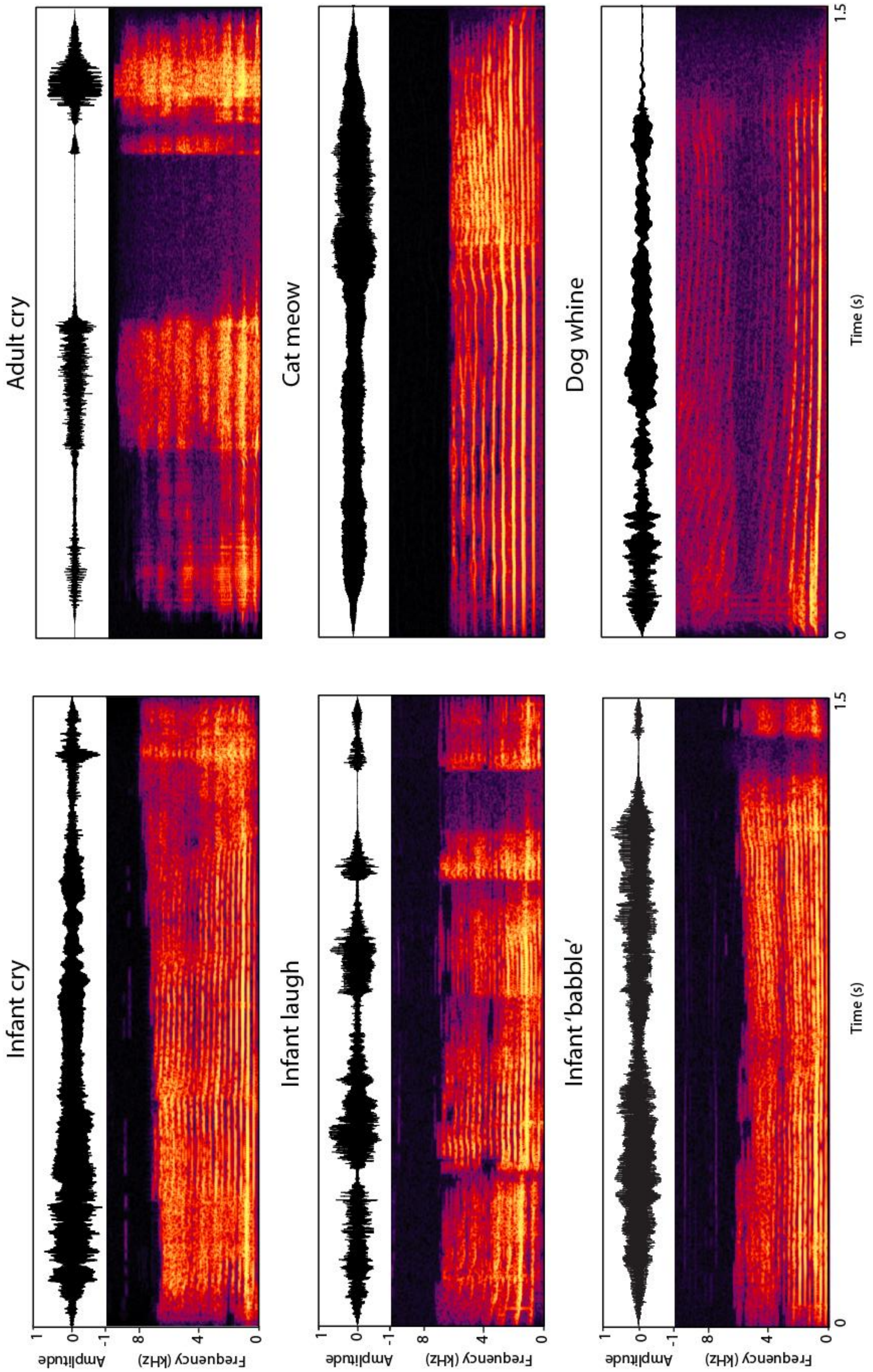

Figure A1. Example waveforms (upper) and spectrograms (lower) for representative stimuli from each category of vocalisation used in the development of the 'OxVoc' database (Chapter 2, section 2.2)

Table B8. *The effects of gender on experimental measures of effortful motor performance*

Experimental Measure	*df*	*t*	*r*	*p*
Game score	56	0.59	.08	.45
Minimum pressure	52	0.51	.07	.48
Maximum pressure	52	0.51	.07	.48
Mean pressure	54	0.62	.08	.44

Note. Statistical results were obtained from independent samples t-tests. There were no effects of gender on any measures of effortful motor performance (see Section 3.3.4)

Table C1. *Experimental measures of effortful motor performance with participants who were parents removed*

Experimental Measure	*df*	*F*	*r*	*p*
		Between Subjects		
Game score	52	9.59*	.39	.003
Minimum Pressure	49	6.20*	.34	.02
Maximum pressure	49	6.20*	.34	.02
Mean pressure	49	1.20	.15	.28
		Within Subjects		
Game score	104	5.66*	.23	.005
Minimum Pressure	98	10.53*	.31	.001
Maximum pressure	98	10.54*	.31	.001
Mean pressure	98	1.39	.12	.25

Note. Statistical results were obtained from one-way ANOVAs. The between-subjects factor was participant group (healthy adults vs. adults with depression), within-subjects factor was stimulus category (infant cries, adult cries, bird sounds). The same pattern of effects was observed with participants who were parents removed, as when they were included (see section 3.3.5). *indicates a significant effect ($p < .05$)

Appendix D. Statistical comparison of individual channel event-related LFP activity in the PAG across stimulus categories

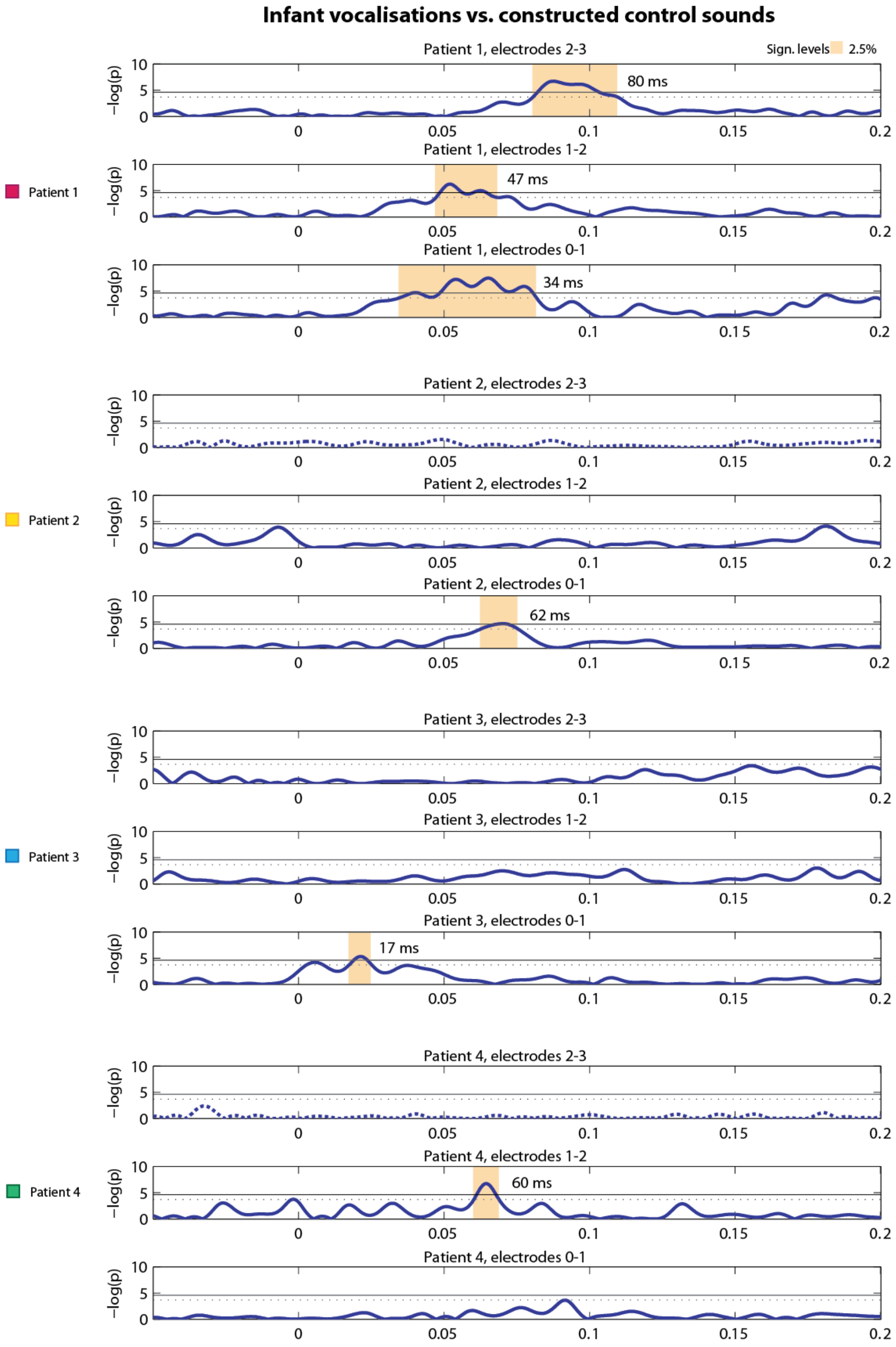

Figure D1. Individual channel data for event-related analysis of LFP recordings from the PAG. Plots show normalized *p*-values obtained from ANOVA tests comparing differences in LFP activity in response to infant vocalisations compared with constructed control sounds. Results demonstrate significant differences in all patients, with varying onset between 17 and 80ms. Dashed horizontal lines indicate an alpha level of 0.025, solid horizontal lines indicate an alpha level of 0.01. Orange shading indicates significant differences in LFP activity (α = 0.025). Dashed blue lines indicate channels in which data was not recorded due to an equipment failure (see Section 5.3.2)

Appendix D. Statistical comparison of individual channel event-related LFP activity in the PAG across stimulus categories

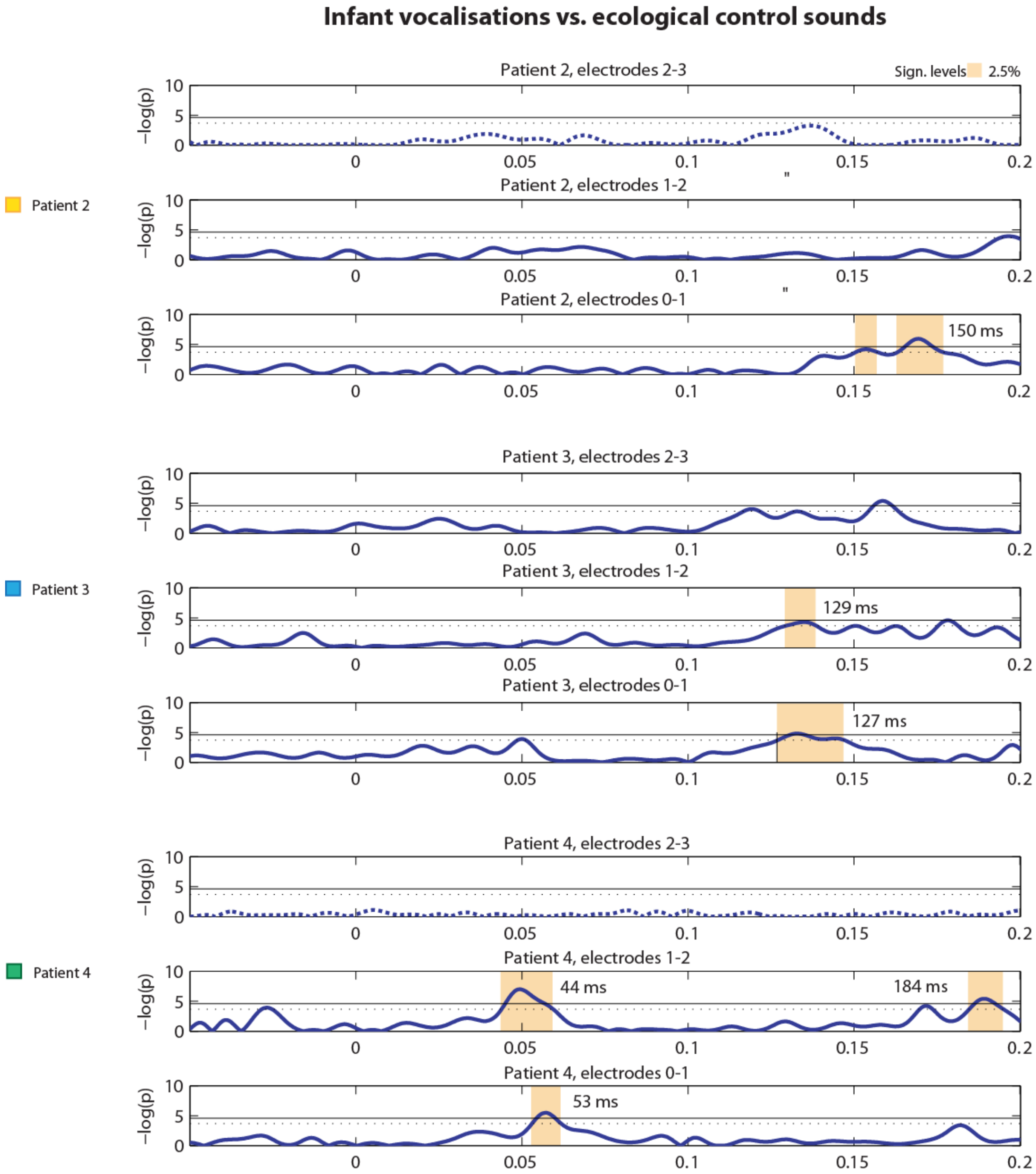

Figure D2. Individual channel data for event-related analysis of LFP recordings from the PAG. Plots show normalized *p*-values obtained from ANOVA tests comparing differences in LFP activity in response to infant vocalisations compared with ecological control sounds. Results show significant differences in all patients, with a variable onset beginning between 44 and 150ms. Dashed horizontal lines indicate an alpha level of 0.025, solid horizontal lines indicate an alpha level of 0.01. Orange shading indicates significant differences in LFP activity (α = 0.025). Dashed blue lines indicate channels in which data was not recorded due to an equipment failure (see Section 5.3.2)