Capturing the Attention of Caregivers: Variability in Infant Vocalizations

Catharine A. Castelluccio de Diesbach

The Graduate Center, City University of New York
Capturing the attention of caregivers: Variability in infant vocalizations

by

Catharine Castelluccio de Diesbach

This manuscript has been read and accepted for the Graduate Faculty in Speech-Language-Hearing Sciences in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy

__________________________________________
Date
John L. Locke
Chair of Examining Committee

__________________________________________
Date
Klara Marton
Executive Officer

Supervisory Committee:
Martin R. Gitterman
Glenis R. Long
Sethu Karthikeyan
Valerie L. Shafer, External Reviewer

THE CITY UNIVERSITY OF NEW YORK
Abstract

CAPTURING THE ATTENTION OF CAREGIVERS: VARIABILITY IN INFANT VOCALIZATIONS

by

CATHARINE CASTELUCCIO de DIESBACH

Advisor: Dr. John L. Locke

The effect of variability in infant vocalizations on potential caregivers’ heart rate variability (HRV), facial expressions, and subjective ratings on emotional reactions and desire to approach the baby was examined in an evolutionary context. Recordings of non-canonical, canonical, fussing, and crying vocalizations were utilized to elicit physiological and self-reported reactions from sixty participants. Breastfeeding mothers, non-mothers at high estradiol point in menstrual cycle, non-mothers at low estradiol point in menstrual cycle, fathers, and non-fathers were included in the study. Participants wore Polar RS800 heart rate monitors, were video recorded for facial expression analysis, and filled out 11 point self-rating forms on emotional reactions to the infant vocal stimuli. It was expected that participants would show higher HRV for the canonical vocalizations as compared to non-canonical, fussing and crying vocal stimuli. Overall HRV as measured by SDNN (standard deviation of NN, or “normal-to-normal” interbeat intervals), was highest for the recorded babbling, however these differences were not significant. Most raters considered crying and fussing to be strong indicators of a need for interaction. Participants showed the greatest percentage of happy facial expressions (evaluated via analysis of video recordings) and also self-reported the babbling vocalizations high on “happiness” and “most liked”, as predicted. Although the predicted directions for the differences between mothers and non-mothers at two different assumed estradiol levels in menstrual cycle were not significant, breastfeeding mothers did show higher facial expressions of happiness while listening to the babbling stimuli, gave higher scores of self-rated sadness when listening to crying, and rated their irritation levels lower and the desire to pick up the baby higher for the fussing stimuli.
The square root of the mean squared difference of successive NN intervals were significantly higher in fathers than non-fathers while listening to the babbling stimuli. Fathers had significantly higher self-reported happiness levels and higher scores towards the “most liked” end of the rating scale for the babbling stimuli. The results are discussed within an evolutionary framework considering the potential influence of parental selection of vocal behaviors, an attraction to complexity of sounds across species, as well as the possible influence of hormones on potential caregivers’ responses to infant needs.
Acknowledgments

I would like to thank my mentor, Dr. John Locke, for his support, advice, encouragement and patience during the dissertation process, and for sharing his passion for an evolutionary focus in speech and language science. I was hooked on day one of his seminar on the evolution and development of language, and have since been fascinated with evolutionary developmental linguistics. Thank you, John, for your inspirational teaching that has changed the way I think about speech and language for the rest of my career.

Thank you also to Dr. Martin Gitterman and Dr. Glenis Long for agreeing to be a part of my dissertation committee, and helping me to incorporate your suggestions into my research proposal, for taking time to read and edit my dissertation draft, and for supporting me and my interests throughout the PhD program. Thank you Dr. Sethu Karthikeyan for reading my dissertation and providing helpful comments and suggestions. I am grateful to Dr. Valerie Shafer for accepting the role as external reviewer.

I would like to thank Dr. Klara Marton, the executive officer at The Graduate Center, CUNY, for her support and excellent advice; she always encouraged me to continue moving forward.

I am grateful to all the participants who volunteered to be a part of this study.

Thank you Dr. Loraine Obler and the Neurolinguistics lab at The Graduate Center, CUNY for introducing me to research with Dr. Erika Levy (Teachers College, Columbia University) and Dr. Mira Goral (Lehman College, CUNY). This experience contributed tremendously to my growth as a doctoral student.

Gary Chant, senior college lab technician at The Graduate Center, CUNY, answered my technical questions thoroughly and patiently.
I am thankful to Saren Seeley, former research assistant at Hunter College, CUNY, Psychology Lab, for her help with HRV analysis.

Dr. Stephen Porges and Dr. Sue Carter graciously answered all of my questions via e-mail.
I am very grateful that Dr. Kimbrough Oller, Dr. James Green, and Dr. Francine Lang Porter allowed me to use their recordings of infant vocalizations for the stimuli in this study.

Dr. Sarah Hrdy provided very helpful comments on my thoughts and guided me during the formation of my proposal.

Thank you, Guillaume, for all of your support.

I would not have been able to complete my dissertation without my family’s love and encouragement throughout the whole process. Thank you Mom, Dad, and to my two greatest inspirations, Alex and Amelie.
# Table of contents

Abstract .......................................................................................................................... iv

Acknowledgments ........................................................................................................ vi

List of tables and figures ............................................................................................... xi

**Chapter I. INTRODUCTION** ......................................................................................... 1

**Chapter II. REVIEW OF LITERATURE** ...................................................................... 2

2.1 Evolutionary developmental linguistics ................................................................ 2

    2.1.1 Developmental plasticity .............................................................................. 2

2.2 Parental investment ............................................................................................... 3

2.3 Cooperative breeding ............................................................................................ 3

2.4 The environment of evolutionary adaptedness ...................................................... 4

2.5 The evolution of vocal complexity ....................................................................... 5

    2.5.1 Infant vocal development ............................................................................. 5

    2.5.2 Vulnerability of human infants .................................................................... 6

    2.5.3 Kindchenschema ......................................................................................... 6

    2.5.4 Parental selection hypothesis ...................................................................... 7

2.6 Comparative research ............................................................................................ 7

    2.6.1 Parent-offspring conflict .............................................................................. 7

    2.6.2 Begging displays .......................................................................................... 8

    2.6.3 Carryover to mating context ....................................................................... 8

    2.6.4 Babbling-like vocal behavior in a non-human primate ................................ 9

    2.6.5 Unpredictable acoustic features attract attention ....................................... 9

2.7 Potential influence of hormones on caregiving .................................................... 10

    2.7.1 Oxytocin and prolactin ............................................................................... 11

    2.7.2 Estradiol ....................................................................................................... 12

    2.7.3 Testosterone ................................................................................................. 13
2.7.4 Physical contact

2.8 Heart rate variability
   2.8.1 Respiratory sinus arrhythmia and vagal tone

2.9 Summary and experimental setup
   2.9.1 Goals

2.10 Hypotheses

Chapter III. METHODS
   3.1 Participant selection
   3.2 Stimuli
   3.3 Procedure
   3.4 Measurement and scoring of physiological and subjective response data
      3.4.1 HRV measures
      3.4.2 Facial expression measures
      3.4.3 Subjective responses

Chapter IV. RESULTS
   4.1 Dependent and independent variables
   4.2 Statistical analysis
   4.3 Comparisons between stimuli
   4.4 Comparisons between mothers and non-mothers
   4.5 Comparisons between fathers and non-fathers
   4.6 Comparisons between all groups (additional analyses)
   4.7 Summary

Chapter V. DISCUSSION
   5.1 Accepted hypotheses
      5.1.1 Evolutionary relevance of human facial expressions
      5.1.2 Conscious preference for canonical infant vocal signals: subjective ratings
5.1.3 Possible hormonal and parental/nonparental influences on potential caregivers 43

5.1.4 Other effects possibly related to hormone levels 44

5.2 Rejected hypotheses 45

5.2.1 HRV measures 45

5.2.2 Possible explanations 46

5.2.3 Facial expressions 47

5.2.4 Possible explanations 47

5.2.5 Subjective ratings 48

5.2.6 Possible explanations 48

5.3 Limitations and future directions 50

Chapter VI. APPENDICES 53

6.1 Questionnaire for females 53

6.2 Questionnaire for males 55

6.3 Self-assessment form 57

Chapter VII. REFERENCES 59
List of Tables and Figures

Tables

Table 1. Differences in Heart Rate Variability (HRV) between Stimuli
Table 2. Non-parametric Testing Comparing Facial Expressions of Happiness between Stimuli
Table 3. Non-parametric Testing Comparing Ratings of Happiness between Stimuli
Table 4. Non-parametric Testing Comparing Ratings of Least Liked/Most Liked between Stimuli
Table 5. Differences between Breastfeeding Mothers (BF) and Non-mothers (NMH and NML) for Babbling
Table 6. Differences between Breastfeeding Mothers (BF) and Non-mothers (NMH and NML) for Cooing
Table 7. Differences between Breastfeeding Mothers (BF) and Non-mothers (NMH and NML) for Crying
Table 8. Differences between Breastfeeding Mothers (BF) and Non-mothers (NMH and NML) for Fussing
Table 9. Differences between Fathers and Non-fathers for Babbling
Table 10. Differences between Fathers and Non-fathers for Cooing
Table 11. Differences between Fathers and Non-fathers for Crying
Table 12. Differences between Fathers and Non-fathers for Fussing
Table 13. Comparisons of Heart Rate Variability (HRV) values adjusted by baseline values
Table 14. Baseline Heart Rate Variability (HRV) values

Figures

Figure 1. Self-rated happiness scales for all stimuli
Figure 2. Mean ratings on the Least liked/Most liked scale for all stimuli
I. INTRODUCTION

Vocal flexibility and complexity are unique and profound differences that may possibly separate human from non-human species. It is conceivable that in our evolutionary history a small increase in vocal complexity caused a significant increase in the ability to vocalize with more complexity (Locke, in press; Harvey & Arnold, 1982). Parental investment (PI) theory predicts that parents should invest more in high-quality offspring (Trivers, 1972). Parent-offspring conflict and the associated need of the offspring to elicit parental attention (Trivers, 1974) may be a contributing factor to the need for complexity and variability in vocal output. Complexity and variability in the vocal output of offspring may be an essential component of signaling health and eventual reproductive success to caregivers. Listeners able to subconsciously pick up on differences in vocal ability, particularly vocalizations of infants and children competing for care and attention in order to survive, may have received fitness information (Locke & Bogin, 2006; Locke, 2008, 2009). Vocalizations provide listeners with many cues, including multiple physical traits (e.g., Pisanski et al., 2016). A possible general attraction to greater complexity of sound across species both in a mating context (e.g., Ryan & Rand, 1990; Catchpole & Slater, 1995; Mountjoy & Lemon, 1996; Anolli & Ciceri, 2002; Puts, 2010; Charlton, Filippi & Fitch, 2012) and within the context of infant care (e.g., Bloom & Lo, 1990; Bloom et al., 1993), may suggest that infants vocalizing using interesting patterns will attract the attention of caregivers and signal fitness.

Infants with more complex babbling structure are perceived more favorably than those with less complex vocalizations (Bloom & Lo, 1990; Bloom, D’Odorico, & Beaumont, 1993). In fact, there may be a general preference for complexity and variability in vocalizations; adult females generally prefer more complex than simple musical sounds (Charlton, Filippi, & Fitch, 2012). Babbling at age 7-10 months has been suggested to be a fitness indicator; infants who babble on time may be in better health (Locke, 2004), because late babbling is associated with developmental issues including sensory, cognitive and language disorders (Oller, Eilers, Neal, & Cobo-Lewis, 1998; Oller, 2014; Patten, Belardi, Baranek, Watson, Labban, & Oller, 2014) and high-risk populations of late babblers show significantly lower expressive vocabulary scores at 18, 24 and 30 months (Oller, Eilers, Neal, & Schwartz, 1999).
The goal of my research was to gain insight into any evolutionary component leading to the development of vocal complexity in infants. The review of the literature will explain the possibility of vocal complexity in human communication evolving as a subconscious strategy used by infants to gain attention from caregivers. A brief introduction to evolutionary developmental linguistics and a parental selection hypothesis, as well as a summary of comparative research, will provide a foundation for understanding the attraction of potential caregivers to vocal complexity. Potential influence of hormonal states on caregiving will also be discussed. Oxytocin and prolactin associated with breastfeeding, estradiol levels in menstruating females, and testosterone in males may each affect caregiver responses to infant vocalizations in different ways.

II. REVIEW OF LITERATURE

2.1 Evolutionary developmental linguistics

Evolutionary developmental linguistics (EDL) (Locke, 2009), gives us a theoretical framework for studying the evolution of language. EDL highlights the importance of development in evolution (Locke, 2009), and describes the significance of the four human stages of development: infancy, childhood, juvenility, and adolescence (Locke, 2009; Locke & Bogin, 2006), and the role they may have played in the evolution of language. Locke describes an evo-devo-evo process which includes the evolution of developmental aspects that either alone, or together with other changes in the environment, facilitated the appearance of language in our species (Locke, 2012). Variation in vocal complexity may have appeared in the infancies of our ancestors, with parental selection shaping preferred behaviors that were later evaluated by peers and others (Locke, 2012). Infancy spans the interval from birth to approximately 30—36 months, at which time most infants are typically weaned in traditional (e.g., hunter-gatherer) societies. General characteristics of human infancy include helplessness, rapid growth, breastfeeding as the main source of nutrition, and eruption of primary teeth.

2.1.1 Developmental plasticity
Developmental plasticity is a phenotypic plasticity that organisms possess to allow for the development of functional phenotypes despite variation and environmental change through accommodation (West-Eberhard, 2003). Changes in development create novel phenotypes, which are then subject to natural selection (Lickliter, 2008). New behaviors appearing during early stages of development, a time of significant plasticity and variability, continue in some form. During evolution these earlier occurring traits that facilitated later ones were reinforced automatically by selection (Garstang, 1922; West-Eberhard, 2003; Locke, 2012).

2.2 Parental investment

Parental investment is any investment by the parent (e.g., feeding, attention, grooming, holding, guarding, etc.) in an individual offspring that increases its’ chance of survival, and therefore reproductive success, either at the cost of the parent’s ability to invest in or even produce other offspring (Trivers, 1972). According to Trivers’ theory of parental investment, parents should invest more in offspring that they expect to survive. For example, in bird species that bring single food items to begging nestlings, and only one gets fed per visit, there is a relatively stronger investment in larger and healthier nestlings. Offspring are expected to demand more parental investment than is optimal for the parent, and may use manipulation to gain more attention and care than is necessary for survival. This conflict between parents and offspring continues throughout the period of parental investment, and may intensify when siblings are present (Trivers, 1974). It can also increase at the end of the period of parental investment leading to weaning conflict. Behavioral tactics and manipulation may involve vocalizations; an example of vocalizations with manipulative intent in birds is exaggerated begging calls.

2.3 Cooperative breeding

The cooperative breeding hypothesis takes into account traits that are shared by humans and other cooperative breeders (Hrdy, 2007). Approximately 3% of mammals and 12% of bird species breed cooperatively (Hrdy, 2007). The only true cooperative breeders among primates are humans, marmosets and tamarins, although primates generally are extremely social and do exhibit some degree of shared
caregiving along with an attraction to infants (Hrdy, 2007). The prolonged dependence that coincides with cooperative breeding creates completely dependent and vulnerable human newborns that need to successfully elicit care from mothers and other potential caregivers. Add to this the substantial need of older infants and juveniles for care, and the groundwork is laid for caregiver responses that are contingent and care more conditional.

Hominids became bipedal over four million years ago (Ward, 2002), and approximately 2.4 million years ago there was a trend toward increasing brain size (Wittman & Wall, 2007). A narrowed pelvis occurring with bipedalism meant a smaller birth canal, and with an enlarged fetal head size there was a shift of brain development into the postnatal period in order to help solve this obstetrical dilemma (Washburn, 1960). This means newborns were extremely helpless, and evolution would have favored babies able to elicit care and parents able to read cues to infant state (Locke & Bogin, 2006; Locke, in press). Cooperative breeding would have helped to alleviate some of the heavy burden of parental investment required to care for such needy offspring, and in turn lessen parent-offspring conflict. Parents and other caregivers in the group (e.g., aunts, grandmothers, and older siblings) would be rewarded via inclusive fitness for the ability to detect fitness cues in vocal signals of infants, as evolutionary success relies on reinforcement of genes.

2.4 The environment of evolutionary adaptedness

The environment of evolutionary adaptedness (EEA) is the combination of selection pressures that formed the design of an adaptation; it is not one place or time, as different traits were formed at various times depending on the problem(s) that needed to be solved (Tooby & Cosmides, 1990). Behaviors observed today exist in part because they solved adaptive problems in the ancestral environments in which humans evolved. When speculating about evolution of human traits it is necessary to think in terms of problems that needed to be solved in traditional hunter-gatherer foraging societies. Hadza women, hunter-gatherers in Tanzania, whose lifestyle resembles that of our ancestral generations, take nursing infants along when foraging for fruit and digging wild tubers, yet toddlers are left in camp and cared for by older siblings, maternal grandmothers, and other non-related caretakers (Marlowe, 2005). In
Ju/'hoansi (!Kung) society, former foragers of Botswana and Namibia, children grow up in a public space in close contact with other village members (e.g., Konner, 1976; Draper & Hames, 2000). The environment that hominids evolved in was very different from that of modern humans. Over 99% of our species' evolutionary history was spent living in small nomadic groups who gathered plants and hunted animals daily (Tooby & Cosmides, 1990). Therefore, infants and children in the band grew up together, and, even without siblings, there was still the possibility of competition with others for care from multiple potential caregivers in the group.

2.5 The evolution of vocal complexity

2.5.1 Infant vocal development

During the first two months of life human infants produce quasivowels, sounds emitted with normal phonation yet with a vocal tract at rest (Oller, 1995). The emergence of speech abilities in infants is distinguished from vegetative sounds (burping, coughing, etc.) and fixed vocal signals (e.g., crying, laughter, etc.). By approximately two to three months of age, in the primitive articulation stage, infants "goo", or coo, and thus commence very limited articulation while vocalizing (Oller, 1995). This stage of going is followed by the expansion stage, consisting of full vowels, raspberries, and marginal babbling, and finally by the canonical stage with well-formed canonical syllables and reduplicated sequences (Oller, 1995). Canonical babbling consists of rapid formant transitions from consonant-like elements to vowel-like elements, something that does not occur in marginal babbling (Oller et al., 1999; Oller, 2014).

Canonical babbling typically appears from six to ten months, at the same time as sitting and crawling (Cobo-Lewis, Oller, Lynch, & Levine, 1996). It does not include squealing, growling, or raspberries, and, by definition, must have at least one full vowel-like element and one consonant-like element; examples include [ba], [nunu], [dada] (Oller et al., 1999). Canonical babbling includes non-reduplicated and reduplicated productions and caregivers may occasionally perceive these forms as words, such as dada or mama (Oller, 1980). As repetitive sound clusters decrease and variety of consonants and vowels increase, there is often a shift from canonical to variegated babbling with different combinations of syllables (e.g., [bada]) (Mitchell & Kent, 1990), however reduplicated and variegated
productions continue to co-occur from approximately ten months onward (e.g., Smith, Brown-Sweeney, & Stoel-Gammon, 1989). In addition to the possibility that babbling may provide vocal practice, babbling is also capable of eliciting social attention, which we will address now.

2.5.2 Vulnerability of human infants

As discussed previously, human infants are born completely dependent on others for care, and they remain nutritionally dependent much longer than nonhuman apes. It has been argued that cooperative breeding in our evolutionary history was necessary for child survival, allowing for production of such costly offspring without changing inter-birth intervals (Hrdy, 2007). Mothers needed support, and conditional maternal investment prompted newborns to successfully elicit care from mothers and older offspring to attract attention of others (Hrdy, in press). Hominin infants and children needed to monitor mothers and other potential caretakers and appeal to them in various ways – via physical attractiveness, gestures, and vocalizations. If there was not another sibling to compete with, there was still the possibility of an unborn sibling who would decrease resources. In a cooperative breeding species, all dependent young members in the tribe could potentially pose a threat as well.

2.5.3 Kindchenschema

Darwin believed that infants possess qualities that prompt adults to respond to and care for them in order to increase individual fitness via reproductive success. Konrad Lorenz described “kindchenschema” as infantile features including large eyes, a big head, chubby cheeks and body, and a small nose and mouth, which are perceived as cute and motivate human caregiving (Lorenz, 1971). A recent fMRI study in women without children showed that baby schema activate the nucleus accumbens, a component of the mesocorticolimbic system that mediates reward processing, motivation and pleasure, suggesting a neurophysiologic mechanism that promotes human caregiving (Glocker et al., 2009). Both parents and non-parents displayed brain activity evaluated by magnetoencephalography (MEG) in the medial orbitofrontal cortex (mOFC), an area implicated in reward behavior, in response to attractive infant faces but not to attractive adult faces (Kringelbach et al., 2008). These findings suggest a neural basis for
this imperative evolutionary process for stimulation of care for infants. This appeal of infant characteristics gives strong support to this claim, considering these findings (Hrdy, in press).

2.5.4 Parental selection hypothesis

What may or may not belong to a vocal kindchenschema? High frequency hyperphonated infant cries are rated by adults as more aversive, distressing, urgent, arousing, and sick than “normal” phonated cries, and skin conductance levels are higher for these abnormal cries as well (Crowe & Zeskind, 1992). As mentioned earlier, adults prefer infant vocalizations that are syllabic versus vocalic, and rate the infants producing the more complex sounds as more pleasant, friendly, fun, likeable and cuddly (Bloom & Lo, 1990; Bloom et al., 1993). The parental selection hypothesis (Locke, 2006) suggests that hominin parents made decisions about allocation of care partially based on their infants’ vocal behavior. Abnormal, inconsolable and/or constant crying may have signaled atypical development and reduced care to infants, while cooing and babbling increased attention and social interaction (Locke, 2006). A baby who was able to vocalize in novel ways would have captured the attention of caregivers, and may have signaled fitness and the ability to learn complex behaviors (Locke, 2006, 2008). Over time these slight increases in vocal variability may have facilitated the evolution of vocal control and complexity.

2.6 Comparative research

2.6.1 Parent-offspring conflict

The parental selection hypothesis (Locke, 2006) can be tested in humans as well as other species with parental care and vocal young. As birth spacing in human history decreased, competitors for care increased. Comparative research provides a practical opportunity to observe solutions to issues such as parent-offspring conflict (Trivers, 1974) and sibling competition, and may provide evidence of complexity and variability in vocal output of offspring as a way to elicit parental attention. Convergent evolution (e.g., vocal flexibility in birds, cetaceans, pinnipeds) can help us to understand the types of problems that certain traits are designed to solve (Fitch, 2000). An elaborate vocal repertoire in birds helps to attract a mate (e.g., Catchpole & Slater, 1995) and defend territories (e.g., Krebs, 1977). Male
European starlings (Sturnus vulgaris) with more complex songs and larger song repertoires are chosen more often by females, and these males also happen to be in better health (Mountjoy & Lemon, 1996). Male Tungara frogs (Engystomops pustulosus) with more complex calls are preferred by females (Ryan & Rand, 1990).

2.6.2 Begging displays

Passerine nestlings use complex begging displays involving posturing, revealing brightly colored gapes, jostling for the best position in the nest, and calling loudly to solicit food (e.g., Redondo & Castro, 1992; Pycraft, 1907; McRae, Weatherhead, & Montgomerie, 1993; Haskell, 1999; Kilner, 2002). Across species, there are also facial, gestural and whole body behaviors that alone or in combination work to elicit the care and attention of caregivers. In avian species, the vocal component seems to become more reliable than other components of the display with nestling age (Kilner, 2002), perhaps due to increased predation cost of increased calling rate (Haskell, 1994) or physical constraints of calling changing with age (Kilner, 2002). Begging is thought to aid in resolving parent-offspring conflict (Godfray, 1995; Trivers, 1974). The benefits of gaining parental attention swiftly may have selected for the most distinctive features of the begging display (Dawkins & Guilford, 1997). Caregivers’ ability to discern honest signals of fitness also plays an important role in the evolution of vocal complexity.

2.6.3 Carryover to mating context

The big brown bat (Eptesicus fuscus) uses different types of vocalizations when in the presence of their mothers presumably to attract attention (Monroy, Carter, Miller, & Covey, 2011). The frequency range of these vocal signals increases throughout development, and adult big brown bats use similar vocalizations during interactions including mating (Monroy et al., 2011). It seems that the bat pup vocal repertoire develops into adult social vocalizations with similar structure yet different uses (Monroy et al., 2011). There may be a link between use of vocal complexity to attract parental attention and later on in mating context to attract attention of potential mates. Some species of birds incorporate begging calls into adult song that is used to attract mates (e.g., Payne, Payne, & Woods, 1998).
Females also use song learning precision in males as an honest cue to developmental history and quality (Lachlan & Nowicki, 2012; Nowicki & Searcy, 2014). Songs learned by well-nourished male swamp sparrows elicited significantly higher levels of courtship display from females than the songs of under-nourished male swamp sparrows. Well-nourished males produced longer songs with a higher trill rate, greater stereotypy, and more notes per syllable (Peters, Kipper, Nowicki, & Searcy, 2011). It is possible that there is also a connection between signal complexity and mating success in primates. In gelada monkeys, females display a preference for the more complex vocalizations of males (Gustison & Bergman, 2016).

### 2.6.4 Babbling-like vocal behavior in a non-human primate

We may also find similarities in traits of more closely related non-human primates. Infant and juvenile pygmy marmosets emit long sequences of mixed calls that are associated with attention and social interactions with other group members (Snowdon & Elowson, 2001). Parents and other caretakers respond quickly to these sounds and approach and carry the infants. Infants with more complex structure in vocalizations attain adult levels of vocal structure sooner than those with less complex vocalizations (Snowdon & Elowson, 2001). It may be possible that these infants capable of vocal variability are also gaining more attention than others not as skilled in vocal behavior.

### 2.6.5 Unpredictable acoustic features attract attention

Another means of eliciting caregiver attention may be through unexpected acoustic features in vocalizations, serving to gain attention and decrease habituation. Nonlinear vocal phenomena are highly complex and unpredictable vocalizations that change normal spectral structures (Townsend & Manser, 2010). In the highly altricial giant panda cub, an increase in nonlinear phenomena (NLP) in vocalizations was associated with high-arousal contexts (Stoeger, Baotic, Li, & Charlton, 2012). An increase in “chaos”, subharmonics, as well as increased call duration, rate and mean F0 were noted in these instances in which it may be quite important to elicit attention from and convey arousal state to caregivers.

Acoustic properties of vocalizations can increase attention (Davis, 1984) and NLP may also
function to make calls less predictable and less likely to reduce responding due to habituation (Fitch, Neubauer, & Herzel, 2002). Cries and screams of human infants contain unpredictable spectral and temporal characteristics that elicit attention from caregivers (Soltis, 2004), which is also seen in some nonhuman primates (e.g., Todt, 1988) and elephants (Stoeger, Charlton, Kratochvil, & Fitch, 2011). In addition to lack of predictability in vocalizations, it is also thought that calls of longer duration and with higher F0 should elicit greater attention from receivers (Protopapas & Lieberman, 1997; Blumstein, Richardson, Cooley, Winternitz, & Daniel, 2008). Examples of how other species (e.g., insects and frogs) make themselves heard when others are also calling may give us some clues as to how nestlings may do the same (Leonard, Horn, & Parks, 2003; Wright & Leonard, 2002). In addition to increasing duration, intensity, output or rate of calls, perhaps offspring change frequency or overall form (e.g., shape on spectrogram) to make their calls less similar to siblings (Horn & Leonard, 2002).

### 2.7 Potential influence of hormones on caregiving

There is a possibility for a general preference for vocal variability in infant signals, or it may be a specific preference that comes with being a caregiver. Experience and/or hormones of parenthood may affect sensitivity to an increase in variability or complexity in infant vocalizations. Perhaps the hormone shift that occurs in parents heightens perception and draws attention to slight increases in complexity and variability as a fitness indicator in offspring.

For all of our evolutionary history as hunter-gatherers infants were breastfed; adaptations to this method of feeding over millions of years has resulted in profound physiological effects on both mother and infants (Macadam & Dettwyler, 1995). In hunter-gatherer societies the main source of nutrition for well over the first year of life is breast milk (Truswell, 1977). Breastfeeding involves the release of oxytocin and prolactin, and these hormones are not only important in the milk ejection reflex and the production of milk, but also in attachment and bonding for both mothers and fathers (e.g., Gordon, Zagoory-Sharon, Leckman, & Feldman, 2010).
Lowering of testosterone is another important factor in paternal caregiving and bonding. A decrease in testosterone once men have had children, with the most dramatic decrease for men who spend more than three hours daily in childcare (Gettler, McDade, Feranil, & Kuzawa, 2011), may be an evolutionary adaptation to help fathers respond more sensitively to the needs of their offspring.

It has been suggested that highly altricial human infants need to elicit care via multiple signals, including vocalizations (e.g., Soltis, 2004; Locke & Bogin 2006; Locke, 2009), and both mothers and fathers experience hormonal shifts that may prime them to caregiving (Hrdy, 2007). Hormones may play an important role in caregiver responses to infant vocal signals.

2.7.1 Oxytocin and prolactin

Hormones associated with breastfeeding may have evolved not only to produce milk, but also to assist women in the ability to respond sensitively to their infants. Mothers who breastfeed exclusively have significantly higher levels of oxytocin and prolactin than those who give supplemental formula (Uvnäs-Moberg, Widström, Werner, Matthiesen, & Winberg, 1990).

The highest levels of oxytocin in breastfeeding mothers are found during the pre-feeding period, regardless of the time of feeding (White-Traut, Watanabe, Pournajafi-Nazarloo, Schwertz, Bell, & Carter, 2009). Oxytocin and prolactin keep mothers calm, relaxed and ready to care for their babies (e.g., Uvnas-Moberg, 2003). Breast-feeding mothers respond more quickly to infant crying (Bernal, 1972), so it seems possible that oxytocin is one of the contributing factors to increased maternal interaction and bonding to the infant (Uvnas-Moberg & Petersson, 2005).

Breastfeeding is also tied to stronger neural maternal response to own infant’s cry (Kim et al., 2011); fMRI was used to reveal a correlation between breastfeeding and greater response in brain regions implicated in empathy and mother/infant bonding in the early postpartum period, along with higher maternal sensitivity at three to four months postpartum. Mothers who breastfeed also show a differential pattern of cardiac response to infant stimuli, and are more inclined to want interaction with infants in subjective responses to their infants’ facial expressions of varying emotions (Wiesenfeld, Malatesta, Whitman, Granrose, & Uili, 1985). The cardiac rate of the nursing mothers was characterized by an initial
deceleration followed by an acceleration above baseline, yet the formula feeding mothers’ heart rates only showed a sharp acceleration comparatively.

What is interesting about this pattern of deceleration (parasympathetic activation) followed by acceleration (parasympathetic withdrawal or sympathetic activation) is that it is the cardiac orienting response, which is associated with greater attentional deployment toward novel, highly-arousing, or motivationally salient stimuli. Ratings of subjective emotional responses included self-reports of happiness, sadness, anxiety, irritability and helplessness, along with readiness to pick up the baby. Skin conductance response has been found to be greater with more stimulating emotions (Khalfa, Isabelle, Jean-Pierre, & Manon, 2002). Interestingly, high-empathy women have larger electrodermal responses, more extreme happiness and sadness reactions, and also report a stronger desire to pick up infants when shown silent videos of smiling, neutral and crying infants (Wiesenfeld, Whitman, & Malatesta, 1984). Oxytocin has been shown to increase empathy (Bartz et al., 2010; Hurlemann et al., 2010).

Similar hormonal effects are not only observed with method of feeding or oxytocin administration. Vaginal delivery involves a natural release of oxytocin, which has a profound effect on maternal behavior in animals (Kendrick, 2000). In the early postpartum period following a vaginal delivery (VD) versus elective caesarean section delivery (CSD), an fMRI study showed VD mothers' brains were significantly more responsive than CSD mothers' brains to recordings of their own baby cry, observed in the superior and middle temporal gyri, superior frontal gyrus, medial fusiform gyrus, superior parietal lobe, as well as regions of the caudate, thalamus, hypothalamus, amygdala and pons (Swain, Tasgin, Mayes, Feldman, Todd Constable, & Leckman, 2008). Research in non-human animals has shown that networks of hypothalamic-midbrain-limbic-paralimbic-cortical circuits act along with hormones including oxytocin and prolactin in caregiver response to offspring (e.g., Leckman & Herman, 2002; Numan, 2007).

2.7.2 Estradiol

The female reproductive hormone estrogen includes estradiol. Maternal responsiveness to young increases during late pregnancy in sheep and rats along with an increase in circulating estradiol and in the estradiol to progesterone ratio (Rosenblatt, Mayer, & Giordano, 1988; Poindron and Levy,
A six-week treatment with estradiol significantly increased the amount of infant handling in ovariectomized female rhesus macaques (Maestripieri & Zehr, 1998). Human mothers reporting highest attachment feelings in the postpartum period are those who show an increase in the estradiol to progesterone ratio mid to late pregnancy (Fleming, Ruble, Krieger, & Wong, 1997). Estrogen and progesterone levels decrease right after birth, generally to pre-pregnancy levels, however, and oxytocin and prolactin increase during lactation. Postpartum depression, which occurs in approximately one out of seven women after childbirth along with the significant decrease in estrogen at birth, is successfully treated with transdermal estradiol in some patients (Moses-Kolko, Berga, McCord, & Wisner, 2009). Changes in estrogen also occur in non-pregnant females. Two estradiol peaks occur during normal menstrual cycles; the first (and larger) one precedes the second by five days or more, with progesterone peaking within two days (+ or -) of the second estradiol peak (Gandara, Leresche, & Mancl, 2007). Shifts in estradiol levels that occur with menstruation, pregnancy, childbirth, and menopause may affect mood so intensely that it causes depression (e.g., PMS, postpartum depression, menopausal depression). With the potential of such powerful effects, hormones should also play a role in perception of vocal signals (e.g., mating calls, maternal responsiveness to infant vocal signals). Hormonal correlates do not occur in isolation, however, and a complex mix of contributing factors, such as maternal experience and social and physical contexts, may interact, accentuate or mask these associations (e.g., Fleming, Ruble, Flett, & Shaul, 1988; Corter & Fleming, 1995).

2.7.3 Testosterone

Male testosterone levels are positively correlated with mating effort via male-male competition and mate-seeking behavior (e.g., Roney, Mahler, & Maestripieri, 2003), and decreased levels of testosterone have been linked with affiliative pair bonding and paternal care (e.g., Gray, Yang, & Pope, 2006). In fact, males need not be married to display this adaptive lowering of testosterone; males who are in committed romantic relationships showed the same 21% decrease in salivary testosterone levels as married men (Burnham, Chapman, Gray, McIntyre, Lipson, & Ellison, 2003). Fathers have significantly lower testosterone levels than both unmarried males and married non-fathers (Gray et al., 2006), and
expectant fathers have lower testosterone levels than non-fathers (Berg & Wynne-Edwards, 2001). Males with lower testosterone levels feel more sympathy and need to respond to infant cries, and fathers with higher prolactin levels are more alert and positive in response to infant cries (Fleming, Corter, Stallings, & Steiner, 2002).

Experience also seems to play a role in paternal responsiveness, as experienced fathers had greater increase in prolactin when listening to cries compared to first-time fathers (Fleming et al., 2002). Many studies of testosterone in males have been cross-sectional, raising the question of whether becoming a father decreases testosterone or males with lower testosterone are more likely to become fathers. Longitudinal data have shown that high testosterone predicts mating success, and the hormone declines significantly with fatherhood, and most drastically in men with heavier childcare involvement (e.g., Gettler et al., 2011).

2.7.4 Physical contact

Parental touch also affects hormone levels in mothers and fathers. Mothers who display high levels of affectionate physical contact with their infants have an increase in oxytocin levels, yet mothers showing low levels of affectionate contact do not (Feldman, Gordon, Schneiderman, Weisman, & Zagoory-Sharon, 2010). Fathers also show an increase in oxytocin after high levels of stimulatory contact with their babies, indicating typical variations in styles of maternal and paternal care (Feldman et al., 2010). The authors suggest the possibility of encouragement of paternal physical bonding activities to assist in issues arising from decreased maternal-infant contact, such as postpartum depression.

Administration of oxytocin nasal spray to fathers increased respiratory sinus arrhythmia (RSA), social gaze, and touch during interactions with their infants, as well as the infants’ salivary oxytocin, RSA, and social behaviors such as exploration, social gaze and social reciprocity (Weisman, Zagoory-Sharon, & Feldman, 2012). RSA involves fluctuations in heart rate that are linked with breathing. The changes in infants’ hormonal state, RSA response, and bonding behaviors occurred without hormonal administration to the infant. Successful infant vocal requests for attention seem to reinforce physical bonding and beneficial hormonal responses.
2.8 Heart rate variability

2.8.1 Respiratory sinus arrhythmia and vagal tone

Heart rate variability (HRV) has been associated with acute emotional responses. HRV has become the accepted term to quantify variation over consecutive cardiac cycles. Other terms may include cycle length variability, heart period variability, RR variability (in which R is a point corresponding to the peak of the QRS complex - the combination of the Q, R, and S waves seen on a typical electrocardiogram, or ECG; and RR is the interval between consecutive Rs), RR interval tachogram, yet all refer to variability in the interval between successive beats, and not merely heart rate (Task Force, 1996). The length of the time between beats, or inter-beat intervals, naturally fluctuates with the respiratory cycle: during exhalation, the inter-beat interval lengthens due to increased activity of the parasympathetic nervous system, and during inhalation, the inter-beat interval shortens. This fluctuation is known as respiratory sinus arrhythmia (RSA), which is a natural, normal fluctuation that is not pathological. The parasympathetic nervous system (PNS) and the sympathetic nervous system (SNS) are both branches of the autonomic nervous system, which controls largely involuntary processes such as breathing, digestion, and heart rate. The SNS is a fast acting system that gets our body ready for action ("fight or flight") and the PNS is a slower moving system that helps bring our body to a state of calm and allows it to relax and repair ("rest and digest"). The respiratory component in heart rhythms mediated by the vagus nerve provides an indirect measure of parasympathetic activity (Porges, 1995). Resting RSA has been assumed to indirectly measure vagal tone, or passive activity of the tenth cranial nerve (the vagus nerve) (Grossman, Stemmler, & Meinhardt, 1990; Saul & Cohen, 1994). This view of RSA providing information on the electrochemical activity of the vagus nerve has been challenged, however (Pyetan & Akselrod, 2003), therefore it is more accurate to say it is an estimate of parasympathetically mediated heart rate variability rather than an index of vagal tone (Appelhans & Luecken, 2006). Higher RSA is associated with more ideal maternal care and infant social engagement (Moore & Calkins, 2004; Moore, Hill-Soderlund, Propper, Calkins, Mills-Koonce, & Cox 2009; Feldman et al., 2010). The parasympathetic branch of the autonomic nervous system is important in bonding in mammals (Porges,
Heart rate deceleration and corresponding increases in RSA are thought to be associated with relaxation and positive mood states, and decreases in RSA are assumed to indicate physiological responses to stressors and negative mood states (Beauchaine, 2001; Porges, 1995; Thayer & Lane, 2000).

2.9 Summary and experimental setup

The capacity for vocal flexibility permits great linguistic creativity in our species. Early in evolution, it is possible, as mentioned earlier, that a slight increase in vocal complexity, once selected for, could eventually cause an expansion of vocal ability (Harvey & Arnold, 1982; Locke, in press). It is speculated that listeners able to subconsciously evaluate these increasingly complex vocalizations would receive fitness information (Locke, 2008). Considering that evolution of organisms involves a change during development, it makes sense that vocalizations of infants and children would be an important factor in our evolutionary history of speech and language (Locke & Bogin, 2006; Locke, 2009). Offspring with greater fitness, and therefore greater potential reproductive value, may receive more care and attention from caregivers.

EDL identifies the environmental changes that caused phenotypic variation as an increase in helplessness in human infants, along with heightened competition for attention from alloparents in traditional foraging societies (Locke, 2009). Parental selection may have reinforced some of this variation, thereby increasing reproductive success (Locke, 2006). The remodeled stage of infancy, and the ability of infants to vocalize with greater vocal complexity, would be the first step in a sequence of adaptive responses to environmental changes, ultimately helping to enhance fitness and human perception, memory and control systems that are necessary for modern human speech (Locke, 2012).

2.9.1 Goals

The goal was to examine adult evaluation of and physiological response to infant vocalizations of varying complexity. Independent variables included audio recordings of infant vocalizations including crying, fussing, cooing (non-canonical vocalizations) and babbling (canonical vocalizations). Dependent
variables measured were cardiac and facial activity, along with self-report of emotional reactions to the vocal signals. It was hypothesized that adults would show more positive responses in all measures to babbling as compared to the other three classes of vocal stimuli. It was hypothesized that higher HRV, thought to be linked with positive mood states (Beauchaine, 2001; Porges, 1995; Thayer & Lane, 2000) would be associated most with babbling as compared to the other types of infant vocal signals. Facial expressions of pleasure would be associated most with babbling. Human emotions are associated with innate and universal facial expressions (Darwin, 1872/1965; Izard, 1994). Subjective emotional responses to the babbling vocalizations would be more likely to include desire to interact with infant, higher self-ratings of happiness, and the vocal recordings of babbling should be given the highest ratings of 'most liked'.

It was also expected that mothers would respond differently from non-mothers to infant vocal signals; increased HRV would be associated with cooing and babbling, with concordant facial expressions (e.g., happiness for babbling, sadness for crying). Non-mothers, especially at the low point in estradiol levels in their menstrual cycles (Rosenblatt, Mayer, & Giordano, 1988; Poindron and Levy, 1990; Fleming, Ruble, Krieger, & Wong, 1997; Maestripieri & Zehr, 1998), were expected to show more stress and negative mood states when listening to fussing and crying, with decreased HRV overall. Subjective emotional responses of the mothers, as compared to non-mothers, to the vocalizations were expected to be more likely to include desire to interact with infant. Breastfeeding mothers were expected to demonstrate the most positive/sensitive cardiac, facial and subjective responses (e.g., see Wiesenfeld et al., 1985; Kim et al., 2011).

Fathers were expected to respond differently from non-fathers to the infant vocal signals (e.g., see Fleming et al., 2002; Gray et al., 2006; Gettler et al., 2011); increased HRV measures were assumed to be associated with cooing and babbling, with corresponding facial expressions (e.g., happiness during babbling). Non-fathers were expected to show more stress and negative mood states when listening to fussing and crying, with decreased HRV overall. Subjective emotional responses of fathers, as compared to non-fathers, to the vocalizations were assumed to be more likely to include desire to interact with infant.
2.10 Hypotheses

H1: Participants will show higher HRV, higher incidence of facial expressions and subjective ratings of happiness, and higher ratings indicating a desire to pick up the baby as well as higher ratings towards the “most liked” end of the scale for the canonical vocalizations as compared to non-canonical vocalizations, fussing, and crying.

H2: Breastfeeding mothers will demonstrate higher HRV, higher incidence of facial expressions of happiness, higher means for subjective ratings of happiness, and higher scores towards the “most liked” end of the rating scale for babbling (canonical) and cooing (non-canonical) as compared to non-mothers at low estradiol and non-mothers at high estradiol time period in their menstrual cycle. Non-mothers at the low estradiol point will show the lowest scores as compared to the other two groups.

H3: Breastfeeding mothers will demonstrate higher HRV, higher incidence of facial expressions of sadness, lower means for self-rated happiness, higher means for subjective ratings of sadness, lower means for self-rated irritation, and higher means for self-rated desire to pick up the baby when the crying and the fussing stimuli were played as compared to non-mothers at low estradiol and non-mothers at high estradiol period in their menstrual cycle. Mothers will also show higher means for self-rated anxiety for the crying stimuli as compared to non-mothers. Non-mothers at the low estradiol point will show the most stress and negative responses to stimuli, including lowest HRV overall, and higher levels of reported irritation to crying as compared to the other two groups.

H4: Fathers will demonstrate higher HRV, higher incidence of facial expressions of happiness, higher means for subjective ratings of happiness, and higher scores towards the “most liked” end of the rating scale for babbling (canonical) and cooing (non-canonical) as compared to non-fathers.
H5: Fathers will demonstrate higher HRV, higher incidence of facial expressions of sadness, lower means for self-rated happiness, higher means for subjective ratings of sadness, lower means for self-rated irritation, and higher means for self-ratings for desire to pick up the baby for the crying and fussing stimuli as compared to non-fathers. Fathers will also show higher means for self-rated anxiety for the crying stimuli as compared to non-fathers.

III. METHODS

3.1 Participant selection

Sixty participants were recruited via fliers posted in local universities in the NYC area, at La Leche League meetings, in new parent community groups, and through word of mouth. All were heterosexual, native speakers of American English, between ages 18 and 35 years. Participants included primiparous exclusively breastfeeding new mothers in the early postpartum period (six months or less), non-mothers at the high estradiol primary peak in their menstrual cycles, non-mothers at the low estradiol point in their menstrual cycles, new fathers, and non-fathers (12 per group). Breastfeeding mothers participated in the listening task during pre-feeding periods, at the time when oxytocin levels would have been the highest (White-Traut et al., 2009). No mother was menstruating at the time of the study. Menstruating non-mothers reported a normal (e.g., 28 day) cycle, and either participated at day 7-10 (low estradiol) or 17-20 (high estradiol) (Gandara, Leresche, & Mancl, 2007). Non-mothers were not taking oral contraceptives, and reported not having used any type of contraceptive pills, implants or patches within the previous six months. Non-mothers and non-fathers were not included if they were currently involved in a long term relationship. New fathers of babies age six months or younger were included if they reported a minimum of three hours daily contact time with their infants. Participants were not included if they were taking medication, reported cardiac or any other serious physical or mental health issues, or had a history of speech, language or hearing disorders. All participants reported normal
hearing. Parents were excluded if they had infants who were premature or ill, or had experienced birth-related complications.

3.2 Stimuli

Infant vocal stimuli included audio recordings of two types of crying (one recording of crying during a circumcision, another recording of crying/fussing in a tired baby), cooing and babbling

Recordings were made using the Buder/Stoel-Gammon vest with a high-fidelity wireless microphone 7 cm from infants’ mouths in order to control mouth-to-microphone distance (Oller & Ramsdell, 2006; Buder & Stoel-Gammon, 2002). Recordings were conducted in 20 minute segments, signals were digitized at sampling rates of 44.1 or 48 kHz, with sixteen-bit quantization, and a signal-to-noise ratio of up to 96 dB (Oller & Ramsdell, 2006). Circumcision cries were recorded in the newborn nursery at Barnes Hospital, St. Louis, MO (Porter, Miller, & Marshal, 1986). Segments of two minute samples representing crying, fussing, non-canonical (cooing) and canonical (babbling) vocalizations were randomly presented. Samples were played twice. The first presentation of stimuli were played to permit the physiological measures of participants, and the second presentation were played to obtain subjective ratings. The author wishes to acknowledge Dr. Kimbrough Oller, Dr. James Green, and Dr. Francine Lang Porter for the use of their recordings of infant vocalizations for the stimuli in this study.

3.3 Procedure

In the first part of the experimental session participants filled out a questionnaire regarding issues such as health status and caregiving experience. Heart rate data were collected using the Polar RS800 heart rate monitor with Polar Pro Trainer software. Participants were seated for a five minute baseline HRV measurement prior to stimuli presentation. A video camera had been placed in the room to monitor and record facial expressions during the first presentation of stimuli. Verbal instructions were given: “Please sit still and look straight ahead while you listen to these sounds. You will be recorded during this experiment.” The participants were informed of the importance of looking up at the video camera and remaining as still as possible. Experimental sessions were conducted in a quiet room free from
background noise; participants wore Sennheiser HD 520 headphones and listened to recordings of infant vocalizations while being continuously monitored for physiological responses during the first presentation of stimuli. A Lenovo G510 i5 with Conexant SmartAudioHD (Dolby Advanced Audio) speakers and a Conexant sound card was used to play recordings. Total harmonic distortion for the headphones is <0.3%, and 0.0003% for the Conexant sound card as measured using RightMark Audio Analyzer 6.4.0 at a sampling mode of 16-bit, 44 kHz. Vocal stimuli were played in random order. After the sequence was completed, participants were instructed to remove the heart rate monitor and the video camera was shut off. The audio recordings were played again with a new randomization, and participants were asked to rate subjective emotional reactions to each stimulus. Participants remained seated at the same table as for the first part of the experiment in which the video recordings and heart rate data were collected, and filled out different self-rating scales for each stimulus. Raters were asked to make use of the entire scale based on variations in vocal samples, and were informed that there were no correct or incorrect answers. The investigator read the rating form aloud and explained that the scales were based on how the participant felt while listening to the sounds, not their perception of the infant’s emotions.

3.4 Measurement of physiological and subjective response data

3.4.1 HRV measures

HRV may be obtained via an electrocardiogram (ECG) or a heart rate monitor (HRM). The Polar RS800 HRM is considered a reliable and valid device to measure HRV (e.g., Quintana, Heathers, & Kemp 2012). Physiological measures were recorded continuously during the first presentation of stimuli using the Polar RS800 heart rate monitor. Heart rate data was processed by Polar Pro Trainer software, and inter-beat intervals (also termed RR intervals) were used to determine HRV. Inter-beat interval (IBI) data was extracted from Polar Pro Trainer software for further analysis.

HRV analysis and artifact removal was completed using ARTiiFACT, a software tool for processing electrocardiogram and IBI data with both automated and manual artifact detection and correction (Kaufmann, Sütterlin, Schulz, & Vögele, 2011). ARTiiFACT provides time- and frequency-
based HRV analyses and descriptive statistics. Appropriate handling of artifacts in IBI data is imperative; a single artifact may cause unreliable HRV results (Berntson & Stowell, 1998, Kaufmann, Sütterlin, Schulz, & Vögele, 2011). HRV measures consisted of SDNN, RMSSD, NN50 and pNN50. SDNN is the standard deviation of NN (or “normal-to-normal” IBI) values, RMSSD is the square root of the mean squared difference of successive NN intervals, NN50 is the number of pairs of adjacent NN intervals differing by more than 50 ms, and pNN50 is the proportion derived by dividing the NN50 by the total number of NN intervals (Appelhans & Luecken, 2006). SDNN reflects overall HRV, while RMSSD, NN50 and pNN50 are believed to represent parasympathetically mediated HRV (Task Force, 1996). HRV, SDNN, RMSSD, pNN50 and RSA have been found to be highly correlated with one another (Allen, Chambers, & Towers, 2006). Thus, SDNN, as overall HRV, and RMSSD as parasympathetically mediated HRV, were included in the statistical analysis.

### 3.4.2 Facial expression measures

Participants’ video recordings from the first presentation of stimuli and facial expressions were later analyzed by FaceReader 6.1 software, Noldus Technology. FaceReader 6.1 automatically analyzes faces in video recordings for the facial expressions of “happy”, “sad”, “angry”, “surprised”, “scared” and “disgusted”. These six facial expressions are considered to be representative of basic and universal emotions (Ekman, 1970). The software also rates facial expressions as “neutral”. One may not choose additional facial expressions when using the program. The software uses the Viola-Jones algorithm to detect the presence of the face, followed by a modeling of the face using an algorithmic approach based on the Active Appearance Method (see Cootes & Taylor, 2000). The actual classification of the facial expressions was done via artificial neural network training (for details refer to Bishop, 1995). Facial expressions included for analysis in this study were “neutral”, “happy”, and “sad”. The additional expressions automat-

---

1 The author expresses her gratitude to the Orangeburg NY Library for graciously supplying their conference room for experimental sessions.
ically recorded by the software of “angry”, “surprised”, “scared”, and “disgusted” were not included in statistical analyses because these categories were rarely obtained.

3.4.3 Subjective responses

Subjective emotional reactions were measured via 11 point rating scales and included self-assessments of happiness, sadness, anxiety, irritation, and desire to pick up the baby for interaction (see Wiesenfeld et al., 1985). The rating scales were printed on paper, and included one page per recording with a total of four pages (please see Self Assessment Form in Appendices). The packet was placed in front of each participant on the table, along with a pencil, and verbal instructions were provided to rate feelings for each recording presentation and to use the entire scale with a rating of zero for the least and ten the most of each particular emotion. The participants were also asked to rate how much they liked each vocal recording on an 11 point Likert scale labeled ‘least liked’ (0) to ‘most liked’ (10), marked and labeled with all 11 numbers (Charlton, Filippi & Fitch, 2012). The packets were later coded for stimulus order. All participant information was kept confidential, and questionnaires along with all data were coded and locked in a file cabinet in the principal investigator’s office.

IV. RESULTS

4.1 Dependent and independent variables

Dependent variables included heart rate variability (HRV), facial expressions (FE), and subjective emotional ratings (RT)/self-reports. Independent variables, or fixed factors, included breastfeeding mothers, non-mothers at high estradiol level in cycle, non-mothers at low estradiol level in cycle, fathers, and non-fathers.

4.2 Statistical analysis

Statistical analyses were completed using SPSS Statistics version 23.0 (2015) and Stata version 12 software (2011). Levene’s Test of Equality of Error Variances, which tests the null hypothesis that the
error variance of the dependent variable is equal across groups, was run prior to each analysis to ensure the homogeneity assumption was met. The homogeneity assumption was accepted for all analyses (p>.05) with the exception of the few that are mentioned below.

4.3 Comparisons between stimuli

**HRV Measures**

ANCOVAs were run to assess changes in HRV measures between stimuli adjusting for baseline levels. All groups were pooled for this analysis. No significant differences were found in SDNN or RMSSD between stimuli, however SDNN was slightly higher for babbling than cooing, crying and fussing (see Table 1). The covariate baseline SDNN was significant (p<.001) and correlated with SDNN (B=.74), and the covariate baseline RMSSD was significant (p<.001) and correlated with RMSSD (B=.819).

A linear mixed model was used to assess differences between stimuli adjusting for baseline values of the HRV measures while accounting for the repeated measures (e.g., the fact that the same individual heard all four stimuli). Although no differences in SDNN were found between stimuli (F=.83, p=.481), baseline SDNN was significant (p<.001) and correlated with SDNN (B=.74). There were also no significant differences in RMSSD between stimuli (F=.42, p=.732), and again the covariate BL_RMSSD was significant (p<.001) and correlated with RMSSD (B=.82).

Table 1

Differences in Heart Rate Variability (HRV) between stimuli

<table>
<thead>
<tr>
<th>Stimuli, Mean (SD)</th>
<th>HRV measure</th>
<th>Babbling</th>
<th>Cooing</th>
<th>Crying</th>
<th>Fussing</th>
<th>F*</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDNN</td>
<td></td>
<td>52.6 (22.8)</td>
<td>51.6 (21.3)</td>
<td>48.7 (22.8)</td>
<td>50.9 (22.1)</td>
<td>.827</td>
<td>.480</td>
</tr>
<tr>
<td>RMSSD</td>
<td></td>
<td>34.4 (17.9)</td>
<td>35.1 (17.8)</td>
<td>35.5 (18.1)</td>
<td>36.0 (18.4)</td>
<td>.421</td>
<td>.738</td>
</tr>
</tbody>
</table>

*F value corresponding to a linear mixed model with a random intercept for subject and baseline values of the HRV measure as covariate

SDNN: standard deviation of NN (or “normal-to-normal” inter-beat intervals)
RMSSD: square root of the mean squared difference of successive NN intervals

**Facial Expressions of Happiness**
The Kruskal-Wallis test revealed significant differences between stimuli for facial expressions of happiness (Chi-Square=28.1, p<.001) (see Table 2).

Table 2
Non parametric Testing Comparing Facial Expressions of Happiness between Stimuli

<table>
<thead>
<tr>
<th>Stimuli</th>
<th>Median</th>
<th>Range</th>
<th>Mean Rank</th>
<th>Chi-square</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Babbling</td>
<td>0</td>
<td>47.4</td>
<td>145.74</td>
<td>28.1</td>
<td>.000</td>
</tr>
<tr>
<td>Cooing</td>
<td>0</td>
<td>26.9</td>
<td>123.43</td>
<td>184.1</td>
<td>.000</td>
</tr>
<tr>
<td>Crying</td>
<td>0</td>
<td>10.2</td>
<td>105.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fussing</td>
<td>0</td>
<td>39.4</td>
<td>107.72</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Subjective Ratings of Happiness

To compare ratings of happiness between stimuli the Kruskal-Wallis test was run because of the abnormal distribution for this subjective rating scale (see Table 3). There were significant differences in self-rated happiness between stimuli (Chi-Square=184.11, p<.001).

Table 3
Non parametric Testing Comparing Ratings of Happiness between Stimuli

<table>
<thead>
<tr>
<th>Stimuli</th>
<th>Median</th>
<th>Range</th>
<th>Mean Rank</th>
<th>Chi-square</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Babbling</td>
<td>9.5</td>
<td>4</td>
<td>196.4</td>
<td>184.1</td>
<td>.000</td>
</tr>
<tr>
<td>Cooing</td>
<td>8.0</td>
<td>9</td>
<td>161.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crying</td>
<td>0</td>
<td>7</td>
<td>65.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fussing</td>
<td>0</td>
<td>7</td>
<td>58.8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A linear mixed model was used to account for the repeated measures (the variances were corrected to account for the repeated measures). Stimuli was a significant factor (F=457.92, p<.001) indicating that there were significant differences in self-rated happiness between stimuli. Pairwise
comparisons again showed the average happiness for babbling was significantly higher compared to the other three stimuli ($p<.001$). Cooing’s average happiness was significantly higher compared to crying ($p<.001$) and fussing ($p<.001$), and finally there were no significant differences in reported happiness between crying and fussing ($p=.442$).

![Figure 1. Self-rated happiness scales for all stimuli. The error bars show the relative standard deviations.](image)

**Ratings for Desire to Pick up the Baby**

Again a linear mixed model was used to account for the repeated measures. There were significant differences in ‘desire to pick up baby’ between stimuli ($F=15.77$, $p<.001$). The estimated means and pairwise comparisons showed babbling and cooing with similar means 5.93 and 5.37 ($p=.320$), which were significantly different (all $p$-values $<.001$) from crying and fussing which were perceived as equal ($p=.913$) in terms of desire to pick up baby, with higher means 8.08 and 8.13, respectively.

**Ratings of “Most Liked”**

The least liked/most liked scale did not follow a normal distribution, therefore the Kruskal-Wallis
test was run. As expected, babbling and cooing showed a higher concentration of values in the high scores (most liked) and crying and fussing showed a higher concentration in the low values (least liked). The Kruskal-Wallis test showed that there were significant differences between stimuli in the least liked/most liked scale (Chi-square=199.67, p<.001).

Table 4

<table>
<thead>
<tr>
<th>Stimuli</th>
<th>Median</th>
<th>Range</th>
<th>Mean Rank</th>
<th>Chi-square</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Babbling</td>
<td>10.0</td>
<td>4</td>
<td>202.2</td>
<td>199.67</td>
<td>.000</td>
</tr>
<tr>
<td>Cooing</td>
<td>7.0</td>
<td>9</td>
<td>158.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crying</td>
<td>0</td>
<td>4</td>
<td>64.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fussing</td>
<td>0</td>
<td>4</td>
<td>57.8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A linear mixed model was fitted to adjust for the repeated measures. There were significant differences in ratings of least liked/most liked between stimuli (F=981.37, p<.001). The pairwise comparisons showed that the most preferred was babbling, which was significantly higher than the score for cooing (p<.001). Both babbling and cooing had scores significantly higher (towards “most liked” end of scale) than crying and fussing (all p-values for these comparisons were <.001). Finally, there were no significant differences between crying and fussing in terms of “least liked” (p=.311).
4.4 Comparisons between mothers and non-mothers

**Babbling**

The analysis in this section was stratified by stimuli in order to compare mothers versus non-mothers for each of the stimuli separately. ANCOVAs were used with mothers and non-mothers as factors and the baseline value of each HRV variable as the covariate. There were no differences between the mothers and non-mothers in SDNN levels for babbling (see Table 5). The covariate BL_SDNN was significant (p<.001) and positively associated with SDNN (B=.72). There were no significant differences between the mother and non-mother groups in RMSSD during babbling (see Table 5). The covariate BL_RMSDD was significant (p<.001) and positively correlated with RMSSD (B=.95). The mothers had higher baseline levels (SDNN, RMSSD) at rest as well, as compared to both non-mother groups (see additional analyses section below).

There were also no significant differences found between mothers and non-mothers in facial expressions of happiness, self-reported happiness, or for the least liked/most liked scales when listening. 

Figure 2. Mean ratings on the Least liked/Most liked scale for all stimuli. The error bars show the relative standard deviations.
to babbling (see Table 5).

Table 5

<table>
<thead>
<tr>
<th>Measure</th>
<th>BF</th>
<th>NMH</th>
<th>NML</th>
<th>Chi-square/F-test</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDNN, Mean (SD)²</td>
<td>68.5</td>
<td>44.7</td>
<td>52.6</td>
<td>.480</td>
<td>.623</td>
</tr>
<tr>
<td>RMSSD, Mean (SD)²</td>
<td>44.1</td>
<td>32.1</td>
<td>39.4</td>
<td>.239</td>
<td>.789</td>
</tr>
<tr>
<td>Happiness FE</td>
<td>12.1</td>
<td>0</td>
<td>0</td>
<td>5.49</td>
<td>.064</td>
</tr>
<tr>
<td>Neutral FE</td>
<td>50.4</td>
<td>70.3</td>
<td>61.3</td>
<td>2.19</td>
<td>.335</td>
</tr>
<tr>
<td>Happiness RT</td>
<td>10 (3)</td>
<td>9 (4)</td>
<td>9.5 (3)</td>
<td>2.52</td>
<td>.284</td>
</tr>
<tr>
<td>Least liked/most liked</td>
<td>10 (2)</td>
<td>9 (3)</td>
<td>9.5 (3)</td>
<td>3.36</td>
<td>.186</td>
</tr>
</tbody>
</table>

¹p-values correspond to the non-parametric Kruskal Wallis test  
²p-values correspond to an ANCOVA adjusting for baseline value of the corresponding HRV measure  
SDNN: standard deviation of NN (or “normal-to-normal” inter-beat intervals)  
RMSSD: square root of the mean squared difference of successive NN intervals  
FE: facial expression  
RT: rating

Cooing

ANCOVAs were used with mothers and non-mothers as factors and the baseline value of each HRV variable as the covariate. There were no significant differences between the mothers and non-mothers in SDNN or RMSSD while listening to the cooing stimuli (see Table 6). Both covariates (baselines) were significant (p<.001) and positively correlated with SDNN (B=.65) and RMSSD (B=.90).

No significant differences were found in the comparisons of facial expressions of happiness, self-rated happiness, or for the least liked/most liked ratings between mothers and non-mothers while listening to the cooing stimuli (see Table 6). An additional analysis comparing neutral facial expressions was run and no significant differences were found there either.
Table 6

Differences between Breastfeeding Mothers (BF) and Non-mothers (NMH and NML) for Cooing

<table>
<thead>
<tr>
<th>Measure</th>
<th>BF</th>
<th>NMH</th>
<th>NML</th>
<th>Chi-square/F-test</th>
<th>p-value¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDNN, Mean(SD)²</td>
<td>66.9 (13.2)</td>
<td>44.6 (21.3)</td>
<td>51.9 (21.6)</td>
<td>.527</td>
<td>.595</td>
</tr>
<tr>
<td>RMSSD, Mean (SD)²</td>
<td>44.7 (15.2)</td>
<td>33.6 (19.8)</td>
<td>40.0 (17.7)</td>
<td>.178</td>
<td>.838</td>
</tr>
<tr>
<td>Happiness FE</td>
<td>0 (15.2)</td>
<td>0 (26.1)</td>
<td>0 (17.9)</td>
<td>2.43</td>
<td>.296</td>
</tr>
<tr>
<td>Neutral FE</td>
<td>66.2 (51.9)</td>
<td>64.7 (64.8)</td>
<td>59.4 (80.4)</td>
<td>.713</td>
<td>.700</td>
</tr>
<tr>
<td>Happiness RT</td>
<td>8 (7)</td>
<td>7 (5)</td>
<td>8 (9)</td>
<td>.335</td>
<td>.846</td>
</tr>
<tr>
<td>Least liked/most liked</td>
<td>6 (7)</td>
<td>6.5 (5)</td>
<td>7 (9)</td>
<td>.140</td>
<td>.932</td>
</tr>
</tbody>
</table>

¹p-values correspond to the non-parametric Kruskal Wallis test
²p-values correspond to an ANCOVA adjusting for baseline value of the corresponding HRV measure
SDNN: standard deviation of NN (or “normal-to-normal” inter-beat intervals)
RMSSD: square root of the mean squared difference of successive NN intervals
FE: facial expression
RT: rating

Crying

HRV variables comparing mothers and non-mothers were looked at using ANCOVAs with mothers and non-mothers as factors and the baseline value of the particular HRV variable as the covariate. There were no significant differences found in SDNN or RMSSD between mothers and non-mothers (see Table 7). The covariate BL_SDNN was significant (p<.001) and positively associated with SDNN (B=.78), and the covariate BL_RMSSD was significant (p<.001) and positively associated with RMSSD (B=.91).

Levene’s Test of Equality of Error Variances showed the homogeneity assumption was not met (p=.016) for the comparison of facial expressions of sadness during the crying stimuli. No significant differences were found between mothers and non-mothers (see Table 7). Comparisons for happy facial expressions during the crying stimuli were not run due to abnormal distributions of the data (mostly zero values). An additional analysis was run for neutral facial expressions and no differences were found between mothers and non-mothers.

No differences were found between mothers and non-mothers in self-reported happiness, anxiety, irritation, or self-rated desire to pick up the baby when listening to the crying stimuli (see Table 7).
Table 7

Differences between Breastfeeding Mothers (BF) and Non-mothers (NMH and NML) for Crying

<table>
<thead>
<tr>
<th>Measure</th>
<th>BF</th>
<th>NMH</th>
<th>NML</th>
<th>Chi-square/F-test</th>
<th>p-value1</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDNN, Mean (SD)2</td>
<td>56.4 (15.3)</td>
<td>43.5 (25.9)</td>
<td>49.4 (20.3)</td>
<td>.634</td>
<td>.537</td>
</tr>
<tr>
<td>RMSSD, Mean (SD)2</td>
<td>44.0 (13.2)</td>
<td>34.9 (21.1)</td>
<td>40.4 (17.1)</td>
<td>.915</td>
<td>.411</td>
</tr>
<tr>
<td>Sadness FE</td>
<td>22.3 (71.1)</td>
<td>24.4 (46)</td>
<td>27.7 (67.8)</td>
<td>.403</td>
<td>.817</td>
</tr>
<tr>
<td>Neutral FE</td>
<td>66.2 (51.9)</td>
<td>65.5 (59.9)</td>
<td>60.7 (67.5)</td>
<td>1.37</td>
<td>.504</td>
</tr>
<tr>
<td>Happiness RT</td>
<td>0 (3)</td>
<td>1 (5)</td>
<td>0 (5)</td>
<td>2.02</td>
<td>.364</td>
</tr>
<tr>
<td>Sadness RT</td>
<td>8 (7)</td>
<td>6 (10)</td>
<td>5.5 (7)</td>
<td>4.62</td>
<td>.099</td>
</tr>
<tr>
<td>Anxiety RT</td>
<td>7.5 (7)</td>
<td>6 (8)</td>
<td>6.5 (10)</td>
<td>2.87</td>
<td>.238</td>
</tr>
<tr>
<td>Irritation RT</td>
<td>1 (10)</td>
<td>5.5 (9)</td>
<td>5 (10)</td>
<td>3.16</td>
<td>.206</td>
</tr>
<tr>
<td>Desire to pick-up baby</td>
<td>10 (10)</td>
<td>9.5 (9)</td>
<td>9 (9)</td>
<td>.806</td>
<td>.668</td>
</tr>
</tbody>
</table>

1p-values correspond to the non-parametric Kruskal Wallis test
2p-values correspond to an ANCOVA adjusting for baseline value of the corresponding HRV measure
SDNN: standard deviation of NN (or “normal-to-normal” inter-beat intervals)
RMSSD: square root of the mean squared difference of successive NN intervals
FE: facial expression
RT: rating

Fussing

HRV variables for mothers and non-mothers were compared using ANCOVAs with mothers and non-mothers as factors and the baseline value of each HRV variable as the covariate. There were no significant differences in SDNN or RMSSD between mothers and non-mothers (see Table 8). The covariate BL_SDNN was significant (p<.001) and positively associated with SDNN (B=.73) and the covariate BL_R MSSD was significant (p<.001) and positively associated with RMSSD (B=.91).

The differences did not reach significance between mothers and non-mothers in facial expressions of sadness when listening to fussing (see Table 8). Comparisons for happy facial expressions during the fussing stimuli were not run due to abnormal distributions of the data (mostly zero values). An additional analysis was run to examine neutral facial expressions, and similar to what was found for crying, no differences were seen in FE neutral between mothers and non-mothers when listening to fussing. There were no differences between mothers and non-mothers in self-reported happiness, self-reported sadness, or self-reported irritation when listening to fussing. The homogeneity of variances assumption was not met (p<.001) for the comparison of self-reported desire to pick up baby
while listening to the fussing stimuli. There were significant differences between mothers and non-mothers (see Table 8).

Table 8

Differences between Breastfeeding Mothers (BF) and Non-mothers (NMH and NML) for Fussing

<table>
<thead>
<tr>
<th>Measure</th>
<th>BF</th>
<th>NMH</th>
<th>NML</th>
<th>Chi-square/F-test</th>
<th>p-value¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDNN, Mean(SD)²</td>
<td>61.1 (17.2)</td>
<td>47.3 (20.1)</td>
<td>52.1 (27.8)</td>
<td>.186</td>
<td>.832</td>
</tr>
<tr>
<td>RMSSD, Mean (SD)²</td>
<td>44.9 (15.2)</td>
<td>33.5 (18.9)</td>
<td>39.2 (18.6)</td>
<td>.165</td>
<td>.849</td>
</tr>
<tr>
<td>Sadness FE</td>
<td>23.2 (80)</td>
<td>19 (56.5)</td>
<td>19.5 (68.3)</td>
<td>.016</td>
<td>.992</td>
</tr>
<tr>
<td>Neutral FE</td>
<td>57.8 (63.1)</td>
<td>69.1 (60.5)</td>
<td>63.3 (61.4)</td>
<td>2.03</td>
<td>.363</td>
</tr>
<tr>
<td>Happiness RT</td>
<td>0 (5)</td>
<td>0 (2)</td>
<td>0 (3)</td>
<td>.458</td>
<td>.795</td>
</tr>
<tr>
<td>Sadness RT</td>
<td>8 (6)</td>
<td>6.5 (10)</td>
<td>7 (10)</td>
<td>2.99</td>
<td>.224</td>
</tr>
<tr>
<td>Anxiety RT</td>
<td>8 (6)</td>
<td>7 (10)</td>
<td>6 (10)</td>
<td>1.92</td>
<td>.382</td>
</tr>
<tr>
<td>Irritation RT</td>
<td>2.5 (9)</td>
<td>5.5 (9)</td>
<td>6.5 (8)</td>
<td>3.51</td>
<td>.173</td>
</tr>
<tr>
<td>Desire to pick-up baby</td>
<td>10 (1)</td>
<td>9 (8)</td>
<td>9 (8)</td>
<td>7.01</td>
<td>.030</td>
</tr>
</tbody>
</table>

¹p-values correspond to the non-parametric Kruskal Wallis test
²p-values correspond to an ANCOVA adjusting for baseline value of the corresponding HRV measure
SDNN: standard deviation of NN (or “normal-to-normal” inter-beat intervals)
RMSSD: square root of the mean squared difference of successive NN intervals
FE: facial expression
RT: rating

4.5 Comparisons between fathers and non-fathers

Babbling

The analyses in this section was stratified by stimuli in order to compare fathers versus non-fathers for each of the stimuli separately. ANCOVAs for HRV variables were run with group (father, non-father) as factors and the baseline value of each particular HRV variable as covariate. There were no significant differences between the fathers and non-fathers in terms of SDNN during babbling (see Table 9). The covariate BL_SDNN was significant (p<.001) and positively associated with SDNN (B=.77). There were significant differences between fathers and non-fathers in terms of RMSSD during babbling; fathers’ RMSSD was higher than non-fathers (31.22 vs. 25.19). The covariate BL_ RMSSD was significant (p<.001) and positively correlated with RMSSD (B=.66).
No significant differences were found in facial expressions of happiness for the babbling stimuli between fathers and non-fathers (see Table 9). An additional analysis was run to compare neutral facial expressions, however the homogeneity assumption was not met (p=.022), and there were no significant differences between fathers and non-fathers in FE neutral when listening to babbling.

There were significant differences between fathers and non-fathers in self-reported happiness when listening to babbling, with fathers showing higher self-reported happiness (see Table 9).

The results obtained for the least liked/most liked scale were similar to the ones obtained for self-rated happiness. Fathers showed significantly higher scores compared to non-fathers (see Table 9). The homogeneity test was not accepted in this case though (p=.002), which could affect the robustness of the estimates.

Table 9

<table>
<thead>
<tr>
<th>Measure</th>
<th>Fathers</th>
<th>Non-fathers</th>
<th>Mann-Whitney/F-test</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDNN, Mean(SD)²</td>
<td>53.0 (16.1)</td>
<td>44.0 (29.4)</td>
<td>2.04</td>
<td>.168.</td>
</tr>
<tr>
<td>RMSSD, Mean (SD)²</td>
<td>30.9 (11.8)</td>
<td>25.5 (19.1)</td>
<td>5.12</td>
<td>.034</td>
</tr>
<tr>
<td>Happiness FE</td>
<td>0 (37)</td>
<td>0 (47.4)</td>
<td>59.5</td>
<td>.390</td>
</tr>
<tr>
<td>Neutral FE</td>
<td>72.9 (51.5)</td>
<td>60.9 (82)</td>
<td>51.5</td>
<td>.236</td>
</tr>
<tr>
<td>Happiness RT</td>
<td>10 (3)</td>
<td>8 (4)</td>
<td>38.5</td>
<td>.034</td>
</tr>
<tr>
<td>Least liked/most liked</td>
<td>10 (2)</td>
<td>9 (4)</td>
<td>36.5</td>
<td>.025</td>
</tr>
</tbody>
</table>

1p-values correspond to the non-parametric Mann Whitney U test
2p-values correspond to an ANCOVA adjusting for baseline value of the corresponding HRV measure
SDNN: standard deviation of NN (or "normal-to-normal" inter-beat intervals)
RMSSD: square root of the mean squared difference of successive NN intervals
FE: facial expression
RT: rating

**Cooing**

ANCOVAs were run for the HRV measures between fathers and non-fathers. Differences in SDNN or RMSSD did not reach significance for the cooing stimuli between fathers and non-fathers (see Table 10). The covariate BL_SDNN was significant (p<.001) and positively associated with SDNN (B=.68), and the covariate BL_RMSDD was significant (p<.001) and positively associated with RMSSD.
No significant differences were found in the comparison of facial expressions of happiness for cooing (see Table 10). An additional analysis was run and no significant differences were found between fathers and non-fathers in terms of neutral facial expressions.

Differences between fathers and non-fathers in self-rated happiness or least liked/most liked ratings did not reach significance when listening to cooing stimuli (see Table 10).

### Table 10

**Differences between Fathers and Non-fathers for Cooing**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Group, Median (Range)</th>
<th>Fathers</th>
<th>Non-fathers</th>
<th>Mann-Whitney /F-test</th>
<th>p-value¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDNN, Mean(SD)²</td>
<td>49.5 (11.4)</td>
<td>45.1 (29.3)</td>
<td>.757</td>
<td>.394</td>
<td></td>
</tr>
<tr>
<td>RMSSD, Mean (SD)²</td>
<td>31.4 (8.8)</td>
<td>25.5 (19.1)</td>
<td>4.13</td>
<td>.055</td>
<td></td>
</tr>
<tr>
<td>Happiness FE</td>
<td>0 (16.8)</td>
<td>0 (26.9)</td>
<td>61.5</td>
<td>.351</td>
<td></td>
</tr>
<tr>
<td>Neutral FE</td>
<td>74.3 (73.6)</td>
<td>66 (70.5)</td>
<td>58.0</td>
<td>.419</td>
<td></td>
</tr>
<tr>
<td>Happiness RT</td>
<td>8 (4)</td>
<td>7 (5)</td>
<td>54.0</td>
<td>.281</td>
<td></td>
</tr>
<tr>
<td>Least liked/most liked</td>
<td>7 (4)</td>
<td>7 (6)</td>
<td>58.5</td>
<td>.424</td>
<td></td>
</tr>
</tbody>
</table>

¹p-values correspond to the non-parametric Mann Whitney U test
²p-values correspond to an ANCOVA adjusting for baseline value of the corresponding HRV measure
SDNN: standard deviation of NN (or “normal-to-normal” inter-beat intervals)
RMSSD: square root of the mean squared difference of successive NN intervals
FE: facial expression
RT: rating

### Crying

HRV variables for fathers and non-fathers were compared using ANCOVAs with group as the factor and the baseline value of the particular HRV variable as covariate. There were no significant differences in SDNN or RMSSD between fathers and non-fathers for the crying stimuli (see Table 11). The covariate BL_SDNN was significant (p<.001) and positively associated with SDNN (B=.82), and the covariate BL_RMSSD was significant (p<.001) and positively associated with RMSSD (B=.74).

No differences were found between fathers and non-fathers in facial expressions of sadness (see Table 11). ANOVAs were not run for FE happy because the distribution showed a very high concentration of zeroes. An additional analysis comparing neutral facial expressions was run and no statistically significant differences were found between fathers and non-fathers (see Table 11).
The homogeneity assumption was not met (p=.016) for the comparison of self-reported happiness between fathers and non-fathers when listening to the crying stimuli. No differences were found between fathers and non-fathers (see Table 11).

No differences were found in self-reported sadness, anxiety, irritation or desire to pick up the baby between fathers and non-fathers when listening to crying (see Table 11).

Table 11
*Differences between Fathers and Non-fathers for Crying*

<table>
<thead>
<tr>
<th>Measure</th>
<th>Group, Median (Range)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fathers</td>
<td>Non-fathers</td>
<td>Mann-Whitney /F-test</td>
<td>p-value¹</td>
</tr>
<tr>
<td>SDNN, Mean(SD)²</td>
<td>42.8 (11.8)</td>
<td>51.7 (34.4)</td>
<td>3.01</td>
<td>.097</td>
</tr>
<tr>
<td>RMSSD, Mean (SD)²</td>
<td>31.1 (14.8)</td>
<td>27.2 (20.5)</td>
<td>1.91</td>
<td>.181</td>
</tr>
<tr>
<td>Sadness FE</td>
<td>0 (75.0)</td>
<td>5.7 (65.5)</td>
<td>70.0</td>
<td>.900</td>
</tr>
<tr>
<td>Neutral FE</td>
<td>73.7 (76.9)</td>
<td>58.1 (68.6)</td>
<td>57.0</td>
<td>.386</td>
</tr>
<tr>
<td>Happiness RT</td>
<td>.5 (7)</td>
<td>.5 (4)</td>
<td>62.5</td>
<td>.557</td>
</tr>
<tr>
<td>Sadness RT</td>
<td>5 (10)</td>
<td>4.5 (8)</td>
<td>58.0</td>
<td>.409</td>
</tr>
<tr>
<td>Anxiety RT</td>
<td>5 (10)</td>
<td>5 (8)</td>
<td>69.5</td>
<td>.883</td>
</tr>
<tr>
<td>Irritation RT</td>
<td>1.5 (7)</td>
<td>3 (8)</td>
<td>47.0</td>
<td>.140</td>
</tr>
<tr>
<td>Desire to pick-up baby</td>
<td>9 (7)</td>
<td>6.5 (8)</td>
<td>40.5</td>
<td>.063</td>
</tr>
</tbody>
</table>

¹p-values correspond to the non-parametric Mann Whitney U test  
²p-values correspond to an ANCOVA adjusting for baseline value of the corresponding HRV measure  
SDNN: standard deviation of NN (or “normal-to-normal” inter-beat intervals)  
RMSSD: square root of the mean squared difference of successive NN intervals  
FE: facial expression  
RT: rating

**Fussing**

ANCOVAs were run comparing HRV variables between fathers and non-fathers using group as factor and the baseline value of the particular heart rate variable as the covariate. There were no significant differences in SDNN or RMSSD between fathers and non-fathers for the fussing stimuli (see Table 12). The covariate BL_SDNN was significant (p<.001) and positively associated with SDNN (B=.76), and the covariate BL_RMSSD was significant (p<.001) and positively associated with RMSSD (B=.71).

There were no significant differences between fathers and non-fathers in facial expressions of sadness when listening to fussing (see Table 12). The facial expressions of happiness distribution contained a very high concentration of zeroes, therefore statistical analysis was not done. An additional
analysis of neutrality of facial expressions was run and no differences were found between fathers and non-fathers when listening to fussing.

Differences between fathers and non-fathers did not reach significance in self-reported happiness, sadness, irritation or desire to pick up the baby when listening to fussing (see Table 12).

Table 12

<table>
<thead>
<tr>
<th>Measure</th>
<th>Fathers</th>
<th>Non-fathers</th>
<th>Mann-Whitney /F-test</th>
<th>p-value¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDNN, Mean(SD)²</td>
<td>45.9 (9.1)</td>
<td>48.2 (30.1)</td>
<td>.705</td>
<td>.411</td>
</tr>
<tr>
<td>RMSSD, Mean (SD)²</td>
<td>33.6 (16.4)</td>
<td>28.2 (20.6)</td>
<td>1.85</td>
<td>.189</td>
</tr>
<tr>
<td>Sadness FE</td>
<td>6.4 (83.8)</td>
<td>6.1 (88.3)</td>
<td>66.0</td>
<td>.716</td>
</tr>
<tr>
<td>Neutral FE</td>
<td>77.2 (89.2)</td>
<td>66.6 (77.8)</td>
<td>62.0</td>
<td>.564</td>
</tr>
<tr>
<td>Happiness RT</td>
<td>1 (7)</td>
<td>.5 (5)</td>
<td>58.5</td>
<td>.410</td>
</tr>
<tr>
<td>Sadness RT</td>
<td>5 (9)</td>
<td>4 (9)</td>
<td>62.0</td>
<td>.557</td>
</tr>
<tr>
<td>Anxiety RT</td>
<td>6 (10)</td>
<td>3.5 (9)</td>
<td>57.5</td>
<td>.398</td>
</tr>
<tr>
<td>Irritation RT</td>
<td>2 (9)</td>
<td>2.5 (8)</td>
<td>66.0</td>
<td>.724</td>
</tr>
<tr>
<td>Desire to pick-up baby</td>
<td>9 (5)</td>
<td>7 (6)</td>
<td>45.0</td>
<td>.111</td>
</tr>
</tbody>
</table>

¹p-values correspond to the non-parametric Mann Whitney U test
²p-values correspond to an ANCOVA adjusting for baseline value of the corresponding HRV measure
SDNN: standard deviation of NN (or "normal-to-normal" inter-beat intervals)
RMSSD: square root of the mean squared difference of successive NN intervals
FE: facial expression
RT: rating

4.6 Comparisons between all groups (additional analyses)

ANCOVAs were also run for each stimulus to assess for differences in HRV measures between all groups (the other analyses examined differences between females and males), adjusting for baseline values. There were no significant differences in SDNN between groups during presentation of the babbling, cooing, crying or fussing stimuli (see Table 13). All covariate baselines for SDNN were significant and with positive coefficients (B=.75, p<.001 for babbling, B=.67, p<.001 for cooing, B=.80, p<.001 for crying, B=.75, p<.001 for fussing), indicating that higher levels of baseline SDNN were associated with higher levels of SDNN during the stimuli presentation, as expected.

ANCOVA revealed that there were no significant differences in RMSSD between groups during presentation of the babbling stimuli (F=2.03, p=.103). The covariate baseline RMSSD was significant.
(p<.001) with a positive coefficient (B=.80), indicating that higher levels of baseline RMSSD were associated with higher levels of RMSSD while listening to babbling. Looking at the parameter estimates, despite having an overall non-significant test for group, there was a significant p-value (p=.014) for non-fathers for these stimuli. The estimated marginal means for group revealed that non-fathers had the lowest RMSSD average (mean=28.91) after adjusting for BL_RMSSD. The pairwise comparisons show that non-fathers had significantly lower RMSSD average while listening to babbling compared to mothers (p=.034), fathers (p=.051), non-mothers at high estradiol point in cycle (p=.042), and non-mothers at low estradiol point in cycle (p=.014). Power was calculated for the pairwise comparisons and ranged from 47.9% to 53.6%, thus indicating there was not enough data in this instance in order to determine if there was an effect or not.

The differences in RMSSD for cooing between groups did not reach significance (F=2.09, p=.095). The covariate, BL_RMSSD reached significance, as expected, and had a positive coefficient (B=.78, p<.001). The coefficient corresponding to non-fathers is significant (p=.014). The estimated marginal means show the average RMSSD for non-fathers during the cooing presentation (mean=29.33) was the lowest. The pairwise comparisons show that non-fathers had significantly lower RMSSD average while listening to cooing compared to mothers (p=.033), non-mothers at high estradiol point in cycle (p=.026), and non-mothers at low estradiol point in cycle (p=.014). The difference did not reach significance for the comparison of non-fathers with fathers (p=.061).

The F-test indicates that there were no significant differences between groups in RMSSD while listening to crying (F=1.85, p=.133); baseline RMSSD was correlated with RMSSD (p<.001, B=.82). The individual pairwise comparisons between the estimated marginal means show that non-fathers had significantly lower RMSSD (mean=30.74) compared to non-mothers with high estradiol (mean=38.22, p=.019) and non-mothers with low estradiol (mean=37.79, p=.027). Post-hoc power analysis for the pairwise comparisons ranged from 51.2% to 52.4%, indicating more data is needed to establish if there is a difference. Group was clearly not a significant factor on HRV responses to the fussing stimulus; there were no differences in RMSSD between groups. The covariate BL_RMSSD was significant and correlated with RMSSD (B=.81, p<.001).
Table 13. Comparisons of Heart Rate Variability (HRV) values adjusted by baseline values

<table>
<thead>
<tr>
<th>Participant Group</th>
<th>Mean</th>
<th>Std. Error</th>
<th>F</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Babbling</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SDNN</td>
<td>BF mothers</td>
<td>68.5</td>
<td>13.7</td>
<td>1.37</td>
</tr>
<tr>
<td></td>
<td>Fathers</td>
<td>53.0</td>
<td>16.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-fathers</td>
<td>44.0</td>
<td>29.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-mothers-high</td>
<td>44.7</td>
<td>23.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-mothers-low</td>
<td>52.6</td>
<td>22.4</td>
<td></td>
</tr>
<tr>
<td>RMSSD</td>
<td>BF mothers</td>
<td>44.1</td>
<td>13.3</td>
<td>2.03</td>
</tr>
<tr>
<td></td>
<td>Fathers</td>
<td>30.9</td>
<td>11.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-fathers</td>
<td>25.5</td>
<td>19.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-mothers-high</td>
<td>32.1</td>
<td>20.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-mothers-low</td>
<td>39.4</td>
<td>17.9</td>
<td></td>
</tr>
<tr>
<td>Cooing</td>
<td>BF mothers</td>
<td>66.9</td>
<td>13.2</td>
<td>1.08</td>
</tr>
<tr>
<td></td>
<td>Fathers</td>
<td>49.5</td>
<td>11.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-fathers</td>
<td>45.1</td>
<td>29.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-mothers-high</td>
<td>44.6</td>
<td>21.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-mothers-low</td>
<td>51.9</td>
<td>21.6</td>
<td></td>
</tr>
<tr>
<td>RMSSD</td>
<td>BF mothers</td>
<td>44.7</td>
<td>15.2</td>
<td>2.09</td>
</tr>
<tr>
<td></td>
<td>Fathers</td>
<td>31.4</td>
<td>8.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-fathers</td>
<td>26.0</td>
<td>21.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-mothers-high</td>
<td>33.6</td>
<td>19.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-mothers-low</td>
<td>40.0</td>
<td>17.7</td>
<td></td>
</tr>
<tr>
<td>Crying</td>
<td>BF mothers</td>
<td>56.4</td>
<td>15.3</td>
<td>1.27</td>
</tr>
<tr>
<td></td>
<td>Fathers</td>
<td>42.8</td>
<td>11.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-fathers</td>
<td>51.7</td>
<td>34.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-mothers-high</td>
<td>43.5</td>
<td>25.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-mothers-low</td>
<td>49.4</td>
<td>20.3</td>
<td></td>
</tr>
<tr>
<td>RMSSD</td>
<td>BF mothers</td>
<td>44.0</td>
<td>13.2</td>
<td>1.85</td>
</tr>
<tr>
<td></td>
<td>Fathers</td>
<td>31.1</td>
<td>14.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-fathers</td>
<td>27.2</td>
<td>20.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-mothers-high</td>
<td>34.9</td>
<td>21.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-mothers-low</td>
<td>40.4</td>
<td>17.1</td>
<td></td>
</tr>
<tr>
<td>Fussing</td>
<td>BF mothers</td>
<td>61.1</td>
<td>17.1</td>
<td>.784</td>
</tr>
<tr>
<td></td>
<td>Fathers</td>
<td>45.9</td>
<td>9.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-fathers</td>
<td>48.2</td>
<td>30.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-mothers-high</td>
<td>47.3</td>
<td>20.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-mothers-low</td>
<td>52.1</td>
<td>27.8</td>
<td></td>
</tr>
<tr>
<td>RMSSD</td>
<td>BF mothers</td>
<td>44.9</td>
<td>15.2</td>
<td>.891</td>
</tr>
<tr>
<td></td>
<td>Fathers</td>
<td>33.6</td>
<td>16.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-fathers</td>
<td>28.2</td>
<td>20.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-mothers-high</td>
<td>33.5</td>
<td>18.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-mothers-low</td>
<td>39.8</td>
<td>18.6</td>
<td></td>
</tr>
</tbody>
</table>

SDNN: standard deviation of NN (“normal-to-normal” inter-beat intervals); RMSSD: square root of mean squared difference of successive NN intervals
The breastfeeding mothers had higher HRV baseline levels (SDNN, RMSSD) at rest as compared to both non-mother groups and both male groups (see Table 14 below).

Table 14. Baseline Heart Rate Variability (HRV) values

<table>
<thead>
<tr>
<th>Participant Group</th>
<th>Mean</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>BL_SDNN</td>
<td>BF mothers</td>
<td>74.81</td>
</tr>
<tr>
<td></td>
<td>Fathers</td>
<td>60.96</td>
</tr>
<tr>
<td></td>
<td>Non-fathers</td>
<td>59.11</td>
</tr>
<tr>
<td></td>
<td>Non-mothers-high</td>
<td>49.61</td>
</tr>
<tr>
<td></td>
<td>Non-mothers-low</td>
<td>59.65</td>
</tr>
<tr>
<td>BL_RMSSD</td>
<td>BF mothers</td>
<td>48.85</td>
</tr>
<tr>
<td></td>
<td>Fathers</td>
<td>33.34</td>
</tr>
<tr>
<td></td>
<td>Non-fathers</td>
<td>34.29</td>
</tr>
<tr>
<td></td>
<td>Non-mothers-high</td>
<td>34.56</td>
</tr>
<tr>
<td></td>
<td>Non-mothers-low</td>
<td>41.68</td>
</tr>
</tbody>
</table>

SDNN: standard deviation of NN (or “normal-to-normal” inter-beat intervals)
RMSSD: square root of the mean squared difference of successive NN intervals

One way ANOVAs were run for each baseline variable. The ANOVA omnibus test indicated that there were no significant differences across the five listener groups for SDNN (F=1.86, p=.131) and RMSSD (F=1.41, p=.242). Pairwise comparisons were also generated and mothers showed significant differences in SDNN from non-mothers at high estradiol point in cycle (p=.009), however differences from non-fathers did not reach significance (p=.099). Differences in RMSSD for mothers as compared to fathers (p=.055), non-fathers (p=.071), and non-mothers at high estradiol point in cycle (p=.076) did not reach significance. More data is needed to determine if there was an effect or not.
4.7 Summary

No significant differences were found between stimuli in HRV. There were significant differences between stimuli for facial expressions of happiness, with babbling receiving the greatest amount of happy facial expressions, followed by cooing, fussing, and last crying. There were significant differences in self-rated happiness between stimuli, with the average happiness for babbling significantly higher compared to the other three stimuli; cooing’s average happiness was significantly higher compared to crying and fussing, and finally there were no significant differences in reported happiness between crying and fussing. There were significant differences in ‘desire to pick up baby’ between stimuli, with babbling and cooing scoring significantly lower than crying and fussing. There were significant differences in ratings of least liked/most liked between stimuli; the most preferred was babbling, which was significantly higher than cooing. Both babbling and cooing had scores significantly higher (towards “most liked” end of scale) than crying and fussing, and there were no significant differences between crying and fussing in terms of “least liked”.

There were significant differences between mothers and non-mothers while listening to the fussing stimuli, with mothers reporting a greater desire to pick up the baby. Fathers had significantly higher RMSSD than non-fathers while listening to babbling. There were significant differences between fathers and non-fathers in self-reported happiness and on the least liked/most liked scale when listening to babbling, with fathers showing higher scores. Additional analyses were also run for each stimulus to assess for differences in HRV measures between all groups; the other analyses examined differences between females and males. ANCOVAs revealed no significant differences, yet pairwise comparisons did show a few significant differences. Non-fathers had significantly lower RMSSD while listening to babbling compared to all other subject groups. Non-fathers had significantly lower RMSSD while listening to cooing compared to mothers and both non-mother groups. Non-fathers had significantly lower RMSSD when listening to crying compared to both non-mother groups. Pairwise comparisons were also generated to compare baseline HRV levels, and mothers showed significantly higher SDNN than non-mothers at high estradiol point in cycle. Post-hoc power analysis for the pairwise comparisons ranged from 47.9% to 53.6%, indicating more data is needed to establish if there truly is a difference for the HRV
measures.

V. DISCUSSION

5.1 Accepted hypotheses

5.1.1 Evolutionary relevance of happy facial expressions for babbling

It was predicted that participants would show higher incidence of facial expressions of happiness for the canonical vocalizations as compared to the non-canonical, fussing and crying stimuli. Babbling lead to significantly more facial expressions of happiness than cooing, crying and fussing. The responses to the stimuli were significantly different from each other except for crying and fussing, which did not differ significantly different in evoking happy facial expressions. The incidence of happy facial expressions to babbling was 72% lower for cooing, 95% lower for crying and 93% lower for fussing. Joy as communicated via smiling is considered one of six basic facial expressions that is cross-culturally consistent (Darwin, 1872; Izard, 1994). Along with the original studies on the universality of facial expressions (Ekman, 1972, 1973; Ekman & Friesen, 1971; Ekman, Sorenson, & Friesen, 1969; Izard, 1971) more than 30 others have replicated the findings of this universal recognition of emotion in the face (reviewed in Matsumoto, 2001).

Another aspect of facial expression is its’ use from a signaling perspective. During parent/infant interactions babies often start with a non-Duchenne smile and then progress to an energetically more expensive Duchenne smile by contracting the orbicularis oculi around the sides of the eyes (Messinger, D., Fogel, A., & Dickson, K.L., 1999). It is possible that this change in facial expression could be due to the change in benefits of the more costly smile (Schmidt, K. L., & Cohn, J. F., 2001). Positive fitness consequences would include increased attention and care. It seems likely that the same concept would be applied to infant babbling.

5.1.2 Conscious preference for canonical infant vocal signals: subjective ratings
It was also expected that participants would assign higher subjective ratings of happiness to the canonical vocalizations as compared to the non-canonical, fussing and crying vocal stimuli. There were significant differences in self-rated happiness evoked by the different stimuli (see Figure 1). Babbling and cooing were the stimuli higher in self-rated happiness, and crying and fussing elicited lower scores. Average happiness for babbling was significantly higher than the other three stimuli. Average happiness for cooing was significantly higher compared to crying and fussing, and finally there were no significant differences in reported happiness evoked by crying and fussing.

Finally, it was also hypothesized that study participants would give ratings towards the “most liked” end of the scale for the canonical vocalizations as compared to the non-canonical, fussing and crying vocal stimuli. As expected, babbling and cooing produced higher concentrations of values in the high scores (most liked) and crying and fussing showed a higher concentration in the low values (least liked). There were significant differences in ratings of least liked/most liked between stimuli; the most preferred stimulus was babbling followed by cooing (see Figure 2). The rating of babbling was significantly higher than the rating to cooing. Both babbling and cooing were more likely to evoke responses towards the “most liked” end of the scale than crying and fussing. Finally, there were no significant differences between crying and fussing in terms of “least liked”.

These results support the parental selection hypothesis (Locke, 2006), and suggest that babbling and cooing elicit attention and positive responses, with the canonical vocalizations being even more favorable than the noncanonical ones. As mentioned previously, adults prefer infant vocalizations that are syllabic versus vocalic (Bloom & Lo, 1990; Bloom et al., 1993), and there exists a possibility of a general attraction to complexity of sound across species (e.g., Bloom & Lo, 1990; Ryan & Rand, 1990; Bloom et al., 1993; Catchpole & Slater, 1995; Mountjoy & Lemon, 1996; Charlton, Filippi & Fitch, 2012). A preference for complexity and variability has been observed in adult human vocalizations (Anolli & Ciceri, 2002), and for music (Charlton, Filippi, & Fitch, 2012). Novel vocalizations capturing the attention of caregivers, thus signaling fitness and the ability to learn complex behaviors (Locke, 2006, 2008) seems likely when reflecting upon the results in facial expressions and subjective rating of the canonical vocalizations.
5.1.3 Possible hormonal and parental/nonparental influences on potential caregivers

Mothers showed a higher mean facial expressions of happiness compared to non-mothers while listening to the babbling stimuli, although it did not reach significance. Mothers also self-rated sadness higher than non-mothers when listening to crying. Mothers rated their anxiety levels as greater than non-mothers and irritation levels lower for the crying stimuli, and mothers had higher means for self-reported sadness for the fussing stimuli as compared to non-mothers at high estradiol point in their menstrual cycles, again approaching but falling short of significance. It is possible with a larger sample these difference may have reached significance, as observed in the post-hoc power analyses discussed earlier.

These results support the idea that hormones associated with breastfeeding may have evolved partially to assist women in the ability to respond sensitively to their infants, and not only to produce milk. Mothers who breastfeed exclusively have significantly higher levels of oxytocin and prolactin (Uvnäs-Moberg, Widström, Werner, Matthiesen, & Winberg, 1990), and oxytocin and prolactin keep mothers calm, relaxed and ready to care for their babies (e.g., Uvnas-Moberg, 2003).

Non-mothers at low estradiol point in cycle rated their irritation levels higher than mothers for the fussing stimuli. Mothers and non-mothers differed significantly in their desire to pick up the baby while listening to the fussing stimuli, with mothers giving the highest and non-mothers at low estradiol level in cycle the lowest. Differences approaching but falling short of significance in the desire to pick up the baby were obtained between mothers and non-mothers with assumed high levels of estradiol. These patterns of results were expected because of previous findings of the links between increased estradiol and increased infant handling (e.g., Maestripieri & Zehr, 1998) and reports of highest attachment feelings in human mothers (Fleming, Ruble, Krieger, & Wong, 1997).

As predicted, fathers also had significantly higher self-reported happiness levels and higher scores towards the “most liked” end of the rating scale for the babbling stimuli as compared to non-fathers. Physiological measures from fathers (RMSSD levels) were significantly higher than non-fathers while listening to the babbling stimuli, after controlling for baseline levels. Fathers had higher RMSSD levels while listening to cooing as compared to non-fathers, approaching but not reaching significance.
They also had higher SDNN levels while listening to crying as compared to non-fathers, which once again did not reach significance. HRV means were higher for fathers; possibly high variability in a small sample contributed to the failure to reach significance. Fathers showed a higher self-rated desire to pick up baby after listening to crying and fussing samples compared to non-fathers; this also did not reach significance.

The results for the male groups were as expected given that lower levels of testosterone have been linked with affiliative pair bonding and paternal care (Gray, Yang, & Pope, 2006), and fathers have significantly lower testosterone levels than both unmarried males and married non-fathers (Gray et al., 2006). Males with lower testosterone levels feel more sympathy and need to respond to infant cries (Fleming, Corter, Stallings, & Steiner, 2002). Fathers included in this study reported a minimum of three hours of daily contact with their infants. The decrease in testosterone in fathers, with the most dramatic decrease for men who spend more than three hours daily in childcare (Gettler, McDade, Feranil, & Kuzawa, 2011), may be an evolutionary adaptation to help fathers respond more sensitively to the needs of their offspring.

### 5.1.4 Other effects possibly related to hormone levels

Additional analyses also revealed some relevant differences, possibly related to testosterone level. RMSSD average in non-fathers was significantly lower while listening to babbling compared to mothers, fathers, non-mothers at high estradiol point in cycle, and non-mothers at low estradiol point in cycle. Non-fathers also had significantly lower average RMSSD average while listening to cooing compared to mothers and both non-mother groups. The difference between RMSSD during the cooing stimuli between non-fathers and fathers approached however did not reach significance. Non-fathers had significantly lower RMSSD while listening to crying compared to both non-mother groups.

The additional analyses also revealed some interesting differences possibly related to hormones linked with breastfeeding. The breastfeeding mothers had higher means for baseline levels of SDNN and RMSSD at rest as compared to both non-mother groups and both male groups (see Table 2). Pairwise comparisons for mothers revealed significant differences in SDNN from non-mothers at high estradiol point in cycle, but differences from non-fathers showed a possible trend, yet did not reach significance.
Mothers had higher baseline levels of RMSSD as compared to fathers, non-fathers, and non-mothers at high estradiol point in cycle, although this fell short of significance, perhaps due to the small sample size. It is interesting that the breastfeeding mothers showed higher HRV baseline levels because higher resting HRV has been associated with greater adaptive coping strategies in college students (Fabes & Eisenberg, 1997), lower levels of distress in young children (Fabes, Eisenberg, & Eisenbud, 1993), greater social competence in young children (Fabes, Eisenberg, Karbon, Troyer, & Switzer, 1994), and higher measures of active coping and acceptance for people who are recently dealing with a death (O’Connor, Allen, & Kaszniak, 2002).

5.2Rejected hypotheses

5.2.1 HRV measures

It was expected that participants would show higher HRV measures for the canonical vocalizations as compared to non-canonical, fussing and crying vocal stimuli. Increases in HRV, thought to be linked with positive mood states (Beauchaine, 2001; Porges, 1995; Thayer & Lane, 2000) were assumed to be associated most with babbling as compared to the other types of infant vocal signals. SDNN was slightly higher for babbling (mean=52.56) than cooing (mean=51.59), crying (mean=48.74) and fussing (mean=50.91), however these differences were not significant.

Breastfeeding mothers were expected to demonstrate higher HRV measures for all four stimuli as compared to non-mothers at low estradiol and non-mothers at high estradiol time period in their menstrual cycle, and non-mothers at the low estradiol point were expected to have the lowest scores as compared to the other two groups. Although there were no significant differences found in the HRV measures, mothers had higher means (SDNN, RMSSD) and also had higher means for resting baseline levels in HRV measures (refer to Table 14). Pairwise comparisons showed significant differences in mothers’ baseline SDNN from non-mothers at high estradiol point in cycle, and non-fathers, and in RMSSD as compared to fathers, non-fathers, and non-mothers at high estradiol point in cycle.

It was predicted that fathers would show higher HRV measures overall for all stimuli as compared to non-fathers. Although fathers showed higher measurements than non-fathers, there were no
significant differences between the fathers and non-fathers in terms of SDNN during babbling. There were significant differences between fathers and non-fathers in terms of RMSSD during babbling, with fathers’ RMSSD being higher than non-fathers as predicted. There were no significant differences in SDNN or RMSSD for the cooing stimuli between fathers and non-fathers, although means for fathers were higher and the difference in RMSSD approached but fell short of significance. The differences in SDNN and RMSSD between fathers and non-fathers for the crying stimuli also did not reach significance, even though means were once again higher for fathers. There were no significant differences in SDNN between fathers and non-fathers for the fussing stimuli, and mean values were almost the same. There were also no significant differences in RMSSD between fathers and non-fathers, although the values were higher for fathers.

5.2.2 Possible explanations

It is possible that significant differences for the infant vocal stimuli were not seen in all the HRV measures as predicted due to small sample size and high variability. Other factors like body mass index (BMI) and exercise can affect HRV and these demographic characteristics may not have been equally distributed across the groups. Power estimates discussed above indicate more data is needed to determine if there is an effect.

Fathers generally had higher means for most HRV measures with all four stimuli as compared to the non-fathers, and although many differences were not significant several approached significance. This is possibly due to the high variability and small subject sample size. Power estimates ranged from 47.9% to 53.6%, thus indicating more participants are needed in order to determine if there was an effect or not. Another consideration to make for this result, and the other HRV measures that were expected to be significantly different and were not, is that lack of group differences may be due to high correlations between the covariate and dependent variable. In ANCOVA, the covariate is supposed to remove variance due to factors that are statistically independent/unrelated to the dependent variable. When controlling for respiration in group differences in RSA, removing the effect of the covariate on the dependent variable that are correlated can remove relevant variance caused by group difference or
experimental manipulation, resulting in misleading conclusions that no meaningful differences exist (Allen, Chambers, & Towers, 2007). It is possible that the same principle applies here. When controlling for baseline, which is the standard approach in HRV research, the research question is “if these groups all started at the same level, do they differ in their response to the stimuli?” which is a different question from “do some groups have higher HRV during specific stimuli?".

5.2.3 Facial expressions

There were no differences observed in facial expressions of happiness for the cooing stimuli or facial expressions of sadness between mothers and non-mothers for the crying or fussing stimuli. No significant differences were found in facial expressions of happiness for the babbling stimuli between fathers and non-fathers, however, the means for fathers were higher as compared to non-fathers. Both fathers and non-fathers showed very low levels of facial expressions of happiness during the cooing stimuli. Facial expressions of sadness for crying stimuli were very similar for both groups. Fathers had higher means for sad facial expressions when listening to fussing, although they were not significantly different from the non-fathers.

5.2.4 Possible explanations

Females reacted similarly as revealed by facial expressions for all stimuli except for babbling. Female hormonal state and experience (motherhood) did not affect facial expressions for cooing, crying or fussing, indicating the possibility that canonical vocalizations may have evolved partially due to sensitivity of perception of breastfeeding mothers’ favorable response and care. It is possible that new breastfeeding mothers, probably the most important caregiver in the newborn’s life, is equipped through changes in hormonal state to respond most sensitively to the changes in vocalizations from noncanonical to canonical. Considering most mothers in our evolutionary history were breastfeeding ones, this extra sensitivity could be heightened via the experience of hormonal changes that accompany childbirth and nursing. The newborn’s main “audience” should be one that has the capability of perceiving these differences in vocal ability, resulting in continued or enhanced attention and elicitation of care.
Again, it is possible that small sample size and high variability may have been a factor in the lack of expected significant differences between fathers and non-fathers. Similar physiological responses in males may also be possible, and differences may be greater between more experienced fathers, rather than new first time dads, and non-fathers.

**5.2.5 Subjective ratings**

The hypothesis that participants would show higher ratings for desire to pick up the baby for the canonical vocalizations as compared to the non-canonical, fussing and crying vocal stimuli was not accepted. Babbling and cooing both scored lower in ratings for desire to pick up the baby, and were both significantly different from crying and fussing, which scored higher and were perceived as equal.

There were no significant differences between mothers and non-mothers in self-reported happiness and ratings of least liked/most liked when listening to babbling or cooing, or self-rated desire to pick up baby when listening to crying. No significant differences were found between mothers and non-mothers in self-reported happiness when listening to fussing.

Although babbling did result in significant differences as predicted for self-rated happiness and most liked scores, with fathers having higher ratings than non-fathers, no significant differences were found between fathers’ and non-fathers’ ratings for happiness or least liked/most liked scales when listening to cooing. Fathers showed higher overall scores for reported happiness and rated the cooing stimuli closer to the “most liked” end of the scale than non-fathers. No significant differences were found in self-reported happiness, sadness, anxiety, or irritation between fathers and non-fathers when listening to crying, although non-fathers reported higher levels of irritation. Although means for the fathers were higher than for the non-fathers, no differences were found between fathers and non-fathers in self-reported happiness or sadness when listening to fussing. There were also no significant differences in subjective irritation with fussing between fathers and non-fathers, although non-fathers rated this higher than fathers.

**5.2.6 Possible explanations**
Although the participants were verbally instructed to rate their desire to pick up the baby for “whatever reason, whether it be to interact, play, or soothe the infant”, it appears that most raters considered crying and fussing to be stronger indicators of need for interaction. It is possible that if the question were worded differently, perhaps rating the “desire to socially interact with baby” rather than “desire to pick up baby”, the canonical vocalizations may have elicited higher ratings. Participants showed the greatest percentage of happy facial expressions and also rated the babbling vocalizations with the highest scores for self-reported “happiness” and “most liked”. This makes it seem likely that the ratings for “desire to pick up baby” were interpreted by the participants in the sense of urgency for care because something was wrong, elicited most by crying and fussing, despite the verbal explanations provided.

Mothers and non-mothers reacted similarly in a few of the self-ratings for babbling, cooing, crying and fussing. It is possible significant differences were not found due to small sample size and/or high variability among participants. Another possibility is that babbling and cooing are universally pleasing (in females), thus resulting in similar ratings for happiness and how well liked the stimuli are. The canonical and non-canonical vocalizations were significantly different from fussing and crying for all participants in happiness ratings and received significantly higher ratings overall towards the “most liked” end of the scale as well.

Crying elicited strong responses from all participants in ratings of desire to pick up the baby, despite the fact that they were verbally instructed to rate this for “whatever reason, whether it be to interact, play, or soothe the infant”. This sense of urgency for care in a signal that relays something must be wrong, elicited most by crying, seemed to override any differences in subject’s hormonal states or experience with children. Fussing also affected all females in their self-rated happiness in the same way.

Non-canonical vocalizations along with crying and fussing did not elicit significant differences in many of the self-ratings between the two male groups. The direction of the differences in the means were as predicted, but not significant. This may be due to the small sample size and high variability within each group. Another possible explanation may be that the canonical vocalizations were the only stimuli able to elicit differences due to their importance in the evolution of the complexity of human vocalizations;
listening to babbling resulted in significant differences as predicted for self-rated happiness and most liked scores, with fathers having higher ratings than non-fathers.

5.3 Limitations and future directions

The results of the current study support the hypothesis for a general preference for vocal variability in infant signals, as observed in facial expressions across all participants while listening to babbling recordings, as well as in the subjective ratings of happiness and “most liked” stimuli. Evidence of the potential effects of hormones was also seen in some of the differing responses of new mothers and fathers as compared to participants without children. These findings have contributed to our understanding of the role of different infant vocal signals in elicitation of attention and care from potential caregivers.

Limitations of the present study include small sample size (n=60) with high variability among individuals, and a potential issue with the wording of one of the subjective rating items. It may have been more effective to reword “desire to pick up baby” as “desire to interact with baby”, as it seems the crying stimuli elicited a very strong response from all participants to soothe the infant. The hypothesis predicting that participants will show higher ratings for desire to pick up the baby for the canonical vocalizations as compared to the non-canonical, fussing and crying vocal stimuli may have been accepted if the rating item suggested social interaction. It is also possible that significant differences were not seen in all HRV measures that were predicted for the infant vocal stimuli due to effects of variation in body mass index (BMI) and exercise. The experimental setting was also far from a normal environment when one is listening to infant vocalizations, which may have had an impact on the results. The number of lines on the rating scales (11 points) may have also posed a challenge, and the wording on these subjective scales may have introduced the issue of different individual perceptions of the words used to describe feelings. For example, one person may perceive “anxiety” as a feeling of irritation, another person may define it as nervousness, and another may even think of it as excitement.

It would be interesting to include mothers strictly using formula feeding for comparison with the breastfeeding mothers in response to various infant vocal stimuli. An attempt was made to do so,
However many formula feeding mothers in the early post-partum period were either menstruating already, or attempting a combination of formula with breastfeeding.

Expansion of the examples of each type of vocalization would be valuable in determining if the same results would be revealed for a larger sample of babbling, cooing, crying and fussing recordings. It would also be clinically relevant to include samples of infant vocalizations from children later diagnosed with autism, speech/language delays, or specific language impairment and compare reactions to vocal samples from typically developing infants. Additionally, comparing parental responses from parents of typically developing children to parental reactions from parents of children diagnosed with disorders in speech and language would add an interesting new perspective.

Future studies might measure the levels of oxytocin, estradiol, and testosterone in saliva or urine in participants rather than relying on proxy measures. It would also be valuable to administer oxytocin via nasal spray to one group of females and males without children and compare them to a control group receiving a placebo nasal spray and measure physiological and subjective reactions to various infant vocalizations. Pre- and post- measurements of oxytocin would be collected as well. Comparing fathers who are with their infants more than three hours per day versus those who spend less than an hour daily in direct care and their reactions to the vocal stimuli would contribute to our understanding of paternal evaluation of infant vocalizations; testosterone levels could be measured before stimuli presentation. Examining physiological and psychological responses to infant vocal recordings after direct physical contact and play with infants will also add further to knowledge of perception of varied infant vocalizations and the effects of oxytocin.

Electrophysiological research evaluating the motor-neuron activity during exposure to babbling, cooing, fussing and crying may also help to provide clues to the evolution of spoken language. Comparative research looking at caregiver responses to offsprings’ vocal output with varying complexity could expand our understanding of its' function in both distant and closely related species, in the arenas of divergent and convergent evolution. Furthermore, additional detailed acoustic analyses of the various human infant vocal stimuli and corresponding facial expressions and HRV measures would be interesting in order to determine possible specific reactions to changes in frequency, intensity and rate, especially
during the babbling segments. Pinpointing the precise moments of greatest facial expressions of happiness and most positive HRV measures could potentially shed light on specific features of infant vocalizations that are preferred by caregivers. For example, if the segments that varied most in frequency, intensity and rate were favored this would help to support the theory of an attraction to vocal complexity, and the possibility that hominin infants varied vocal output to attract positive attention of potential caregivers.
VI. Appendices
6.1 Questionnaire for females

Questionnaire (females)

Participant: __________________________________________ Presentation CODE: ______________

What languages do you speak? ______________________________________________________________

Where are you originally from? ____________________________________________________________

How old are you? ________________________________________________________________

Are you a native speaker of American English? ______________________________________________

What is your sexual orientation? __________________________________________________________

Do you have any medical conditions (including a heart condition)? _____________________________

If yes, what are they? ______________________________________________________________

______________________________________________________________

Do you have a history of speech, language or hearing issues? _________________________________

Is your hearing normal? _________________________________________________________________

Are you currently in a committed relationship? _____________________________________________

Are you currently pregnant? ____________________________________________________________

Have you undergone any hormonal treatments? ______________________________________________

Do you use contraceptive pills, an implant or patches? ______________________________________

If yes, have you used them in the past six months? __________________________________________

Are you currently menstruating? __________________________________________________________

Do you have regular menstrual cycles? _____________________________________________________
Please circle the length of your typical menstrual cycle:
<22  22  23  24  25  26  27  28  29  30  31  32  33  34  35  36  37  >38

What day is this in your cycle? ____________________________________________________________

Do you have children? ________________ If yes, how many children do you have? __________
If you have no children, do you plan on having them? ________ If so, how many? ______________

How important is it to you to have children?

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
</table>
Not important | Somewhat important | Extremely important

Do you regularly take care of children (e.g., babysitting, childcare for a relative)? ______________
Are you a childcare provider (e.g., nanny, work in daycare)? ___________________________________
(Skip the following questions if you have no children)
How old is your child (children)? __________________________________________________________
Are you breastfeeding? ___________________________________________________________________
If yes, how often?
_______________________________________________________________________________________

Are you supplementing with formula? _______________________________________________________
If yes, how often per day? __________________________________________________________________
Are you feeding exclusively with formula? ___________________________________________________
6.2 Questionnaire for males

Questionnaire (males)

Participant: __________________________ Presentation CODE: __________________________

What languages do you speak?
_____________________________________________________________________________

Where are you originally from?
_____________________________________________________________________________

How old are you?
_____________________________________________________________________________

Are you a native speaker of American English?
_____________________________________________________________________________

What is your sexual orientation?
_____________________________________________________________________________

Do you have any medical conditions (including a heart condition)? ____________________
If yes, what are they?
_____________________________________________________________________________
_____________________________________________________________________________

Do you have a history of speech, language or hearing issues?
_____________________________________________________________________________

Is your hearing normal?
_____________________________________________________________________________

Have you undergone any hormonal treatments?
_____________________________________________________________________________

Are you currently in a committed relationship?
_____________________________________________________________________________

Are you a childcare provider (e.g., nanny, work in daycare)?
_____________________________________________________________________________
Do you regularly take care of children (e.g., babysitting, childcare for a relative)? ___________________

Do you have children? _____________ If yes, how many children do you have? ___________________

(Skip the following questions if you have no children)

How old is your child (children)? __________________________________________________________

How many hours per day do you spend in direct contact with your children? ___________________

If you have no children, do you plan on having them? _________ If so, how many? ________________

How important is it to you to have children?

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not important</td>
<td>Somewhat important</td>
<td>Extremely important</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
6.3 Self assessment form

SELF ASSESSMENT FORM

PARTICIPANT:

SELF ASSESSMENT: Please rate your feelings after listening to these sounds. Please remember to use entire scale.

Happiness:

Least                                           Most

Sadness:

Least                                           Most

Anxiety:

Least                                           Most

Irritation:

Least                                           Most

Helplessness:

Least                                           Most
Desire to go and pick up the baby:
VII. References


of the National Academy of Sciences, 106(22), 9115-9119.


Gustison, M. L., & Bergman, T. J. (2016). Vocal complexity influences female responses to gelada male calls. Scientific Reports, 6, 19680; doi: 10.1038/srep19680


64