More than meets the eye: from stress to scaffolding: a microgenesis of infant attention

Author: De Barbaro, Kaya

Publication Date: 2012

Peer reviewed|Thesis/dissertation
More than Meets the Eye: From Stress to Scaffolding,
A Microgenesis of Infant Attention

A dissertation submitted in partial satisfaction of the degree Doctor of Philosophy

in

Cognitive Science

by

Kaya de Barbaro

Committee in charge:

Gedeon O. Deák, Chair
Christine M. Johnson, Co-Chair
Marian Bartlett
Leslie Carver
Andrea Chiba
Edwin Hutchins
Gwen Littlewort
Mohan Trivedi

2012
The Dissertation of Kaya de Barbaro is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____________________________________________________________________

_____________________________________________________________________

_____________________________________________________________________

_____________________________________________________________________

_____________________________________________________________________

_____________________________________________________________________

Co-Chair

_____________________________________________________________________

Chair

University of California, San Diego

2012
DEDICATION

I dedicate this dissertation to myself.

I dedicate this dissertation to the love of learning, to life-long curiosity, to taking care in what we do and how we do it. To process.

I dedicate this dissertation to all of those kept out of educational institutions by systemic and individual oppression, to all of those who didn’t get the breaks they were promised when they started their degrees. To all those who can’t afford even a public education anymore, to all those who want to participate but can’t because of time, or money, or the sense that they just don’t fit in to this (strange) culture. To all those who never considered participating. To all those who weren’t given the time of day because they didn’t have letters after their name, to all those who have struggled through school with limited resources to get those letters to prove to someone, somewhere, that what they had to say was valid. To all those who tried and realized, at some point, that it wouldn’t work, that it wasn’t for them, that it wasn’t going to be possible. To all those whose talents aren’t considered valuable within the institution, to all those whose talents slipped through the cracks of the institution because they were missing one piece, or learning something else, then.
TABLE OF CONTENTS

SIGNATURE PAGE ........................................................................................................ iii
DEDICATION .................................................................................................................. iv
LIST OF FIGURES .......................................................................................................... vi
LIST OF TABLES ........................................................................................................... vii
ACKNOWLEDGEMENTS .............................................................................................. viii
ABSTRACT OF THE DISSERTATION ................................................................. xiv

INTRODUCTION: MORE THAN MEETS THE EYE: FROM STRESS TO SCAFFOLDING, A
MICROGENESIS OF INFANT ATTENTION ................................................................. 1

CHAPTER 1: MICRO-ANALYSIS OF INFANT LOOKING IN A NATURALISTIC SOCIAL
SETTING: INSIGHTS FROM BIOLOGICALLY BASED MODELS OF ATTENTION .......... 29

CHAPTER 2: TWELVE-MONTH “SOCIAL REVOLUTION” EMERGES FROM MOTHER-IN
FANT SENSORY-MOTOR COORDINATION: A LONGITUDINAL INVESTIGATION .. 41

CHAPTER 3: SENSORY MOTOR DECOUPLING CONTRIBUTES TO 12 MONTH SOCIAL
“REVOLUTION”: A LONGITUDINAL INVESTIGATION OF MOTHER-INFANT-OBJECT
INTERACTIONS ........................................................................................................... 97

CONCLUSION .............................................................................................................. 154
LIST OF FIGURES

Figure 1 Vigilance Scores ................................................................. 34
Figure 2 Typical 12 Month Imitation Episode................................. 105
Figure 3 Typical Four, Six, and Nine Month Sessions with Maternal Bids Indicated . 107
Figure 4 Synchronized Three-Camera View of Interaction ............... 112
Figure 5 Maternal Bids Histograms, Full Sample................................. 119
Figure 6 Maternal Bids Identification Parameters .............................. 123
Figure 7 Beta Distribution Range ................................................................. 126
Figure 8 Changes in Infant Sensory-Motor Coupling across first year .......... 130
Figure 9 Maternal Bids Histograms, Immature and Mature GH Couplers ............ 134
Figure 10 Maternal Bids Histograms, Immature and Mature HH Couplers ............ 135
Figure 11 Confidence Intervals for Beta Parameters ................................. 136
LIST OF TABLES

Table 1 Average (and SD) of Four Independent Theoretical Measures of Vigilance ... 33
Table 2 Correlation Coefficients among Four Theoretical Measures of Vigilance...... 34
Table 3 Correlation Results: Vigilance Index and Proportion of Looking Time to
   Different Regions of the Room, and Responses to Experimenter's
   Verbalizations ........................................................................................................... 34
Table 4 Developmental Trajectories ........................................................................ 89
Table 5 Average Median Bid Durations .................................................................... 124
Table 6 Best-fit Beta Parameters for Response to Maternal Bid .............................. 127
ACKNOWLEDGEMENTS

I would like to acknowledge my students, my mentors, my research community, and just generally everyone that has participated with me in this process.

In that order, I’d like to thank the original “Dream Team” and all the other students that followed: Sharon Hughes, Sophie Radford, and Maddi Taylor; Stephanie DeAnda, and Karmen Dykstra; Cathy Nyugen and Ioney Lau; Jonathan Hoang and Melody Lin; Melissa Real, Will Springer, and Noopor Gupta; Alicia Zuniga. Our time together was so fun that I coded more dimensions of mother-infant behavior than a single dissertation could possibly use. You each provided different things to the project: whether in your obsessive precision in helping export, forcing me to justify why (not) to code more, Boba breaks, or your patience with the chaos that is interaction data. Thank you to EACH ONE of you, my students!!!!

I thank my mentors, for their mentoring and the friendships that we’ve developed after so many years. I thank Deborah Forster for her confidence that the cognitive science department was a good place for me. For teaching me d-cog in a way that finally clicked. For years of her thoughts translated into pen scratches at P*te’s that inspired whole ways of conceptualizing my data. For her introductions to anyone doing anything relevant on campus. I thank Chris Johnson for all the fun we had sitting and pointing at a screen, framing forward and back. For teaching me to
see interaction as a multiparty micromodal configuration becoming coordinated through time. Our three hours a week was where my dissertation happened. Your attentiveness to my strengths and weaknesses as a researcher has helped me grow immensely. I thank Gedeon for his openness to all that I proposed, his love of (...occasionally... tangential) scientific banter and his painstakingly BEAUTIFUL dataset. And of course, for all the chocolate... To my mentors: thank you for your passion: for science, for social interaction, for knowledge; thank you for all the process (and pain!) that it took for you to get where you are today. I stood on your shoulders and I wouldn’t be where I am today had you not already come so far.

I thank my committee members !!!!!!

I thank Marta Kutas and Andrea Chiba, for their early cheerleading of my work, my talents, and the economic value of my mind, while always being confident to tell me exactly how I still needed to grow as a researcher. I strive to mentor as you do.

I thank all members of the Sequential Analysis Working Group, including but not limited to my co-organizer Deborah Foster, Chris Johnson, Paul Ruvulo, Jamie Alexandre, Nancy Renner, and Whitney Friedman. Our eclectic blend of computer scientists and ethnographers has been an indispensable part of my process of translating the mess of real world observations into systematic ways of characterizing interaction. I can still remember the sounds I heard from the audience the day I put
my original timeseries version of the “response to maternal bid” up, and the enthusiasm with which we worked to translate it into what it is today. You’ve been a wonderful community for me in the last few years, and I’m hoping I’ll be able to find colleagues like you wherever else my research takes me.

I thank my cohort mates, for early inspiration, and for so many years together. We all continue on, one way or another. Gaston, I miss you and I wish you were here. I thank my department mates, for being there for me, whether to print something, or to lend keys, or to rescue me from a burning car on the side of the freeway. Your comments on my work have pushed me and my research forward. Your supportiveness over the years has made me feel at home.

I thank my family, my mom, my sister, and my dad. For trusting in me, for the genes, for giving me the space to run around and knock into things when I was small, for endowing me with the privilege of being the THIRD GENERATION of women with PhDs in my family... wow, privilege,.. and wow, women ! I thank my extended family, my years of housemates and my friends. For being a safe space in a strange world, for inspiring me with your own ways, for dance parties and black beans.

I thank my partner Alex, who has always been willing to read my drafts, cook me food, open all the mail and not even seem too tortured with my insistence on balance and wild adventures. I love you big time.
I acknowledge a small number of the privileges I was endowed with at birth and those that I experience daily: my white skin, my able body, my healthy family who could provide me so much. For all the ways in which I so comfortably fit in to a North American institution. For the fact that my family could pay for my undergraduate tuition, for being able to tell me “I’d have the rest of my life to work”. I can only hope to take advantage of these privileges in such a way that will bring forth a more equitable world.

Almost last, but not least: I thank all the mother-infant dyads that participated in the study, as well as the staff and students in the UCSD’s Cognitive Development lab for creating the video tapes of the interactions that we used for this study, including but not limited to: Natalie Wong, Marybel Robledo and Jordan Danly!

And of course, last of all: Thank you to NSF for funding the collection of this data, part of the MESA Project (Modeling Emergence of Shared Attention; NSF SES-0527756) as well as NSF grant #SBE0542013 to the Temporal Dynamics of Learning Center, an NSF Science of Learning Center, for multiple times funding my time on the project.

Chapter 1, in full, is a reprint of the material as it appears in Developmental Science 2011. De Barbaro, Kaya; Chiba, Andrea; Deák, Gedeon O., 2011. The dissertation author was the primary investigator and author of this paper.
Chapter 2, in part is currently being prepared for submission for publication of the material. De Barbaro, Kaya; Johnson, Christine M.; Forster, Deborah; Deák, Gedeon O. The dissertation author was the primary investigator and author of this material.

Chapter 3, in part is currently being prepared for submission for publication of the material. De Barbaro, Kaya; Johnson, Christine M.; Forster, Deborah; Deák, Gedeon O. The dissertation author was the primary investigator and author of this material.
VITA

2005 Bachelor of Science, University of Toronto
2007 Master of Science, University of California, San Diego
2007-2008 Visiting Scholar, University of California, Berkeley
2012 Doctor of Philosophy, University of California, San Diego

PUBLICATIONS


FIELDS OF STUDY

Major Field: Cognitive Science

Studies in Developmental Science
Drs. Gedeon O. Deak and Leslie Carver

Studies in Computer Science
Drs. Mohan Trivedi, Marian Bartlett, and Gwen Littlewort

Studies in Ethnography and Distributed Cognition
Drs. Christine M. Johnson and Edwin Hutchins

Studies in Neuroscience
Dr. Andrea Chiba
ABSTRACT OF THE DISSERTATION

More than Meets the Eye: From Stress to Scaffolding, A Microgenesis of Infant Attention

by

Kaya de Barbaro

Doctor of Philosophy in Cognitive Science

University of California, San Diego, 2012

Gedeon O. Deák, Chair
Christine M. Johnson, Co-Chair

A widely accepted view in developmental science is that infants are biologically prepared to learn complex behaviors via their experiences interacting in a dynamic world with social agents who are responsive to their activity. However, current methods for studying the complex processes of interaction are lagging.
Traditional methods identify the *products* of development, and privilege infants’ internal structure as the cause of new social-cognitive skills and the essence of developmental change. In my dissertation, I use methods inspired by embodied and distributed cognition to precisely capture interaction as it unfolds from moment-to-moment. I use these methods to support and develop novel theories of infant development.

In my first study we expand upon classic looking time studies to collect a heterogeneous set of measures of infant looking-behaviors in a naturalistic environment. Our data suggest that looking-behaviors are not strictly a function of the time infants require to encode stimulus properties, but also are affected by factors such as uncertainty and psychological stress.

My second study is based in the tradition of naturalistic observation. In a longitudinal dataset spanning the first year, we precisely track the targets of infants’ and mothers’ multiple sensorimotor modalities—the hands, gaze, and mouth—as the dyads attend to one another and to various toys. From these data we characterize a number of developing behavioral trajectories in mother-infant object coordination across the first year. These trajectories ground an alternative theory to the current representational account of the development of social attention (see, e.g. Striano & Reid, 2006). At their youngest (4 months) infants raptly attend to their mothers and to her actions on toys. Over the first year, these early behaviors combine with
developing haptic articulation, sensory-motor decoupling, emergence of routines, and maternal scaffolding to produce, by 12 months, increasingly complex triadic interactions. Visualizations and analyses of multimodal patterns that preserve the rich temporal structure of the interaction allow us to confirm elements of this theory.

By precisely capturing infants’ interactions we characterize how the dynamics of the infants’ growing sensorimotor capacities, and the structure of the social environment provided by caregivers, jointly contribute to infant development.
References

INTRODUCTION: MORE THAN MEETS THE EYE: FROM STRESS TO SCAFFOLDING, A MICROGENESIS OF INFANT ATTENTION

Introduction

Traditional methods in developmental science are rooted in essentialist or representational accounts of cognition. These accounts have in common that they privilege internal processing or structure as driving cognition in the moment – or, on a longer timescale, as driving developmental change. A widely accepted alternative is that infants are biologically prepared to learn complex behaviors via their experiences interacting in a dynamic world responsive to their activity. On a short timescale the alternative to the essentialist viewpoint is that moment-to-moment changes in infants’ activity are codetermined by a set of heterogeneous factors, including the initiated and feedback actions of caregivers, infants’ own sensorimotor behaviors, and other elements of infants’ environments (e.g. degree of novelty).

While nowadays these are commonplace claims (Fogel & Thelen, 1987; Lewis, 2000; L. B. Smith & Thelen, 2003), our methods for studying these complex processes are lagging. In order to identify the processes that contribute to the dynamics (short timescales) and development (long timescales) of infant behavior, we need new methods that will allow us to capture the rich interactions between infants and their physical and social environments. Developing these methods is important. It is one
thing to say that caregiver, infant, and environmental factors somehow interact to influence infants’ development. It is another thing to prove how they interact, to demonstrate processes and mechanisms of development, and to look for general developmental processes that operate across contexts, across cultures, and across varieties of developmental phenotypes.

For these reasons, in my dissertation I have explored novel methods to supplement two traditional paradigms in developmental science. There paradigms, looking time studies and naturalistic interactions, are described below. My contribution to these paradigms was to track detailed, rich measures of infants’ sensorimotor contact with objects and adults, as those contacts unfolded in real time, in complex environments where infants and parents had considerable freedom to engage one another. These efforts lead to new insights about the nature of infants’ attention, social engagement, and action in complex environments. Those insights, in turn, compel some changes in our theoretical accounts of individual and age difference in infants.

Representational accounts of cognition have a long history in cognitive science. The dominant information-processing paradigm conceptualizes cognition as information processing. Under this framework, cognition occurs within the central nervous system, and it involves the manipulation of internal symbols representing elements of the outside world. The outcomes of these manipulations drive the
outputs of the system: actions. Learning is a matter of mapping inputs to appropriate outputs, and internally manipulating these outputs to generate complex, higher-order secondary output (e.g., action selection).

A central problem with this account is the process of selecting what information is relevant for a given task in a given moment—that is, what to represent—appears to be more difficult than the processing itself (Vervaeke, Lillicrap, & Richards, 2009). Sensory input does not come prepackaged into static “objects”, rather, the ability to recognize elements from the dynamic flux of environment is an accomplishment that involves a history of active participation of the organisms participating in it (Cantwell-Smith, 1998). Moreover, we do not passively receive input from a world that displays itself before us; instead, we move our bodies—especially our arms, necks, and eyes—to gain visual access to targets, and to pick up and manipulate the objects around us (Gibson, 1988; L. Smith & Gasser, 2005). These activities are part of the cognitive work of perception.

Given that we need to coordinate our actions with the real time dynamic changes in the world around us, actions cannot be pre-specified, but rather must be flexible and responsive. Coordinating our actions with the physical environment or the actions and reactions of social partners requires real-time assessment and adjustment (Suchman, 1987). As such, the Distributed Cognition approach (D-Cog, Hutchins, 1995) challenges the assumption that cognition occurs entirely within the
central nervous system (CNS), or even just within the body (see also Clark & Chalmers, 1998). Rather, it suggests that we can also see cognition—or cognitive work—in the way organisms organize their sensorimotor systems in real time during various activities (Clark, 2008). This is an insight that gives novel methodological weight to tracking rich behavior as it occurs.

Dynamical systems theory and developmental systems theory provide some similar insights in their critique of traditional models of cause and effect. Traditional western philosophy has essentialized activity and development as the effects of unilateral forces, such as genes, internal abilities, or choices (Oyama, 1985). For example, referring to development as a process of infants emerging “abilities” reinforces notions that novel behaviors originate from structures or processes internal to the infant. A more neutral formulation is to refer to changing behavioral outcomes as the product of a multiplicity of diverse, stochastic, and interactive factors, both internal and external to infants’ CNS, bodies, and even family system (Taylor, 2000). Like in the distributed approach above, one methodological implication of DST is that we must track an increasing number of elements of the system in order to understand the organization of activity or outcomes. That is, our theoretical insights cannot be any more accurate or comprehensive than the size, scope, and precision of the data that are used to derive those theories.
In this dissertation, I use methods inspired by distributed and dynamic theories of cognition to study the dynamics of infant attention at two timescales: First, as it unfolds in moment to moment “micro-events” nested within episodes of interaction with the world; second, on the much longer timescale of month-by-month development. The jumping-off points for this work are two very common paradigms for studying infant behavior: Looking-time studies, and naturalistic observation of mother infant interactions. These paradigms come from different historical and theoretical roots. Researchers who use looking-time measures typically use the framework of information-processing assumptions; those who use naturalistic studies often espouse dynamic or interactionist views of cognition and development. However, I argue that both traditions, in their prototypical forms, rest on representational or essentialist accounts of cognition. I will further argue that this deeply rooted reliance has stifled the progress of understanding the processes of infant development.

My first study, documented in chapter one, demonstrates the limitations of the classic representational accounts. Specifically, it shows that those accounts cannot predict or explain important features in the micro-behavioral dynamics of infant looking. Our study suggests that infants’ looking-time may not be strictly a function of how quickly they can encode stimulus properties. It may also be a function of factors such as uncertainty, the schedule of events in an environment,
and infants’ tendency to respond emotionally to those cognitive and environmental variables, for example, in terms of stress or anxiety.

Chapters two and three are based in the tradition of naturalistic observation of mother-infant interactions. We precisely track multiple sensorimotor modalities to characterize changes in mother infant object coordination across the first year. These changes ground a theoretical alternative to current representational accounts of the development of social attention (see, e.g. Striano & Reid, 2006).

Chapter 1

Developmental science has come to rely heavily on the use of looking-time studies (LTS). In a classic LTS, the infant sits in the lap of his caregiver, and stimuli are projected on a backlit screen. The dependent measure is the amount of time the infant attends to each stimulus; it is widely considered to be an index of the novelty of the stimulus for the infant (with longer duration looks indicating higher novelty or surprise). Novelty is considered inversely proportional to familiarity, thus LTS are thought to provide an index of what the infant knows (i.e., represents) or does not know, and how quickly they acquire new representations. For example, if the infant looks longer at a block passing through a barrier than at a block going around a barrier, that difference is taken to as evidence of certain spatial expectations or representations (e.g. Baillargeon, 1994). LTS have been used to make discoveries about infants’ looking-time differences to sights that are taken to illustrate a wide
variety of content domains, ranging from number (Wynn, 1998) to morality (Hamlin, Wynn, & Bloom, 2007). These discoveries are universally taken to reveal what infants know—that is, what representations infants “have.”

However, some researchers claim that the field of developmental science has an over reliance on looking-time methods (Aslin, 2008; Kagan, 2008). One concern is that the interpretation of LT data focuses entirely on the products of infant development rather than the dynamic processes of their emergence (Deák, de Barbaro, Goldstein & Schwade, in preparation). Additionally, the LTS paradigm uses one-dimensional visual attention measures to impute other, much more abstract, phenomena of interest (e.g., conceptual knowledge). This relies on the aforementioned theoretical assumptions that infant attention is determined by internal information structures.

Specifically, the relationship between looking-time and complexity or novelty is based on comparator theory (Clifford Jr & Williston, 1993; Sokolov, Worters, & Clarke, 1963), which states that when infants look at an object they are collecting information to build a mental representation of it, and when they have completed ”building” the representation. Thus, the theory depends on the idea that each stimulus has an amount of information to be processed by the infant.

The studies described above compare looking (i.e., fixation) durations within subjects, to assess which stimuli are more novel, complex or surprising to infants as a
group. Other studies compare durations between subjects in order to estimate individual differences in “information processing” (IP) speed. IP speed is thought to be a static (i.e., high-order, general, and internally determined) and stable trait that varies between individual infant. For example, it has often been tested for its potential to predict later measures of cognitive functioning, including IQ (Bornstein & Sigman, 1986; Colombo & Mitchell, 1990).

Although some researchers consider broader neurological (Colombo, Richman, Shaddy, Follmer Greenhoot, & Maikranz, 2001; Richards & Casey, 1992) and ecological (Franchak, Kretch, Soska, & Adolph, 2011) underpinnings of infant looking behavior, LTS continue to be used and interpreted based on the classic theories of encoding.

Moreover, the experimental paradigms of traditional LTS further reinforce the notion that “internal” processes such as encoding determine behavioral outcomes. Traditional LTS remove infants from their natural surroundings, replacing a complex environment with a dark room with a single lighted screen upon which 2-D, often static stimuli are presented. Testing room walls are barren and there are no objects in reach for the infant to manipulate. There is no social partner. By contrast, natural environments for infant attention include many different types of dynamic stimulation (moving people, dogs barking, lights flashing) including objects of various sizes and shapes that the infant can use his full body, including his hands and mouth,
to explore (Rochat, 1989). This rich array creates a situation in which infants must distribute their attention to some elements of the environment to the exclusion of others. In natural environments, caregivers facilitate this process by providing multimodal cues as to what is interesting in the environment. They also can facilitate more differentiated and dynamic states and patterns of attention by physically altering the proximity and motion of objects (Lawson, Parrinello, & Ruff, 1992).

The predominant reliance on a stark testing environment by psychologists suggests an implicit dualism between the agent and its environment. Specifically, underlying this paradigm is 1) the notion that the infant can be separated from its environment with the goal of examining his attention “context-free” and 2) that what we measure in this stripped down environment of the testing room will generalize to attentional behaviors in the real world. Of course, no one would argue that these factors do not affect looking behaviors at all, and indeed, many researchers who use LTS recognize the importance of social and ecological factors for development. However, taking seriously the tenets of an interactionist framework of cognition suggests we can learn substantially more about the dynamics of infant attention by observing it in a richer more naturalistic environment.

In our own study of visual behaviors in the laboratory, we provided infants with visual access to the entire testing room by turning on the lights and displaying stimuli on six monitors located in all directions around the infant. Monitors displayed
a random sequence of musical video stimuli and infants were seated across from an experimenter. We also increased the degrees of freedom of the infants’ attentional behaviors by placing them in a seat designed to allow them to comfortably rotate their gaze to the monitors behind them as well as to incidental stimuli such as the chair rungs and their own toes (which were also available for manipulation).

This novel testing ecology highlights the complexity of infant attentional behavior as it unfolds in time. It allowed us to explore some of the other contextual factors that might influence looking behaviors. In particular, we were interested in the effects of uncertainty—defined not just as a novel picture but as series of events that varied in predictability—on looking behaviors. Moreover, it allowed us to collect many heterogeneous measures of the dynamics of attention. LTS typically report only a single measure of looking behavior, look duration. Thus another unique contribution of the first chapter was to examine the relationships between many heterogeneous measures of infant looking that had not been looked at all together in the past.

In chapter one, we review animal studies and computational neuroscience models that indicate that attending in non-human animals and human adults is a function of the degree of uncertainty in the environment (“Aston-Jones Model of Attention” (AJMA); Gary Aston-Jones, Rajkowski, & Cohen, 1999; Yu & Dayan, 2005). More specifically, AJMA states that as uncertainty or arousal increases, animals shift
along a continuum from more casual exploring and learning about the environment, with the capacity to focus on elements of interest, to a more vigilant, stimulus driven assessment of sources of uncertainty. Neurologically, this is paralleled by moment to moment adjustments in excitatory impulses from projections of norepinergic cells originating in the locus coeruleus to many areas of cortex and basal ganglia (G Aston-Jones, Chiang, & Alexinsky, 1991).

This model indicates a pathway by which characteristics of the physical and social environment can impact infant attention during a laboratory task. Previous studies have indicated that novelty, especially in the absence of a sensitive and responsive caregiver, can lead to infant arousal or stress (Gunnar, Larson, Hertsgaard, Harris, & Brodersen, 1992). AJMA suggests that such increases in arousal might contribute to fast looking. Current infant research has not assessed how changes in infant arousal mediate infant looking behaviors, although it is clear that reductions of arousal are important for focused attending, and that increases in arousal are related to higher distraction (Lansink & Richards, 1997).

Testing the contributions of stress to infant looking behaviors was one of the aims of the research we report in chapter one (published in the journal Developmental Science; de Barbaro, Chiba, & Deák, 2011). While not detailed in the publication, our procedure included a mild social manipulation meant to up-regulate infant arousal and stress: the still face task (Tronick, Als, Adamson, Wise, & Brazelton,
Following the literature on LC activation, we predicted that stress due to reduced social contingency by mothers would increase vigilance as observable in a number of infant looking behaviors, including decreasing the duration of infant looks (Arnsten, 1998). If look duration is a function of arousal it becomes ambiguous how to interpret classic LT results: shorter looks could be due to stable differences individual differences in something akin to processing speed or alternatively, to acute stress. Because we collected a number of heterogeneous behavioral measures of vigilance, we were able to explore the coherence between these measures and test the validity of the AJMA for an infant sample. This was interesting as prior infant studies had never examined such a range of looking behaviors in a single sample of infants. Moreover, this allowed us to test whether the behavioral measures we collected were more consistent with AJMA or with the traditional IP model of attention.

Our measures of vigilance included four measures: how quickly the infant responded to semi-regular changes in his environment (the turning on and off of “target” monitors), how often the infant reoriented to these changes in the environment, and the rate and duration of fixations. These measures showed a high degree of coherence in individual infants, in parallel with findings from AJMA. Results also indicated that less vigilant infants showed greater habituation to the trial

\[ Note \text{ that the still face is mentioned in the methods section as mild social manipulation.} \]
structure and more attentiveness to less salient stimuli during periods of high attentional competition. This pattern of results could not be explained by the standard IP model but can be explained within AJMA. Specifically, the IP model predicts that fast looking infants will habituate faster to the trial structure, while AJMA suggests that high arousal can lead to acute increases in vigilance, including temporary reductions in habituation as the system returns to a threshold of certainty about the environment (Arnsten, 1998).

In sum, our analysis of the indices and correlates of vigilance in infants in a semi-naturalistic paradigm supports and generalizes Aston-Jones et al.’s (1999) attentional theory. By defining indices of vigilance derived from non-human animal studies, we found that human infants show the same patterns of co-occurring indices that are documented in animal literature. This is one of the first human studies, and the first infant study, to investigate micro-indices of vigilance predicted by AJMA. The relations among our collected measures cannot be explained by conventional IP accounts of infant looking. Additionally, the AJMA model provides a framework to integrate universal behaviors ignored in IP theory, including social attentiveness, individual differences in temperament, stress reactivity, and, most generally, distribution of attention in natural (e.g., lit, dynamic, cluttered) environments.
Chapters 2 & 3

The over-representation of highly artificial and oversimplified tasks and settings in the infant literature prevents us from discovering the processes that lead to change in infants’ behaviors. By removing infants from the environments in which they typically interact with people and objects, we cannot investigate how the infant’s environment supports development. For example, removing social partners restricts us from observing how those partners provide novel challenges for infants as their behavioral repertoire expands. Also, by restricting infants’ behaviors in the testing room—either literally or de facto, by measuring so few behaviors so sparingly—we lose access to most of the changes in behaviors through development.

Naturalistic studies of interaction have long been a part of developmental science, perhaps uniquely so among the cognitive sciences. Researchers observing naturalistic interaction have long espoused a more interactive ecological view of development cognition (Gibson, 1988; Vygotsky, 1978) in which development emerges as a product of many nested intersecting factors (Bronfenbrenner, 1979). However, oftentimes these accounts simply provide high-level suggestions or arguments, or qualitative descriptions lacking the methodological rigor to precisely capture the manner in which these factors affect behavior. Sometimes researchers have made quantified observations of naturalistic data, but then coded the data at such high level that most processes of interaction were bypassed, and only the high-level products (phenotypes) are treated as variables of interest. For example,
Bakeman and Adamson (1984) code the types of dyadic coordination that the mother and infant achieve at various time points, but not the particular behaviors by which coordination is achieved.

Without data that captures the processes of interactive behavioral emergence, there is no way to test or develop an ecological account, except by pure speculation. This lack of data leaves researchers in a position to propose theories that very few essential factors cause shifts in structure or process. Those hypothetical causes will typically be construed as internal, because that is the bias of researchers who take a psychological perspective on the world. By contrast, research that tracks multivariate, dense interactions between infants and their physical and social environments opens up the possibility of identifying and testing heterogeneous multivariate processes that lead to development. Technological progress in video (and increasingly, sensor) technology increases the plausibility of such rigorous research without imposing such extreme constraints as are typical of LTS studies. We can use video and sensor technologies, and increasingly machine (automated) processing of video data, to capture the precise activity of the developing infant in relation to the mother and in relation to changes in the environment. We can also use statistical and quantitative models that go beyond the traditional psychologist’s toolkit in order to identify the complex elements of the interaction are changing through development, and which of these elements have temporal and contingent
precedence. Those factors with precedence are the factors most likely to be contributing to development.

This logic forms the basis of the methods in chapters two and three. In these chapters, we analyze a longitudinal dataset of mother infant interaction to study the development of shared actions on objects over the infant’s first year. The period of nine to 12 months is thought to bring a new ability to coordinate with and share the objects of another person’s attention—for example, following the gaze of the mother, or playing a ball-tossing game together. This has been termed “triadic” or “you-me-it” attention and is thought to be the basis for future learning, including early language (Bruner, 1975; M Tomasello & Farrar, 1986). In this work we focus on the development of one type of triadic attention, shared actions on objects, which includes both imitation and games (M. Tomasello, 2007).

To study the development of shared actions on objects, we used video recordings of sessions when the infants were four, six, nine, and twelve months of age. We analyzed the videos frame-by-frame, tracking the details of participants’ attention to one another and to the objects. Specifically, we operationalized attention in infant and mother as comprised of multiple independent sensorimotor modalities, including vision (i.e., direction of gaze) and haptic exploration (i.e., both hands). We also coded changes in affect, and other cues relevant to the dyadic interactions, such as toy presentations by the mother. Previous accounts of the
development of triadic attention find “a curious gap” in measures of social attention based on tracking infants’ “general” (amodal) focus of attention (Adamson & Bakeman, 1991, p. 21). In contrast, by tracking the particulars of attending across modalities and both parties, we can document how the distribution of attentional resources changes with age, as how these changes contribute to social development.

In chapter two we detail the theoretical basis and motivation for our unique methods of studying the interaction. Moreover, we use these methods to identify a set of action trajectories that we claim provide a novel account of the development of shared actions on objects. A few important elements of these methods, and the conclusions we were able to draw, are as follows.

First, at macro-level we identify events that regularly occur within each interaction and across months (Forster & Rodriguez, 2006; Hutchins, 1995). For example, at all months infants directed their sensorimotor modalities to objects. At all months, mothers made bids for their infant’s attention to a toy. These are macro-level variables (i.e., unfolding over 2 to 40 sec), in that we can describe them in terms of a configuration of more micro-level (i.e., moment-by-moment, or within fractions of sec) elements. For example, we identify a maternal toy bids when the mother manipulates a toy that has not been “in play” in front of the infant for at least five sec.
Next, we identify and track the activity of various micro-level elements that participate in the macro-level event. In our account, all micro-level elements are visible behaviors that occur between the mother, the infant, and shared objects.

At each age, the mother and the infant attend to each other and to toys with each of these modalities, but the manner in which the modalities get configured across the ages is different. The word *configuration* refers to the temporal dynamics of how the micro-level sensorimotor actions construct macro-level events. For example, our youngest infant’s gaze and hands were always coupled when attending to objects, and as they got older, gaze and hands were more likely to be directed at different objects. When we detail these configurations at each age, we create a longitudinal history that can provide clues to the development of triadic attention.

Our approach began with an extensive observational study of five dyads, detailed in chapter 2. In that study we identified longitudinal changes in the configurations of micro-level changes in multiple sensorimotor modalities of mother and infant, as they unfolded within in two types of macro-level configurations: while infants attended to toys, and while infants responded to mothers’ toy bids. This was a nested-levels strategy in which we identified changes in the micro-level processes that tended to precede and constitute the macro-level events. We found in these data a number of simple behavioral trajectories that slowly develop across the first year. These trajectories include an increasing number of loci of attention, refinement
of haptic articulation, the emergence of routines, and increasing continuity across episodes of activity. For example, while our youngest infants modalities were nearly always targeted towards a single toy, what we referred to as “coupled”, over the course of the first year, infants began to decouple modalities such that they could attend to multiple toys simultaneously.

Over the first year, as infants produce increasingly complex behaviors with objects, they produce ever more complex responses to their mother. Also, mothers’ behaviors towards the infant and the objects also changed and diversified. By altering the infant’s environment, the mother “co-shapes” each interaction, and thus contributes to the development of triadic attention. More surprising, the data show that the emergence of classic triadic behaviors around nine to 12 month is not sudden or discontinuous, but rather is a culmination of continuous changes along these trajectories.

In the third and final chapter, we formalize one measure of mother-infant-object coordination, and use systematically coded videos from the same longitudinal dataset to capture continuous change in this measure over the first year. Additionally, we create a measure of one of the trajectories that we theorized contributes to the development of triadic attention. This is a measure of how the infant distributes its attention across multiple toys: specifically, the proportion of the session that pairs of the infant’s modalities (gaze and hands, or right and left hand)
are coupled (i.e. in contact with a single target) or decoupled (in contact with two distinct targets). We use converging methods to verify that increased decoupling across sessions or within sessions is associated with changes in the infants’ response to mothers’ bids for attention to objects. In particular, as they decouple more, infants increasingly distribute attention between a bid toy and the objects of ongoing attention in such a way that allows them to visually “check in” on their mother’s actions as she manipulates a set of toys. We observed the same patterns of attending during episodes of imitation, suggesting to us that this type of attention distribution is important for participating in shared actions where the infant incorporates elements of the mother’s manipulation into his or her own actions. Our data suggests that infants’ continuously developing skill at distributing attention across toys and partners increases the potential complexity of interactions with mother, culminating in episodes of imitation or games.

Again, the novel account of the development of triadic attention addressed in chapters two and three challenges the traditional belief that the appearance of imitative behaviors is due to a shift to a novel conceptual/representational capacity to infer the intentions of a social partner (Striano & Reid, 2006; M. Tomasello, Carpenter, Call, Behne, & Moll, 2005). The crux of the quantitative analysis in Chapter 3 is in creating a single index of mother-infant-object coordination that applies throughout the first year.
The studies described in chapters 2 and 3 share certain unique features that imply a novel theory. One feature is the number of dimensions of the interaction that we tracked, including the gaze and hands of both the mother and infant. Typical quantitative studies of social interaction may capture one or two dimensions of behavior, for example, dyadic attending at a macro-level, (e.g. Bakeman & Adamson, 1984) or gaze or facial expressions at a micro-level (Yale, Messinger, Cobo-Lewis, & Delgado, 2003). Further, micro studies of social interaction rarely capture the activity of the hands during social interactions. Our study expanded on these past studies to include many components of micro dynamics of attention, including capturing the activity of the hands, which proved particularly important for capturing changes in the development of shared actions on toys. Additionally we captured both micro- and macro-level elements of the interaction. This allowed us more perspective on the activity of our dyads than has previously been captured.

Another unique feature of our work is the analyses that were designed to capture the expanded dimensions of the interaction. Typical studies of social interaction examine the dynamics of a small range of states within one or two dimensions. For example Messinger et al (2012) model patterns in transitional likelihoods of infants’ looking at versus away from their mother, given how long an infant has currently been looking at the mother. This model has one predictor and one outcome dimension: a nicely parsimonious way to begin, and unlikely to reveal
very comprehensive social dynamics. Yale et al (2003) model how infant gaze
dynamics are affected by whether or not the mother is smiling. Again, this study uses
one predictor and one outcome variable. In our own study we use multiple
computational methods to characterize co-activity of pairs or triplets of dimensions
(gaze and hands, two hands), as well as nested levels analyses to identify important
macro-level elements of the interaction, capture changes in the microdynamics of
mother/infant attention over a longitudinal period.

These features allow us to provide an account of triadic attention that is
grounded in continuous changes in the activity of our participants rather than a
simple shift in internal structures. By focusing on changing processes of attentional
behaviors across the first year, we gain a novel understanding of this important
developmental phenomenon.

Conclusion

In summary, this dissertation investigates a number of concrete adjustments
to typical developmental paradigms. These adjustments are motivated by a view of
development that is focused on learning through interaction rather than by
maturation of internal structure. Additionally, they are motivated by a view of
behavior being determined in the moment by a heterogeneous set of factors both
internal and external to the infant. The argument is not that there is nothing to be
gained from previous studies. However, taking seriously the tenets of a biologically
and ecologically informed, interactionist, D-Cog framework of social cognition suggests new ways to go beyond traditional paradigms for studying infant attention and development.
References


Micro-analysis of infant looking in a naturalistic social setting: insights from biologically based models of attention

Kaya de Barbaro, Andrea Chiba and Gedeon O. Deák

Department of Cognitive Science, University of California, San Diego, USA

Abstract

A current theory of attention posits that several micro-indices of attentional vigilance are dependent on activation of the locus coeruleus, a brainstem nucleus that regulates cortical norepinephrine activity (Aston-Jones et al., 1999). This theory may account for many findings in the infant literature, while highlighting important new areas for research and theory on infant attention. We examined the visual behaviors of n = 16 infants (6-7 months) while they attended to multiple spatially distributed targets in a naturalistic environment. We coded four measures of attentional vigilance, adapted from studies of norepinephrine modulation of animal attention: rate of fixations, duration of fixations, latency to reorientation, and target 'hits'. These measures showed a high degree of coherence in individual infants, in parallel with findings from animal studies. Results also suggest that less vigilant infants showed greater habituation to the trial structure and more attentiveness to less salient stimuli during periods of high attentional competition. This pattern of results is predicted by the Aston-Jones model of attention, but could not be explained by the standard information processing model.

Introduction

Measures of infant looking behavior have long been considered an important indicator of attention and processing speed. More recently, developmental science has begun incorporating biological frameworks to better characterize the function and processes of attention-mediating and information-encoding behaviors. For example, concurrent measurement of physiological data such as heart rate has been successfully used to break gaze into component phases of orienting, maintenance, and shift-preparation (Colombo, Richman, Shaddy, Greenhoot & Maikranz, 2001; Richards & Casey, 1991). Current biological literature also posits that the neuromodulatory state of an organism can drive visual attention. Specifically, theoretical advances posit that several behavioral indices of attentional vigilance are paralleled by the activation of the locus coeruleus, a brainstem nucleus that regulates cortical norepinephrine (NE) activity (Aston-Jones & Bloom, 1981a; Aston-Jones, Rajkowsi & Cohen, 1999). These indices include several 'micro' behavioral features of looking, including rate and duration of individual saccades. To our knowledge, no study has explicitly assessed the validity of this theory in human infants. However, there are parallels between the Aston-Jones model of attention (AJMA) and a more traditional model in developmental studies that describes some of these measures of attention in terms of information processing (IP) constructs. Thus, research on the coupling between the neuromodulator norepinephrine and animals' attentional states is relevant for (1) refining our understanding of infant attention, and (2) deriving a theoretical framework for measuring and testing novel hypotheses about micro features in infants' looking.

The current work is organized as follows. First, we review the animal literature on NE and behavior, and document its relations to the existing infant literature, specifically the traditional IP model of attention. Although Aston-Jones and colleagues' work clearly shows the role of the LC and NE in attention modulation, it is impossible to measure these effects directly in human infants, and difficult even to measure them indirectly. Thus while we do not measure LC/NE activation in infants, we review the animal research because it lays out the empirical and theoretical basis for the analysis of behavioral measures in the current study. Specifically, we can test the validity of this theory for explaining infant looking by examining particular behaviors that should arise in certain circumstances. Some of the relations between these indices are counterintuitive under current information processing models. By testing the coherence of these measures of infant looking behaviors, we determined that the Aston-Jones model can account for patterns of visual attention that were heretofore unexplained in human infants.
Aston-Jones model of LC modulation of attention

The activation of NE neurons is widely thought to function in the modulation of a continuum of attentional behaviors (Amaral & Sinnamon, 1977; Aston-Jones, Chiang & Alexinsky, 1991; Berridge, Page, Valentinio & Foote, 1993). A nucleus of NE neurons in the LC ascends extensively to limbic and cortical areas that are believed to enact and alter attentional states (Morrison & Foote, 1986; Posner & Petersen, 1990).

At the lowest end of the attentional continuum are activities such as sleeping or resting, as well as automatic or habitual behaviors that require minimal attention to environmental stimuli such as grooming or drinking (Aston-Jones & Bloom, 1981a). These activities are associated with low frequency oscillatory release of NE, called phasic LC activity.

Intermediate levels of NE release are associated with ‘selective’ or ‘focused’ attention wherein animals are active but only moderately responsive to stimuli outside of their immediate focal area (Aston-Jones & Bloom, 1981b; Rajkowski, Kubin & Aston-Jones, 1994). These states occur in environments in which goals and stimuli stay relatively stable, where environmental processing demands are low but not trivial (Aston-Jones et al., 1999).

At higher frequency oscillations of NE release by LC neurons, called tonic firing rates, animals show a higher degree of vigilance or ‘anticipatory readiness’ to their surroundings (Aston-Jones & Bloom, 1981a). Behaviorally this can be assessed via an increased likelihood of reorienting to novel stimuli, and decreased latency to reorient. For example, Aston and Jones and Bloom (1981b) found that LC activity spiked when animals redirected their attention, and blocking NE receptors (thereby extinguishing the effects of NE release) reduced rates of stimulus reorientation. Increased NE is also associated with a high rate of short fixations (Rajkowski et al., 1994). Highly responsive animals have low thresholds for responding to novel visual stimulation and therefore do not maintain gaze to any one location for very long. At extreme levels this results in continuous fast scanning.

This decreased threshold to respond to features of the environment leads to increased and broadened sensory access, allowing animals to more quickly identify and adapt to rapidly changing circumstances (Aston-Jones et al., 1999). This is adaptive in situations of novelty and unpredictability, ranging from minor (e.g. a sound coming from an unknown source) to major (e.g. an attack by a predator). Modeling accounts show that the amount of time of NE neurons maintaining the degree of perceived uncertainty (Yu & Dayan, 2005).

However, experimental and observational work shows that states of high vigilance have a reciprocal cost. Specifically, the readiness associated with high vigilance also leads to increased distractibility and difficulties focusing attention to a central location. Rajkowski et al. (1994) found that moderate NE release in old-world monkeys corresponded with steady foveation to a fixed stimulus, and relatively low responses to distractors. By contrast, high NE release was associated with difficulty of central foveation and increased ‘scanning’ eye movements, and a decreased threshold for response to distractor stimuli (i.e. false alarms).

In sum, LC-dependent levels of cortical NE levels are tightly related to a number of attentional behaviors. These relations have also been observed in adult humans (Skosnik, 2000; Clark, Geffen & Geffen, 1989). To our knowledge, however, no study has documented these relations in human infants. AJMA does make predictions that fit the results of some studies of infants; however, in evaluating these we must compare our models and predictions of information processing theory.

AJMA and current information processing models

Infant information processing efficiency is usually assessed in terms of the duration of the longest look (i.e. ‘peak look’) to a novel stimulus. Theoretically, the relationship between looking time and processing speed is based on comparator theory (Clifford & Williston, 1993; Solokov, 1963), which states that when infants look at an object they are collecting information to build a mental representation of it, and when they look away they have completed the representation. This is supported by empirical research showing that infants with shorter peak (i.e. longest) fixation durations, or ‘short lookers’, show similar novelty scores (evidence of learning) as ‘long lookers’, which is interpreted as evidence that they process the same amounts of information in a shorter time.


A number of predictions are shared by the information processing (IP) model and AJMA. For example, infants identified as fast processors disengage more rapidly to orient to perceptually placed stimuli in a subsequent task (Frick, Colombo & Saxon, 1999). Similarly, during a television viewing task with distractors, Richards and Turner (2001) found that long looks were followed by longer latencies to respond to new distractor stimuli, whereas short looks were followed by shorter response latencies. These studies parallel findings in the animal literature that increasing vigilance corresponds with increased speed of disengaging and increased probability of reorienting to a new, peripheral event.

Fast looking infants also show a broader spatial distribution of gaze (Colombo & Jusowik, 1998; Jankowski, Rose & Feldman, 2001; Orlan & Rose, 1997) than longer looking peers. Experimentally broadening the spatial distribution of slower looking infants’ gaze increased their familiarity with stimuli to levels shown by fast lookers (Jankowski et al., 2001). AJMA suggests that this broadening occurs due to the increased readiness to respond to stimuli associated with vigilant states. AJMA
modeling work has also shown that such 'scanning' patterns are functional for certain types of learning (Yu & Dayan, 2005), as Jankowski et al. (2001) identified.

In contrast to findings such as these, there are a number of behavioral elements and relations that are predicted by the AJMA, but are absent from IP theories of infant attention. Studying the differences between the two theories can help to refine models of infant attention.

Two differences between IP theory and AJMA are particularly important. First, IP considers behaviors related to vigilance attentional states (e.g. faster reorientations) as 'better' or more mature. By contrast, AJMA proposes that each end of the continuum of attentional states has reciprocal costs and benefits. That is, the readiness associated with high-tonic activity states (i.e. vigilance) also entails increased distractibility, as the threshold of stimulus salience that will elicit reorientiation from a central gaze location is reduced. In other words, attention is more driven by peripheral sensory information. By contrast, at more moderate levels of vigilance, organisms have increased opportunity to self-direct attention and achieve focus, as interference from peripheral events is relatively dampened (Aston-Jones et al., 1999). This is adaptive for tasks that require sustained attention to a centralized location of low perceptual salience, such as studying or waiting for a hard-to-detect cue (Aston-Jones et al., 1999; Rajkowski et al., 1994).

Its basis in comparator theory leads traditional IP to ignore the tradeoff between focused attention and high vigilance, because the outcome of a look is presumed to always be the same: a complete representation of one focal stimulus. However, models of active vision show that we should not conceive of stimulus information as 'out there' in discrete, prepackaged, perceiver-independent 'chunks'. Instead, organisms elaborate features of the environment as a function of prevailing goals and behavioral demands (see reviews by Aloimonos, Weiss & Boddyvadicay, 1988; Findlay & Gilchrist, 2003; Gibson, 1998; Hayhoe & Ballard, 2005). Under these models, looking away does not signify 'completing a representation', but rather, given the current task, that there is more to be gained by looking elsewhere. Conversely, continuing to look can also suggest additional gains from the objects of attention (see also Courage, Reynolds & Richards, 2006, and Ruff & Saltarrelli, 1993, for additional evidence from infant studies for these claims). One way to resolve this tension between the two models is by investigating the relationship between vigilance and 'attentional tradeoffs', or time spent looking to high vs. low salience stimuli. Specifically, AJMA predicts that time spent attending to low salience targets in the context of peripheral higher salience 'competitors' will be reduced under states of increased vigilance. This reduction would not be predicted by traditional IP accounts.

Second, AJMA highlights the influence of short timescale factors on attention, or task- and state-specific adaptation of attention. Both naturalistic observations and experimental manipulations show that changes in the firing rates of LC/NE cells, and their behavioral consequences, shift from moment to moment. These shifts are due to internal and external factors related to stress and uncertainty. For example, an unexpected noise can cause both a spike in NE and a redirecting of gaze, illustrating the potential for novel or unexpected stimuli to affect levels of vigilance (Aston-Jones & Bloom, 1981). This is also evident in activation of LC neurons by a variety of physical and psychological stressors (Abercrombie & Jacobs, 1987; Morilak, Pornai & Jacobs, 1987a, 1987b). Again, this is unsurprising as increasing vigilance has the effect of decreasing uncertainty about the surroundings (Yu & Dayan, 2005). The typical pattern of LC/NE activation is one of dynamic adaptation in response to stability or change in the environment: as uncertainty is reduced, so is activation of LC/NE (Aston-Jones & Cohen, 2005).

By contrast, the traditional IP model does not incorporate such factors, but hypothesizes that looking times reflect stable individual trait(s) of information processing. For example, the traditional IP model interprets 'fast' shifting of attention as an indicator of intelligence (Frick et al., 1999). Although AJMA does not preclude stable individual differences in cognitive efficiency, it would allow that this behavior might be a sign of a temporary state, such as a response to novelty or acute stress. This contrast between models suggests that it is important to control for or measure environmental factors such as acute stress or novelty to test for their relationships with looking behaviors. For example, it is predicted that upon initial exposure to a novel environment, an animal should show more vigilant behaviors, but with increasing exposure and exploration, attention should become less vigilant, and the animal can deliberately focus attention and action on specific environmental features.

The current study

The goal of the current study was to establish the relations between infant analogs of the behavioral indices of vigilance that have shown high coherence in the animal literature. For this purpose we created an observational paradigm to capture multiple measures of infants' attention to surrounding stimuli. Specifically, we measured vigilance using four measures. Two measures captured degree of responsibility of the infant to peripherally presented target stimuli: reorientiation latencies and reorientiation likelihoods. For these, increased vigilance corresponds to shorter latencies of reorientiation and increased likelihood of reorientiation. Two additional measures captured specific features of fixations related to vigilance. These were the duration and rate of fixations during the time of stimulus presentation, where vigilance entails a high rate of short fixations. These four micro features of attention allowed us to assess the vigilance of
individual infants. By examining the within-subjects coherence of these multiple independent measures of vigilance, we tested whether infant attentional behaviors were consistent with the AJMA model. Additionally, we wanted to assess the relations between these micro measures of vigilance and the two counterintuitive claims of AJMA: the individual differences in attentional tradeoffs and the adaptation of attention over time.

Capturing these four features of looking required a number of specific changes from the typical experimental paradigms used to study infant attention. Typical paradigms keep adults out of the infant’s line of gaze (but see Benasich & Tallal, 1996) and artificially constrain the infant’s direction of attention by selective lighting of a single central stimulus. Instead, our experiment was performed in a well-lit room with six monitors placed around the infant. The video monitors were situated all around the infant, so they were not all visible from any given viewing angle. The monitors would turn on and off in a quasi-randomized sequence to play a colorful video-clip with music. Such stimuli are known to be highly salient to infants (Phillips-Silver & Trainor, 2005; Teller, Civan & Bronson-Castain, 2005). In order to localize the sound and fixate on the video, the infant would have to redirect attention to the current target monitor. An experimenter was seated (facing the infant). This contributed to ecological validity in that an infant would seldom be left alone in an unfamiliar setting. Otherwise, the room was empty and painted with neutral colors. This setup allowed us to observe infants as they attended to a naturalistic scene where sources of salience were spatially distributed and multimodal, as well as dynamically changing.

As the monitors turned on, infant responses to the target stimuli provided our measures of responsivity. The design included six trials, which allowed us to study adaptation of infant looking over time: as different monitors played, infants could vary in their persistence of vigilance. AJMA suggests acute regulation of the LC/NE system; thus we hypothesized decreasing vigilance to the decreasing novelty across trials. Additionally, we measured percent of time spent looking to low salience stimuli (such as floor or walls) during target monitor presentation, as an additional potential correlate of low vigilance states.

Methods

Participants

Sixteen 6- to 7-month-old infants (11 boys, 5 girls; mean age in days = 205, range 174–223 days) were recruited to participate. Most of the infants (N = 12) had participated in an experiment in our lab one month earlier. Infants were recruited from local mother-infant social groups such as playgroups and exercise classes. Average age of parents was 35.5 years (range 26–43) and average education was 4.7 years post high school (range 2–7). Three additional infants could not be coded due to equipment error. Infants were recruited and tested using procedures approved by the Human Research Participants Protection committee (UCSD).

Materials and setting

The testing room was fitted with five unobtrusive cameras (one in each corner and a fisheye lens overhead) and a microphone for auditory data collection. It was also outfitted with six 30 cm flat-screen video monitors, each with stereo speakers, placed in specific locations around the room: three to the left of the infant (to his/her front, periphery and back) and three to his/her right, in similar locations. Infants sat facing an experimenter on their caregiver’s lap in a seat designed to allow a full range of torso rotation. Caregivers wore shaded glasses and sound-isolating headphones to ensure that they could not systematically influence their infant’s responses.

All clips were drawn from the series ‘Baby Einstein’. Each depicted colorful moving toys or animals that were found in a pilot study to be highly and approximately equally interesting to infants. Synthesized classical music clips from the videos were edited to play at a uniform volume, and were synchronized with the onset of the video. The six video-and-sound stimuli were identical across subjects. Prior to participating in the task, infants were familiarized with the lab setting for approximately 10 minutes, and then completed another brief social-interaction task.

Procedure

At each trial, one of the six monitors (i.e. ‘target’) began playing an 8 second musical video clip. The five remaining monitors displayed a static image the color of the surrounding walls meant to neutralize their effects. For reasons unrelated to our questions, 2 seconds after the video onset the experimenter pointed and looked to the target monitor and said [Infant’s name], look! The pointing was held until the end of the clip, after which her arms returned to the center, and she looked back to the infant. Following the offset of each clip was a 2 second inter-trial interval, followed by the next trial, for a total of six trials. Between clips the experimenter smiled at the infant and said ‘Hi baby!’ but otherwise did not physically or verbally engage the infant. Clip order was quasi-randomized (with the constraint that neither side nor latitude was repeated across successive trials).

Every infant received the same order.

Coding

Coding was completed by a single coder who was unaware of the hypotheses. For calculating reliability statistics, the first author additionally coded 20% of the videos.
Sound coding

An audio spectral analysis of video was performed using Audacity sound editing software (http://audacity.sourceforge.net/) to find the onset of each video clip and the experimenter's verbal prompt.

Behavioral coding

The coder coded frame-by-frame (30 Hz) for each trial, for the duration of each video clip (i.e. 8 seconds per trial over six trials for 48 seconds total per infant). The following behavioral variables were coded:

(a) **Reorientation onset (RO)** was the first frame of the infant's saccade or head turn in the direction of the target monitor, in response to the music or video. This did not require that the next fixation was to the target, but merely an initiation of a reactive shift in the correct direction. Motions in another direction or plane (e.g. vertical shifts) were not counted.

(b) **Looking region** was a continuous, mutually exhaustive index of the area of the room where the infant was looking. Location categories were the experimenter's head/torso, experimenter's arm/hand (if extended; e.g. pointing), down (floor; area beneath the seat; wall below monitors), non-target monitors (regions around any of the five monitors not playing the video in any given trial) and target monitor. Thus, the areas designated target monitor and non-target monitor changed in every trial. k = .81.

(c) **Fine-grained fixations.** When infants' eyes were visible in the camera angles, fixations were recorded. Fixations were defined as an infant's eyes remaining static for at least 230 ms (seven frames at 30 fps). Our four-camera system provided appropriate views of the infant's eyes from the four facings. Through observing the quad view of these four cameras, we were able to determine when the infant was making small saccades versus remaining fixated on a location. Due to the dynamic quality of the videos, it was difficult to determine whether fine-grained saccades on a target monitor were examining image details, or tracking moving objects in the video. Thus, this variable was not coded when the looking region was the target monitor. Across trials, the two coders were correlated at r = .85 for number of fixations, and r = .82 for average duration of fixations.

Measures of vigilance

From these coded behaviors we derived several variables that parallel those in the animal literature described above.

(a) **Latency to Reorientation Onset (LRO)** was the latency of the infant's first saccade or head turn in the direction of the target monitor. LRO was calculated by subtracting the RO from the onset of the target video, via the onset of sound. For analyses, we used the median LRO over the six trials.

(b) **Target Hits (TH)** was the percentage of trials in which an infant fixated upon the target monitor.

(c) **Fixation Duration** was the average duration of all fixations to regions other than the target monitor, coded over all six trials.

(d) **Fixation Rate** was the ratio of the count of coded fixations in regions other than the target monitor to the total amount of time spent looking at all of these areas. This normalizes the count across infants such that it is independent of looking durations.

The four measures of vigilance are independent. However, there are potential dependencies between the looking time measures. Looking times are analyzed as a percentage of the trials, and thus at extreme levels could become dependent. Importantly though, unless proportion looks to any one location approach ceiling or floor, there can be a great deal of independence between them. For example, an infant who has a short LRO or high THs can spend a lot of time looking at the experimenter if upon gazng to monitor, she quickly returns to gaze to the experimenter. We therefore examined individual infants as well as sample means for ceiling/floor effects. None were found, indicating that all measures vary independently.

### Results

Descriptive statistics for the four measures of vigilance showed considerable variability among infants (Table 1),

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Average (and SD) of four independent theoretical measures of vigilance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LRO</td>
</tr>
<tr>
<td></td>
<td>(SD)</td>
</tr>
<tr>
<td>Average</td>
<td>1.27 sec</td>
</tr>
<tr>
<td>Range</td>
<td>.46-3.5</td>
</tr>
</tbody>
</table>

Note: LRO = Latency to Reorientation Onset.

---

1. During analyses we examined distributions for average LRO scores and discovered that for many of the subjects a single trial was at least 1.5 SD above the average. To minimize this skewing effect of these trials we used the median latency to initial motion scores for this and all following analyses.

2. Note that our definition of fixation follows that of the animal literature rather than the infant literature. Infant literature often uses the words 'look' and 'fixation' interchangeably to refer to uninterrupted gaze towards a large on-screen stimulus. However, such 'looks' are actually composed of many individual fixations joined by saccades. This distinction is relevant because, as noted above, it is the duration and rate of these fine-grained fixations that have been identified as indices of vigilance in animal samples.
thus allowing for the examination of coherence among measures of vigilance.

**Primary analysis: relations among ‘measures of vigilance’**

To investigate whether individual infants would show coherent patterns of behavioral vigilance we calculated correlations between the four proposed measures (Fixation Duration and LRO scores were reverse-coded so that higher scores indicated higher vigilance). All correlations were in the expected direction, and there were positive correlations between all pairs of measures (see Table 2).

**Vigilance Index**

Given the high correlations among the four predicted measures of attentional vigilance, we created a summary Vigilance Index. For this we averaged the four standardized values (using Z-scores reverse-coded where appropriate). The relatively high degree of correlation between this index and each measure (Table 2) validates that a summary index is appropriate; likewise, the fact that each pair of measures is not perfectly correlated suggests that there is added value in creating such an index. This table also shows that the correlation between each of the individual measures and the Vigilance Index was high. Figure 1 shows standardized scores for each of the four measures of vigilance and the summary index for each infant.

**Additional measures: looking distribution; attention to experimenter; adaptation across multiple trials**

We found interesting correlations between the Vigilance Index and our continuous and mutually exhaustive measures of looking time (Table 3). These measures were summed over all trials (i.e. 48 seconds total) and then calculated as a percentage of total time. To summarize, vigilance was strongly negatively correlated with time spent looking at the experimenter, looking down, and moderately negatively correlated with responding to verbalizations. A large positive correlation was found between the Vigilance Index and Target Monitor Looking time. Non-target monitor looking (i.e. time looking around the room at other monitors) was moderately correlated with the Vigilance Index.

We also investigated vigilance to the experimenter’s cue as an index of sensitivity to social events, specifically verbal and gestural cues. As described above, 2 seconds into the trial, the experimenter extended her arm towards the target monitor and said ‘Infant’s name, look!’ If infants reoriented to the experimenter’s hand or face within 2 seconds of the onset of the verbal cue they were coded as ‘responding to verbalization’. The percentage of trials in which this occurred was negatively correlated with the Vigilance Index (r = -0.44). Thus, more vigilant infants were less likely to attend to social cues.

In order to test for group differences in changes in vigilance measures over time (i.e. adaptation of the vigilant attentional profile) we divided the infants into Low and High Vigilance groups, using a median split of the summary index scores. This method provides a less sensitive analysis than correlations shown above; however, previous studies have used this technique to identify meaningful subgroups for individual difference analyses (Frick et al., 1999).

---

**Table 2** Correlation coefficients among four theoretical measures of vigilance

<table>
<thead>
<tr>
<th>Variable</th>
<th>LRO</th>
<th>Target Hits</th>
<th>RC Fix Duration</th>
<th>Fixation Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target Hits</td>
<td>0.75**</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fixation Duration</td>
<td>0.43**</td>
<td>0.81**</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Fixation Rate</td>
<td>0.58**</td>
<td>0.56</td>
<td>0.26</td>
<td>1.00</td>
</tr>
<tr>
<td>Vigilance Index</td>
<td>0.57**</td>
<td>0.93**</td>
<td>0.77**</td>
<td>0.64**</td>
</tr>
</tbody>
</table>

Note: LRO = Latency to Reorientation Onset; RC = Reversal-coded; Fix = Fixation.

*p < .10; **p < .05; ***p < .01.

**Figure 1** Subjects are arranged along the x-axis in rank order of vigilance index, from left to right. For each subject, the ‘column’ includes the vigilance summary index, and each specific vigilance measure.

**Table 3** Correlation results: Vigilance Index and proportions of looking time to different regions of the room, and responses to experimenter’s verbalizations

<table>
<thead>
<tr>
<th>Variable Controlled (%)</th>
<th>Vigilance Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Experimenter Looking Time</td>
<td>-0.76**</td>
</tr>
<tr>
<td>% Non-target Monitor Looking Time</td>
<td>0.45*</td>
</tr>
<tr>
<td>% Down Looking Time</td>
<td>-0.64**</td>
</tr>
<tr>
<td>% Target Monitor Looking Time</td>
<td>0.86**</td>
</tr>
<tr>
<td>% Response to verbalization</td>
<td>-0.44*</td>
</tr>
</tbody>
</table>

*p < .10; **p < .05; ***p < .01.
On average, across trials less-vigilant infants showed a greater decline in attentiveness to targets, or faster habituation, to the trial structure itself. We compared early trials (1–2) to late trials (5–6).1 Though high vigilance infants spent more time looking to target monitors than their low vigilance peers, all but two infants (in the baseline group) spent the majority of their early trials looking to the target monitors (majority was considered greater than 50%). By contrast, during late trials, only one of eight low vigilance infants met this criterion, while six of eight high vigilance infants did. A t-test shows that this pattern of change is highly significant (p < .01). Thus while high vigilance infants maintained high target attentiveness consistently through the end of the session, low vigilance infants reduced their target looking as the trial progressed.

Discussion

There have been calls in the literature to broaden our metrics of looking behavior beyond the standard measure of duration (Astin, 2007; Kagan, 2008). Our study is the first to code fine-grained measures of gaze distribution to multiple targets that are distributed broadly around a room. Multiple synchronized cameras were used to capture video of individual gaze fixations within longer periods of shifting attention to dynamic ecological events. Trained human coders took standard (NTSC, 30 fps) video recordings and, using commercially available coding software, achieved high reliability in capturing some elements of individual fixations. This is notable because current eye-tracking systems cannot readily capture fixations from infants who are broadly scanning natural environments.

We were able to characterize infants in our sample on a continuum of vigilance, based on a set of heterogeneous measures derived from the non-human animal and adult human attention literature. The work of Aston-Jones and colleagues shows coherence among features of visual attention, specifically a high rate of fixations, short duration fixations, frequent reorientation to peripheral stimuli, and short latencies to respond to peripheral stimuli. In single cell recording studies as well as pharmacological and behavioral experiments, these behaviors have been shown to tightly correspond to the release of norepinephrine from the locus coeruleus (Aston-Jones et al., 1999; Skosnik, 2000).

We know of only one study that has used a fine-grained measure of fixation as it might relate to other indices of looking behavior in an infant sample. Bronson (1991) found that shorter fixation duration of 3-month-olds was related to broader scanning of a photograph. This is consistent with AJMA, but as with the studies reviewed above, it documents a relation between only two of the indices of vigilance. We analyzed both rate and duration of fixations, as well as two measures that had already been used in the infant literature, reorientation latency and likelihood. Overall, we found correlations between a higher rate and shorter duration of fixations, short response latencies, and high number of reorientations to peripheral targets. Most previous research on infant attention has shown relations between summed or maximum looking time and a single index of vigilance. By showing the within-subjects coherence of multiple independent measures of vigilance, our work extends the support to the validity of AJMA framework for an infant sample.

Although there was overall high coherence among measures of vigilance, our measure, rate of fixations, was slightly less (though still reliably) correlated with the others. A more careful examination of the data suggests that this indicates a complex relation between fixation rate and overall vigilance, at least in this paradigm. Specifically, there was greater coherence in infants with higher overall vigilance than those with lower overall vigilance. One explanation is that at low levels of vigilance, other endogenous factors (e.g. the infant’s interest) dominate rates of fixation, but at high levels of vigilance, fixation rates are increasingly determined by a more centralized factor that modulates several behavioral manifestations of vigilance. This hypothesis bears future study.

Additionally, when we investigated the relationships between micro-behavioral measures of vigilance and looking time behaviors that captured attentional tradeoffs and adaptation in looking behavior over the course of the session, we found many strong relationships in the directions predicted by AJMA. These are particularly interesting as they are not predicted by traditional IP models of attention. We discuss these findings in more detail below.

Additional relationships to vigilance

Attentional tradeoffs

AJMA predicts that infants who are more responsive to their surroundings will spend proportionally less time looking at less salient elements of their environments, and more time looking to salient but peripheral locations. To test this prediction, we compared the amount of time spent looking to elements of the room with varying levels of salience.

We designed our paradigm to create a situation of maximal attentional tradeoffs between video clips (high salience stimuli) and other more neutral stimuli in the room. Specifically, the category ‘looking down’
comprised visually accessible but low salience areas (e.g. floors, walls, chair, experimenter’s lap, or infants’ own toes). The contrast between these areas was confirmed in the current data: in early trials, when novelty and uncertainty were highest, infants spent 70% of trial time (but not inter-trial time) looking at target monitors. By contrast, infants spent only 7% of total trial time looking down. Consistent with AJMA, individual proportion of time spent looking to available but low salience areas was tightly linked with overall vigilance: the correlation between the vigilance index and looking down was \( r = -64 \). In other words, infants who were less vigilant were less driven by unpredictable peripheral stimuli, and thus could direct their attention to less salient stimuli.

This potential to ‘self-regulate’ attention, or the ability to systematically focus on more or less salient events in a top-down manner, is critical for cognitive and affective control in a dynamic environment. For example, although we did not specifically study it in our study, children who use focused or exploratory toe-looking was common in ‘down’ looking. Such exploratory activity is a potential benefit of focused, less reactive attention. An increase in such ‘willful elaboration’ is thought to contribute to the increase in looking times to complex stimuli (such as video or objects that can be manipulated by the infant) shown by older infants (e.g. see Courage et al., 2006; Ruff & Saltarelli, 1993).

Gaze to experimenter

We also found that less vigilant infants spent relatively more time looking towards the experimenter and were more likely to look at the experimenter within 2 seconds of the social ‘pointing’ cue. At first glance this seems to suggest an alternative explanation for our data: perhaps infants whom we described as ‘more vigilant’ were in fact less sociable, and this trait (Buss & Plomin, 1984), rather than vigilance, affected the individual differences in looking time. That is, infants who are less interested in the experimenter might instead look around the environment.

However, there are reasons to question this alternative. First, it is not independent. Decreased sociability may co-occur with LC/NE-mediated vigilance as a behavioral manifestation of a mildly stressed infant. Activation of the LC/NE pathway is a major aspect of the sympathetic nervous system response, and gaze aversion co-occurs with sympathetic activation in distressed infants (Gunnaar, Larson, Hertzgerd, Harris & Brodersen, 1992; Halsey & Stansbury, 2003). Potentially stressful aspects of our paradigm could have been the experimenter’s non-response to the infant (akin to the ‘still-face’ paradigm; Tronick, Als, Adamson, Wise & Brazelton, 1978) or the sudden onset of videos. Because reactive infants would show both gaze aversion (Stansbury & Gunner, 1994) and vigilance, sociability differences are not an independent alternative explanation. In ongoing research we are testing this hypothesis using physiological indices of sympathetic activation related to stress and attention.

Another reason why looking pattern results cannot be explained by individual differences in sociability is that our measures are not independent (see methods section above). Thus, the construct of vigilance explains our results more comprehensively and parsimoniously than the construct of sociability.

Nonetheless, as explained, we expect sociability to be altered by vigilance. A potential implication is that highly responsive, stimulus-driven infants may miss perceptually subtle – but important – social cues such as gestures and gaze direction. However, in more naturalistic circumstances, adults might overcome this inattention by increasing the salience of their bids for an infant’s attention (e.g. Brand, Baldwin & Ashburn, 2002). Another implication is that while highly vigilant children might attend to and learn more from unpredictable, perceptually salient events, they will have difficulty maintaining focused attention to less salient events (e.g. a homework assignment or a teacher’s instructions). This implies a neural mechanism that contributes to longitudinal school performance outcomes, not only in individual children but also across ethnic groups and socioeconomic strata (SES). Such outcome differences are also correlated with stress. Many studies show that families of color and/or low SES experience higher levels of stress, and this contributes to infant stress reactivity (Brunner, 1997; Fonagy, 1996; Williams, Yu, Jackson & Anderson, 1997).

Adaptation over session

Vigilance is also adjusted according to changes in uncertainty and stress. We can investigate this by comparing early to later trials of the session, where repetition may decrease the novelty of the target monitor events. We found that less vigilant infants showed a decline in attention to targets across trials, which appears to be a habituation to the trial structure itself. By contrast, most of the vigilant infants continued to spend a high proportion of time looking at the target monitor in later trials. This result directly contradicts the traditional information processing account, which suggests that fast habituating infants should be consistently fast. In addition, the IP account would make the prediction that fast habituators will show other signs of processing speed, such as broad scope of attention and faster reorientation. This prediction was not confirmed; in fact, the opposite pattern was strongly observed.

However, these findings can be explained under AJMA. Specifically, vigilant infants might have been experiencing a mild stress response that caused high responsiveness to exogenous events, and failure to habituate to the occurrence of those events. A stress response entails the activation of the HPA-axis, which can cause prolonged LC/NE-mediated vigilance in
consistently stressful environments (e.g. as a strategy for identifying lurking predators or for reacting to threatening same-species competitors; Arnsten, 1998). By contrast, less vigilant infants showed initial responsibility to the monitors, but after repeated trials could habituate to periodic video onsets. This is a more typical pattern of response to uncertainty where a pattern can be learned, and uncertainty is thereby reduced (Yu & Dayan, 2003).

Thus, within the typical context of infant habituation studies, where the occurrence of semi-predictable, normally distributed (e.g. peripheral) stimuli has been artificially removed, we ask: Are short lookers smart, or are they stressed out? In ongoing tests in our lab we are testing the relationships between arousal, HPA-axis activation and attention (Zavala, Overton, Chiba, de Barbaro, Khandrika & Deák, in preparation).

Conclusions

Our analysis of the indices and correlates of vigilance in infants in a semi-naturalistic paradigm supports and generalizes Aston-Jones et al.'s (1999) attentional theory. By defining indices of vigilance derived from non-human animal studies, we found that human infants show the same patterns of co-occurring indices. This is one of the first human studies, and the first infant study, to investigate micro-indices of vigilance predicted by AJMA. The relations among vigilance indices cannot be explained by conventional IP accounts of infant looking.

Additionally, the AJMA model provides a framework to integrate universal behaviors ignored in IP theory, including social attentiveness, individual differences in temperament, stress reactivity, and, most generally, distribution of attention in natural (e.g. lit, dynamic, cluttered) environments. We do not argue that speed of processing is a non-factor in infant looking: individual differences are well documented, robust, and relevant to many cognitive skills (Kail & Salthouse, 1994). Generalized processing speed must eventually be integrated with a comprehensive neuro-modulatory model.

However, processing speed cannot itself explain many findings. AJMA theory allows for explanations of attentional dynamics on both short (e.g. second-to-second) and longer timescales, socio-emotional influences on attention, and the adaptive importance of slow/sustained attention as well as fast looking. Importantly, it may also help to elucidate mechanisms of stress-related social factors not only in infants but in older children's school outcomes (Lee & Burkam, 2002). By demonstrating the viability of AJMA theory to a developmental audience, we hope to encourage a dialogue about how the ecology and neuroscience of infant attention can explain universal patterns of infant cognition and attention.

Acknowledgements

This project was funded by a grant from the National Science Foundation (NSF SRS-0527756) to G. Deák and by NSF grant #SBE-0542013 to the Temporal Dynamics of Learning Center, an NSF Science of Learning Center. We thank all of the families who participated, as well as Alicia Zuniga, Melissa Karnaze, and Ana Ramundo for assistance in collecting and coding data. We also thank Natalie Wong and Marybel Robledo for data collection and management and training of students.

References


© 2011 Blackwell Publishing Ltd.


© 2011 Blackwell Publishing Ltd.


Received: 14 June 2010
Accepted: 8 March 2011
Chapter 1, in full, is a reprint of the material as it appears in Developmental Science 2011. De Barbaro, Kaya; Chiba, Andrea; Deák, Gedeon O., 2011. The dissertation author was the primary investigator and author of this paper.
CHAPTER 2: TWELVE-MONTH “SOCIAL REVOLUTION” EMERGES FROM MOTHER-INFANT SENSORY-MOTOR COORDINATION: A LONGITUDINAL INVESTIGATION

de Barbaro, K., Johnson, C.M, & Deák, G.

Introduction

The period around an infant’s first birthday marks a qualitative change in how the infant responds to and participates in the activity of adults. While, from much earlier months, infants will share gaze with a partner, or gaze at objects manipulated by a partner (i.e. dyadic states of mother-infant or infant-object play), around 12 months they begin more actively to engage in shared actions on objects, such as imitation and games (Bakeman & Adamson, 1984; Hay, 1979; Piaget, 1962; Ratner & Bruner, 1978; Stern, 1985; Trevarthen & Hubley, 1978). These activities are “triadic” in the sense that they all involve shared attention and activity between the infant, an adult, and an external locus of attention, such as an object or an event.

Previous research has shown that triadic attention is a foundation for later learning, including language development (Bruner & Watson, 1983) social skills (Bornstein & Tamis LeMonda, 1989), and cultural learning (Tomasello, Carpenter, Call, Behne, & Moll, 2005). However, we know less about how triadic attention develops. Specifically, previous methods have puzzled researchers eager for a
coherent developmental account of the processes that bridge to the dramatic “triadic shift” at 12 months (L. Adamson & Bakeman, 1991, p. 34; A Fogel & DeKoeyer-Laros, 2007). We aim, in this longitudinal study of naturalistic mother-infant interaction, to demonstrate how a novel approach to studying interaction can make headway on this putatively intractable developmental question.

Our approach comes out of recent theoretical and empirical work in cognitive science that takes as its premise that cognition is fundamentally embodied and distributed (Hutchins, 1995), and can be observed through the systematic micro-analysis of multi-modal, multi-party interaction. We will first give an overview of methods traditionally used to study triadic attention, and consider how they have shaped researchers’ interpretation of its development. Next, we review micro studies of infant attention, followed by an overview and motivation of an embodied and distributed cognition approach. Finally, we will describe the results of our application of a micro-analysis of multiple attentional modalities in mother-infant interactions, and detail the implications of this approach for the distributed and embodied cognition that develops.

The triadic “shift”: Previous accounts

There is a broad consensus that mother/infant face-to-face interaction follows a three part trajectory (L. Adamson & Bakeman, 1991; Tomasello et al., 2005). First, from about two months, infants begin to be able to engage in dyadic
states of “shared attention” with their caregivers. Here, shared attention is construed as jointly attending to one another’s faces via gaze. Infants both initiate and respond to various facial expressions their caregivers make with increasing amounts of positive engagement. This has been observed both in observational and experimental studies. For example, when mothers are asked to stop responding to their infants by “freezing” their facial expressions, even very young infants will quickly become less positive and animated (E. Tronick, Als, Adamson, Wise, & Brazelton, 1978) and will make active bids to reengage the mother (E. Z. Tronick, Ricks, & Cohn, 1982). Next, starting at around six months, infants show a marked decrease in gaze and positive affect to their mother’s face (Kaye & Fogel, 1980; Trevarthen & Hubley, 1978), and become much more attentive to the objects in front of them (Bakeman & Adamson, 1984; Bruner & Watson, 1983). Thus again it is a dyadic state which predominates the interaction, this time involving the infant and an object of interest. At this stage, infants will occasionally gaze up to the parent while attending to objects. However, they do not make systematic efforts to involve their partners in their object exploration until near the end of their first year (Bakeman & Adamson, 1984).

Summing over years of research, Adamson & Bakeman (1991) describe a “curious developmental gap” (Adamson & Bakeman, 1991b, p21) between the dyadic states described above and the appearance of true “triadic” play around the infants’ first birthday. Given that what they consider to be the two components of triadic
play—that is, infant-mother interpersonal play and infant-object play—each predominate in earlier periods, it is not clear why infants do not readily or smoothly begin engaging in triadic or joint object play with caregivers. One sort of explanation for the late emergence of triadic play emphasizes a qualitative leap, attributable to new conceptual and inferential resources. The explanation stipulates a dawning awareness that other people have “other minds” (Stern, 1985; Tomasello et al., 2005; Trevarthen & Hubley, 1978). One claim of the work we will be reporting here is that the “gap”, or discontinuity in the behavioral data, almost forces previous researchers to invoke an invisible representational shift to account for the development of triadic attention. Further, we suggest that the behavioral discontinuity is based in past researchers’ choice of units of analyses, and how the resulting component parts that they consider limits their access to the processes involved in development.

Our embodied distributed cognitive account of a longitudinal sample of five infants at four, six, nine and 12 months suggests that these new complex action and exploration patterns can emerge without the need for a conceptual “sea change.”

By tracking each of the partners’ access to others’ activities in the world: their motions, words, affect, through space and time, we were able identify a number of distinct action trajectories wherein activity builds off of that seen at earlier sessions. This methodology is a boon to a developmental account in that it focuses on the changing processes of cognition that are visible rather than invisible. In doing so, it
provides a basis with which to compare moments of interaction longitudinally across a developing dyad, prior to and including the infants’ first birthday (Johnson 2001; Johnson, in press, Forster, 2002; Forster & Rodriguez, 2006). Our analyses lead us to argue that 12-month complexity is not sudden or discontinuous but a culmination of continuous changes across the first year.

Before detailing our approach and our account of triadic development, we first review relevant past research into two categories: macro measures and micro measures.

**Macro Analyses**

The majority of previous studies view the development of triadic attention in terms of “macro” level changes in mother/infant face-to-face interactions. These can be considered macro both in the timescale at which phenomena of interest are tracked, as well as in the unit of analysis. For example, Fogel (Fogel & Hsu, 2003) and Bakeman and Adamson (1984), in historically detailed studies of developing mother infant interactions, annotated their video with a single dimension: that is, they used a single “layer” of mutually exclusive state-variables to classify relatively large time-units within the interactions, at the level of the dyad. For example, in their classic longitudinal study, Bakeman and Adamson (1984) use a one dimensional coding scheme that distinguishes between the following six states, each referring to the focus of an infants’ attention during face-to-face interaction with mom: Unengaged,
Onlooking, Persons, Objects, Passive Joint, and Coordinated Joint. Similarly, Fogel’s “Relational System” (e.g., Hsu & Fogel, 2003) codes at the level of the attentional coordination of dyad. From that approach, Fogel notes, for example, whether the attention the dyad displays is Symmetrical (both attending to the same thing “actively”), Asymmetrical (one partner viewing the other actively attending to an object, as when a mother might gaze to her infant as he manipulates a toy), or Unilateral (one partner engaged in an activity and the other attempting to engage them in a second activity) relative to one another.

While these studies reveal systematic changes in attention across development, they do not indicate how such changes occur (for discussion, see also A Fogel & DeKoeeyer-Laros, 2007; Forster & Rodriguez, 2006; Johnson, 2001).

Macro studies indicate the high level products of the interaction: what the dyad accomplishes during each episode of interaction. The categories are qualitatively different from one another, such that we simply don’t see anything similar to triadic interactions at earlier months. This leads to behavioral discontinuities between dyads with infants of different ages. In turn, this behavioral discontinuity provides little traction for explaining the age-related shift from one state to another.

Distributed and embodied cognition provide an alternative. By instead focusing us in on the interaction as a rich temporal configuration of component parts,
we can observe continuous progress in measures that span the entire age range. By providing continuity across behaviors observed across the first year, we no longer need a discontinuous representational shift to explain the development of triadic attention.

**Micro-analysis**

A number of more recent empirical studies track attention at the “micro” level, specifying the particulars of how the mother and the infant attend to one another and available toys. Micro can be differentiated from macro studies in a number of dimensions, including the timing, specificity and comprehensiveness with which they code attentional behavior. Generally, timing refers to the units of time at which changes in attention are identified. By specificity, we refer to the precision with which attentional behavior is indicated via a particular sensorimotor modality, such as gaze. Macro studies often code a high level construct of attention akin to the traditional psychological definition of attention as a unitary amodal “spotlight.” This renders attention as an invisible, internal process. Instead, micro studies often specify the particular sensory-motor modality that comes into contact with the targets of attention. Finally, the majority of micro studies of infant activity specify gaze as the single sensory-motor modality by which their subjects attend. However, the utility of coding a wider range of modalities (e.g. hand, mouth, & gaze) and their targets in fine detail is becoming evident. Comprehensiveness refers to the degree to
which studies code a variety of sensory-motor behaviors. Below, we review micro studies of attention relevant for our study. Overall, we argue that we need high resolution in timing, specificity and the comprehensiveness of modalities in order to capture the development of sharing actions between mother and infant.

Many studies have shown the critical importance of tracking at the millisecond level when accounting for changes across longitudinal time. The majority of the macro-level studies described above code at relatively large temporal scales of once every second or even once every three seconds (Bakeman & Adamson, 1984; Cohn & Tronick, 1987). However, attention to social information can occur at the much finer timescale of tens of milliseconds. This is relevant, for example, in that recent evidence from Yoshida and Smith (2008) suggest that older infants’ gaze to caregivers might shift from longer periods of fixation to shorter but more frequent “checking-in” fixations. More generally, Deák, Krasno, Triesch, Lewis, and Sepeda (in review) found (using a clearer dyadic coding scheme) that dyadic attention states between caregivers and infants from 3 to 11 months of age changed an average of 31.7 times per minute. This suggests that coding intervals of any period longer than 1 sec would certainly miss important changes in social attention; even with the largest acceptable interval of 1 sec, many other events will be missed. However, many previous efforts to code attention-sharing and triadic attention used coding unit durations of several seconds. Thus, developmental changes of attention-distribution
within the ongoing social interaction were simply not captured by many past coding schemes.

A number of studies have started to specify moment to moment changes in the targets of gaze during social interactions. Detailing attention in this way has led to the finding that infants spend relatively little time looking to the face of their caregivers during their interaction with objects. Instead, they spend the majority (up to 80%) looking to hands – either their own hands or their caregivers’ hands, while those hands are in some sort of contact with the objects (Fiser, Aslin, Lathrop, Rothkopf, & Markant, 2006; Krasno, 2007; Yoshida & Smith, 2008). This holds true for infants of a variety of ages and in naturalistic settings with many potential looking targets (Deák, in press).

By combining high resolution gaze coding with fine grained temporal analyses, Deák et al (in press) identified that it is the motion of mom’s hand on a toy–rather than the gaze of the parent – that best predicts infant gaze shifts from one location to another (Deák et al., submitted; Yoshida & Smith, 2008). By precisely coding specific attentional behaviors as they unfold in real time these studies have identified patterns that contradict the conventional wisdom on early infant attention.

However, we know of no studies that detail infant attention in a dyadic context in a comprehensive manner, coding the full range of sensorimotor behaviors by which infant attends to its surroundings. While gaze is certainly the most
commonly identified modality with attention, developmentalists have long commented on the attending functions of other sensory motor modalities. Eleanor Gibson (1988), for example, has written extensively about the importance of the hands to infant sensory exploration: infants can, for example, rotate an object to view different angles, or squeeze it to receive the sensations about its density and internal makeup (see also Streri & Feron, 2005). Additionally, they can use hands to bring an object to the mouth, another key modality in the first year (Rochat, 1989). Manipulation has motor functions but also has attending functions.

A large body of work from the lab of Holly Ruff has determined that multimodal attention to objects has different physiological and cognitive consequences from simply gazing to an object. Ruff and her colleagues differentiate between “casual” attention which involves gaze-only, and “exploratory” attention, which involves concentrated gaze, coordinated with certain types of haptic manipulation such as slow rotating or fingerling an object (for a review, see Ruff & Saltarelli, 1993). In comparison to gaze-only attention, exploratory attention is associated with physiological indices of increased focus, such as heart-rate deceleration (Lansink & Richards, 1997). It also diminishes with object familiarity, decreases the likelihood of distraction (Oakes & Tellinghuisen, 1994), and inversely predicts future distractibility (Lawson & Ruff, 2004). Thus, we anticipate that
specifying between these will be important for characterizing attention development across the first year.

As a final motivation for a comprehensive tracking of attentional modalities, we note that we saw important differences in the way that infants responded to the mothers’ actions on toys, but only when we were looking at the infants’ manual activity in addition to their patterns of gaze. In the study by Deák et al. described above, all infants throughout the study period (from 3-11 months) showed similar patterns of shifting gaze towards a toy following maternal manipulation (Deák et al, submitted). In our own analyses (described below) we have found that there are indeed longitudinal changes from four to 12 months in infants’ responses to maternal bids. However, these changes are in the nature of infants’ multimodal contact with the object, and how that unfolds over timescales of 5-30 seconds, rather than solely on their gaze immediately following the bid.

Embodied affect is another important dimension for characterizing the developing social attention in the dyad (L. Adamson & Bakeman, 1991). Affect is studied via general arousal levels as well as facial expressions, especially relative to similar activity in the other. That affect and attention have an important relationship in infancy is clear from previous work showing, for example, that infants use gaze aversion and negative affect to regulate over-arousal (e.g., Field, 1981). Additionally, affect has been shown to have its own developmental course in coordination with
other components of attention. For instance, by 12 months, infants have developed patterns of affect that are precisely timed with actions on an object, and associated with the gaze to mother (Eckerman, Whately, & McGehee, 1979). Additionally, affect from the mother is important for the development of attention and later learning, perhaps via the social reinforcement that the infant is receiving for particular actions (Dodici, Draper, & Peterson, 2003). In our study, we further observe that affect in the infant shifts from being attached to immediate events (the mom’s smiling face, a toy tossed in its lap) to occurring within a larger routine (at the denouement of a game, at the recognition of an imitation, at the accomplishment of - or frustration with - a task, etc).

Summarizing, the current micro literature shows the relevance of the micro analysis of affect and attention, including the targets of gaze of both participants, as well as their manual actions. However, contemporary work is still too piecemeal to account for the emergence of triadic attention at 12 months. For this we need a distributed analysis that situates analyses of embodied attention in the social interaction, as we detail below.

Our approach: Reconceptualizing interaction

Our definition of attention for this work is based on theories of embodied cognition wherein activity of the body is considered cognitive activity. Perception is not a passive process in which sensory information is displayed on the retinas.
Instead, infants move their bodies in order to identify and explore features of their environment. Furthermore, infants do not just sense wavelengths of light but rather they actively seek out differences in their surroundings via all manner of sensory receptors.

This leads us to define attention as an effortful change in the sensory access of one individual to objects or other individuals (Johnson & Karin-D Arcy, 2006). Consistent with the neurobiological organization of human sensory-motor systems, we distinguish different types of sensory access, including visual, oral, and haptic. Thus, changes in the targets of gaze, as well as in manual and oral contact with toys and partner, were all regarded as important in characterizing the attention of our participants. Furthermore, we distinguished between reaching towards a toy, grasping a toy, and manipulating a toy, since each of these motor activities differs in the type and timing of the tactile and proprioceptive access that they afford.

Above we reviewed a number of benefits to operationalizing attention in this way. Moreover, by tracking dynamics of sensory-motor modalities as they gain access to targets, attention becomes a process that unfolds over moment-to-moment time. Studying the changing process of how infants’ modalities become organized to attend to their caregivers and objects in the world around them provides a new window into the development of the infant’s attention. For example, we can characterize the dynamics of sensory-motor modalities as they become coordinated
with a toy: How many modalities are on the toy? What order do they get there? How long does gaze remain on the toy given concurrent maternal elaboration? Given concurrent infant manual elaboration? Given maternal elaboration of another toy? To give an example, our qualitative results show that gaze typically leads in infants of all ages, but that younger infants maintain gaze contact with the toy for the full duration that any modality is in contact with that toy, whereas older infants (6-9 months) may begin to look away once the hand has made contact with the toy, or even (9-12 months) look away during the reach. Thus, while reaching is visually guided at all ages, it depends on a decreasing level of gaze involvement, as the infant comes to decouple its sensory modalities (see also Bushnell, 1985). In this way, by documenting the microdynamics of how the infant organizes its sensory motor modalities over developmental time, we can observe the variety of cognitive changes involved in the emergence of triadic attention.

Drawing from distributed cognition, we embed this analysis of embodied, multimodal attending in a triadic context of the mother-infant-object. From a distributed perspective (Forster, 2002; Hutchins, 1995; Johnson, 2010), the focus of research is not just on the elements of a system, but on their configuration. A distributed account of ontogeny, then, is one of configural change. Given, for example, the multi-modal data generated by embodied analyses, we can observe a set of elements that re-configure as the infant ages. Many of these elements - such
as touch, eye contact, visually tracking a proffered object, etc. - are common to all ages. However, at each age, these elements organize relative to each other, and to new behaviors (like “reach” or “stack”), in a distinctive way. The above example of the changing relationship between gaze and reach illustrates this notion of attentional reconfiguration. By characterizing such changes in organization, the distributed approach can help provide a coherent account of the transitions involved in the development of triadic attention.

The distributed approach is well adapted not only for data that are multi-modal, but also for those that are multi-party. In the development of triadic attention, there are many critical relationships - such as eye contact, imitation, altering the other’s access to the toy, etc. - that cannot be specified by the behavior of one subject alone. For example, when a mother “presents” a toy, the criteria for scoring that event include both the mother’s grasp and extension of the object, as well as the infant’s available line of sight to that object. Taking interaction as the unit of analysis is a key characteristic of the distributed approach. From this view, the infant’s ontogenetic challenge is not to perform particular, pre-specified actions, but to adapt to the conditions embodied by the mom’s activity, which in turn are adjusted in response to the infant’s current behaviors.

Furthermore, we can describe pronounced individual differences in the mother’s tendency to act contingently with her infant - that help shape how and
when the infant engages. Thus, rather than assigning binary performance scores (e.g. whether the infant does or does not “attend” at a given age), distributed analyses produce “profiles of participation” (Forster & Rodriguez, 2006) that reflect the co-regulation of activity within a dyadic interaction. Assessed longitudinally, such analyses can reveal the developmental course of changes in mother-infant object coordination.

The study of cognition from this approach also highlights a tenet of the distributed cognition approach: cognitive events are multi-scalar - i.e., unfolding simultaneously at the micro, macro- and historic/developmental time scales (Hutchins, 1995). To produce a coherent distributed account of cognitive development, information must be collected at all these time scales. At the micro-second scale, we observe shifts of gaze, facial expression (e.g., gleeful smiles), and hand movement. At the macro-scale, a particular look or grasp is positioned within an ongoing routine, where, for example, it may repeat (as in peek-a-boo), or change, or organize with other events. At the historic/developmental time scale, the dyad’s long-term experience with such routines (e.g., a playful father’s tendency to initiate exciting games) comes into play.

In the study reported here, a monthly six-minute sample of free play was recorded from infant-mother dyads at home. For the current analyses we observed episodes from when the infant was four, six, nine and 12 months of age. By tracking
the details of these interactions across months, we can capture relevant changes at both the micro and macro levels. Of course, making inferences about these historic shifts requires interpolation and induction - for example, we do not know how often, and at what ages, a given mother-infant dyad have played, e.g., “peek-a-boo” or “build-up/knock-down” games. Nonetheless, some historic relations are directly observable, by comparing configural change across months. For example, we can describe long-term changes in the extent and nature of scaffolding by the mom by observing how her bids for the infant to attend to objects vary over developmental time. That is, the level and placement of motor activity by the mom clearly changes as the infant ages, from elaborate movements and expressions repeatedly directed at the infant’s immediate frontal space, in the earlier sessions, to only a distal object touch and glance at the infant by 12 months. Together, assessments at these different time scales enable us not only see cognition in action, but to watch it develop as well.

While a multi-modal, multi-party, multi-scalar account can become very complex very quickly, its grounding in the embodied activity provides a straightforward approach that can reveal both significant pattern shifts and unexpected continuities. At each age, we will describe the same set of variables and participatory event-types. These include the type and number of attentional modalities directed to each target, the order and rate of modality specific transitions
between targets, the contingencies between the participants’ activities, and the type and timing of accompanying affect. We can thereby document how the organization of these elements changes as triadic attention develops. Each description will also include how “maternal bids,” in which the mother directs the infant’s attention to a different object, are coordinated at each age. As we shall see, these interactions and their constituent micro-behavioral elements configure differently over time, as they become organized into routines that are increasingly prolonged, attentionally divided, and eventually embedded. Through this approach, for example, we can observe how constraints on the infant’s motor development—for example, decoupling the hands to simultaneously contact multiple objects—shape how the infant responds to maternal toy bids. In this way, we argue, an embodied developmental change (i.e., manual action behaviors) directly feeds into developmental changes in distributed, dyadic-participatory changes. This provides an illustration of how this approach—using embodied and distributed analysis to parse complex social interactions at multiple temporal and behavioral units—can address the cognitive and behavioral complexity that emerges in very elaborate social interactions, such as infants’ triadic interactions with caregivers.
Methods

Data Corpus

We selected a random sample of five mother infant pairs from the corpus created for the Modeling the Emergence of Shared Attention (MESA) project (G. Deák, PI, NSF SES-0527756). The full corpus included data from 40 mother infant pairs who were assessed twice monthly, once at home and once in the lab, from three months to nine months, and then at 12 months. For the study described here, we used video recordings of home session free play interactions occurring at four, six, nine and twelve months.

Procedure for free play interaction

At months four, six and nine, infants were placed in modified walkers in order to control for individual differences in postural stability and height of young infants. At the twelve month session, mother and infant were both seated on the floor. Floor seating was preferable at this age, since 12 month infants become fussy at being confined to the chair. Furthermore, all infants were able to sit upright at this age so it was no longer necessary to provide additional stability to control for upright posture. This was important for the younger infants, as it affects their ability to reach (A. Fogel, Messinger, Dickson, & Hsu, 1999). In all cases, the mothers were seated on the floor facing and within reach of their infants.
At each session, a set of three toys were placed between the mother and the infant. All mothers were instructed to “play as they normally would” with their infants, using the toys as they felt fit. At four through nine month sessions, two of the toys were placed in specially mounted cup holders at the sides of the walker tray. At these sessions, mothers were instructed to leave only one toy on the tray at a time, and to return the others to the cup holders. At 12 month sessions there were no cup holders and mothers were not given further instructions to constrain the number of toys in active play. Thus while there were some differences in the setup and instructions between the four through nine and the 12 month sessions, at all ages multiple toys were simultaneously accessible to both mother and infant, both visually and manually.

At each session, three cameras recorded the interaction simultaneously: one directed at each of the faces of the participants, and one positioned in order to capture a side or “contextual” view of the dyad interacting with the toys. The free play session was recorded for 6-7 minutes at each session.

**Observational Methods**

For approximately nine months, de Barbaro and Johnson spent between three and four hours each week viewing and discussing the subsample of 20 free play sessions (five dyads x four longitudinal sessions). We observed each video many times, occasionally viewing the sessions of a single dyad in longitudinal order, and
occasionally viewing all dyads at a single session, as befitted clarifying individual observations into a pattern of results. During the process of writing we frequently returned to the videos for additional verification. For the reasons described above, we watched the tapes with the explicit goal of characterizing the interactions as organizations of multimodal, multiparty components in real time, focused on the mothers’ and infants’ attentional resources as they engaged with each other and the toys. In particular, we paid close attention to the targets of attention of gaze, mouth, right hand, and left hand, of the infant and the mother, as well as the position of each of the objects as they were manipulated and transported by the participants. We also noted displays of affect of in both the mother and infant, and to some extent, their vocalizations. We were particularly concerned with identifying variability across the longitudinal sample as well as regularities that occurred within sessions. Our analyses, while qualitative, were highly detailed and thorough, and revealed multiple developmental trajectories of interest. (For a more quantitative account of these interactions, see de Barbaro, et al., forthcoming.)

Results

For each group of sessions at each age, we will first describe the pattern of sensory-motor coordination that we observed in the infants, and then describe the nature of the dyads’ interactions at that age. The latter will focus, in particular, on “maternal bids” and the infants’ responses to them.
4 Month Sensory Motor

Sensorimotor coordination in four month olds can be characterized as “convergent”. That is, all the infant’s attentional streams tended to converge on one single target at a time. Vision generally leads, followed by hands, then mouth. If only one hand made contact, it might rhythmically slap at the object, repeatedly rub all five fingers on it, or sweep it toward the body. Often, two hands working in concert, mirroring the same motions, would clasp the object. Gaze to the object co-occurred for the full duration of such contact. Often the hands then brought the object to the mouth, allowing oral attention to supplant visual. Alternatively, the infant would bend its body to make direct oral contact with an object on the tray, even if one or both hands were already touching the object. Visual fixations on targets were relatively prolonged at this age. When a change in target did occur, the transition was fairly slow, with all of the infant’s modalities shifting to the new target, with eyes leading hands, and the previous target being fully abandoned. If the attentional shift was to the mother’s face, the hands might release a previously held object.

4 Month Dyads

Mothers engaged in their most active scaffolding at this age. This commonly included grasping an object and moving it to loom in the infant’s visual field, near his/her face. The mother would often also pivot or rhythmically shift the object. While mothers showed individual differences in their tendency to engage their
infants, all mothers presented objects, and smiled and talked when they did so. The mothers were also most likely, at this age, to swoop their own faces into their infant's near visual field, soliciting and usually achieving eye contact. Although infants individually varied in how prone they were to look toward their mother’s face, when that did occur, both eye contact and joint positive affect (smiling) tended to be prolonged.

When mothers made objects loom near the infant’s face, this also brought the objects within the infant’s reach. However, at this age infants seldom extended their arms; they kept their hands relatively close to the body, although their hands were continuously active and responsive to opportunities to touch the objects. Thus maternal scaffolding at this age was critical for infant haptic contact with toys. Mothers were most likely at this age to place an object in contact with the infant's hand, or even move the infant’s hand to an object (see Zukow-Goldring & Arbib, 2007). Infants most often responded positively to such bids, immediately clasping the object and sometimes adding visual, haptic, and oral attention to it.

At all sessions, mothers instigated a high number of object switches, in which mother’s bids interrupted an ongoing bout of multi-stream attention to a previous object, which the mothers would often first remove. In response to maternal bids, four-month-old infants would readily shift all of their attentional streams to the novel object. This shift was gradual but complete, even if the infant had gazed towards or
reached for the departing object. The infants showed little of the negative affect that they would display at later ages, in response to such removals. In fact, we characterize flow of attention following the bid at this age as *Well Coordinated*, with mothers making frequent bids and the infants complying.

**6 Month Sensory Motor**

Infants at this age frequently reach, grasp, and retrieve nearby objects. Reaching was visually guided, with eyes leading the reaching hand to within grasping distance. The infant’s capacity to grasp, and manipulate an object was better developed. Once grasped, the object was often brought closer to the body in a bimanual grasp. As at four months, infants often mouthed the object, although typically by using the arms to lift the toy to the mouth rather than bending the torso. As objects were being reached for, infants often opened their mouths, as if in anticipation of oral contact. Other times infants would shake, bang, or drop the object, usually while tracking it visually (see Bourgeois, Khawar, Neal, & Lockman, 2005, for detailed examples). Finally, infants sometimes used their other hand to manipulate or explore the object that they held (and sometimes rotated) with their other hand. Not surprisingly, some of the bimanual “mirroring” and mouthing manipulation that dominated at four months still occurred.

Unlike at four months, infants were not as bound to look at whatever object they were touching. That is, the infants visually guided their reach for an object, but
thereafter could de-couple their gaze (e.g., to a novel object) while they continued to grasp the original object. The grasping hand would often be passive while the eyes and other hand haptically explored a novel object. These infants would not, however, grasp two objects simultaneously. On the relatively few occasions when they both grasped and gazed at a novel object, the original object would tend to be left behind, as the divided system “collapsed” into a new configuration of convergence on the novel target. Thus some visual and exploratory haptic attention could be directed at one target, while a grasping hand maintained contact with another, but there was a strong tendency for such divided attention to shift, such that all the attentional streams returned to the object that was being, or had been, grasped.

Most strikingly, six-month-old infants more consistently tracked and maintained prolonged attention to a target. Even after haptically exploring a novel object, infants tended to return their gaze to the (held) previous object. When previously attended objects were removed by the mother, the infants often tracked the object during removal, and redirected hands (and mouth) to it once relocated.

6 Month Dyads

Infants initiated their own handling of objects, even if their mother had redirected their attention. This impacted the dyad’s social coordination. As at four months, gaze was the first modality to shift to an object that the mother moved in the infant’s frontal field (see Deák, et al, in review). However, infants were less likely
to then reach for that target. Rather, they looked back at their own still-grasped object. As a result, infants less often converged their attention on a novel object presented by the mother. Also, overall, there was less mutual co-gaze between the partners at this age. Finally, infants occasionally showed negative affect when their mothers attempted to remove an object of their attention. Negative affect was expressed as frowning, or by rearing back its body or gaze (Field, 1981).

Mothers did less active scaffolding at this age, perhaps, in response to the infants less often accepting new objects. Mothers also abandoned their bids at this age, and followed infants’ attention back to their current object of attention. Nonetheless, maternal bids sometimes succeeded, typically if the mother persisted, and especially if she managed to remove the original object. In either case, we classified the dyad’s coordination as Disrupted. That is, maternal bids tended to disrupt the infants' ongoing, directed activity, and attentional resources were thus actively and variably negotiated, rather than all smoothly following the mother's lead as they had at four months. This is not an evaluative label—indeed it appears to indicate a more mature strategy by the infant—but rather describes that the flow of attention between mother and infant is no longer smooth as it was at four months.

9 Month Sensory Motor

Sensorimotor coordination at nine months was marked by rapid, fluid transitions, reiterated routines, and an increased tendency to handle two objects at
once. Compared to earlier months, the progression for gaze to reach to grasp to manipulate was fluid and facile. Unlike at six months, infants could divert their attention - including grasp and manipulate - to a novel object, but quickly return full attention to the prior object. Furthermore, infants’ tendency to look back and forth between two objects, then commit attentional streams to one of the objects, appeared more controlled, and less compelled by exogenous, “in the moment” salience.

The exploratory routines at this age also changed in type and redundancy. At six months infants maintained possession of an object for longer, and used a consistent basic set of simple manipulations (e.g. shake, hit, mouth). In contrast, at nine months, although infants switched objects more frequently, they tended to repeat an object-specific routines (e.g. bang the toy, or make it spin) many times in succession, sometimes punctuated by looking at their mother (see below). This is reminiscent of Piaget’s description of secondary circular reactions (Piaget, 1954), but interleaved with gaze to partner. Also, nine-month-olds tended to “follow through” by directing a focused gaze to the objects when their action had an effect (e.g., dropping an object and then leaning to stare at it on the floor; slapping an object to make it rock, then pausing to watch). Also, infants’ manipulations were more diverse and seemingly selective. For example, infants touched one object with another, or
carefully stroked, rubbed, or pinched a part of an object. These actions were suited to the specific affordances of the object (Bourgeois et al., 2005; Lockman, 2000).

Most markedly in the videos at this age, the infants frequently had two hands engaged at once, with a different object in each hand. Similar to six months, one hand would passively maintain contact with (usually grasp) one object while the other actively manipulated the other object. However, while at six months the active hand did not grasp the object being haptically explored, at 9 months the active hand would grasp, lift, and manipulate this object. Gaze was usually directed at the active hand, but shifted to the passive hand soon before they started manipulating that hand’s object. Both hands were simultaneously active only when jointly manipulating one object or bringing it to the mouth.

9 Month Dyads

Infants’ transitions between objects were based both on their own initiative and mothers’ presentations of new objects. When infants took up their mother’s bids, they fully engaged the new object. However, unlike at six months, they did not then abandon the original object. We characterize the flow of attention, in these interactions, as a divergent or distracted trajectory following the mother’s bid. That is, the mother’s object did not disrupt the infant’s momentum in attending to the previous object. Also, infants stopped showing negative affect to mothers’ bids, although they often tried to retrieve the object that she had removed. Thus, infants
could coordinate attention to the mother’s bid without losing track of their previous activity.

Perhaps because of infants’ increased tendency to initiate and repeat particular exploratory actions, mothers presented new objects less frequently. They also tended to shift their own attention to the object of the infant’s engagement, and even imitate or facilitate the infant’s actions. Alternation of actions was more common at this age, and infants sometimes paused to watch their mother and then resume to their own activity.

Although mothers did most of the imitation, infants occasionally imitated mothers. This involved actions the infant had produced earlier (e.g., banging an object), but now repeated immediately after the mother, and sometimes followed by smiling at the mother. Also, bouts of turn-taking were sometimes followed by mutual gaze and positive affect. Although such events were infrequent, every infant had at least one such sequence.

12 Month Sensory-Motor Coordination

At 12 months, infants showed improved bimanual organization, more differentiated articulation, and longer, more elaborate routines.

As at nine months, the two hands/two objects pattern was common, as was the one active/one passive pattern. Infants still sometimes directed both hands to a
single object. However, at 12 months, infants were bi-manually coordinated. That is, at nine months, when both hands were active they were only directed to a single object. By contrast, now each hands would grasp and manipulate separate objects. This resulted in activities such as clapping the objects together, or touching/rubbing them together.

This active bimanual coordination was also visually mediated in a new way. At nine months, when each hand grasped an object, whichever one the infant looked at was the active hand. A 12 months, infants looked rapidly back and forth between the two objects, sometimes during bimanual activity (e.g., object-clapping). Infants also seemed to visually attend to particular parts of the objects. For example, we observed an infant repeatedly “run” a toy ladybug along a surface, then turn it over and visually examine the wheels in its underside. Such focused “inspection” often entailed holding the object with one hand and actively probing it with the other.

Twelve month-old infants continued to repeat action routines, but routines were longer and included more different actions and foci of attention. For example, an infant pick up and squeeze an object, then shake it, and then repeat the more elaborate sequence. Infants at 12 months also began to embed sub-routines into longer manipulative sequences.. For example, we observed an infant who was holding two blocks adjust one to align its face to the other, before clapping them together. Some routines resembled the experiment-like “tertiary circular reactions”
described by Piaget, in which the infant explored a variety of actions, both novel and familiar, for producing specific effect on the object. For example we observed an infant squeezing a rubber toy using one hand, two hands, a full fist, just the fingers, and by pressing it against the floor. Finally, at this age infants often verbalized, both while handling objects and while watching their mothers’ (see below).

12 Month Dyads

The most pronounced change at this age was in the coordination between mother and infant, mediated by changes in both the infant’s and the mother’s behavior.

Mothers’ actions on toys appeared to have a novel appeal. If the mother and infant performed a similar action in synchrony, the infant could show a ratcheting-up of its arousal. For example, one infant lifted an object with two hands. The mother then raised her two (empty) hands, and this mirroring action captured the infant’s attention and led to eye contact and mutual laughter. Thus, at 12 months, synchronous similar co-activity appears pleasant for both participants.

Infants often paused their haptic activity to watch their mother perform an action with an object. This response was now embedded in a new sequence. Whereas nine-month-olds would manipulate their object, pause to watch the mother, and then resume their original manipulation, at 12 months infants sometimes modified their activity upon resumption. For example, the infant might
return their gaze to the held objects and engage in the observed activity with those objects. Occasionally infants engaged in considerable effort, and multiple strategies, seemingly in an effort to reproduce the outcomes of the mothers’ activity. For example one infant spent many minutes trying to stack blocks—horizontally and vertically—before succeeding. Sometimes during these efforts infants stopped to return their gaze to the mother’s hands. When this occurred, mothers tended to respond by repeatedly demonstrating the target action in the infant’s field of view. While frustration in these settings typically led the infant to look at the mother’s hands, success led the infant to look at the mother’s face and show positive affect, or, alternatively, success did not lead to changes in affect or shared gaze (for similar observations, see, e.g. Ross & Lollis, 1987). The infants' persistence at, and variability within, such routines gave them the feel of “projects”, in which the infants now also actively involved the mother.

Projects fit the classic characterization of triadic interactions. Other triadic interactions at this age included a new configuration of activity in which infants grasped an object and then extended it toward either the mother’s hand or, less often, her face (who extended her hands to receive the object. Such “giving” routines never occurred at nine months. Another type of triadic interaction was turn-taking bouts, in which alternating actions would be repeated by the dyad, with the particular actions changing little over the repetitions of the sequence. These often
involved infants replicating the mother’s action. Such actions encompassed object manipulations, changes in posture, and vocalizations (e.g., nonverbal sounds, “sound effects” for the objects). Finally, triadic interactions emerged as complementary actions within a routine. One mother, for example, put a cover over a toy, turned to the infant, and performed a gesture-complex that included spreading her hands, shrugging her shoulders, and asking, “where is it?” The infant watched her, then pulled off the cover, looked up at the mother, and expressed positive affect. The dyad then repeated this sequence.

The increasingly formalized alternation of sequences of coordinated activity, or “roles”. were demonstrated in a number of ways. First, infant’s activity did not go forward unless the mother performed her particular behavior. If the mother failed to repeat her actions, the infants could respond to this situation with negative affect. Alternatively, infants performed their next actions in the routine as a way to solicit the mother’s further participation. For example, we observed a mother squeezed a toy to blow air on the child’s cheek. After several iterations, when the mother refrained, the infant looked toward the mom, then turned his cheek toward her and vocalized a “whooshing” sound.

While the infants were clearly initiating many of these interactions, mothers continued to play an active role. As at nine months, mothers imitated their infants’ actions. Mothers also facilitated infant’s ongoing activity, as in the past, but with new
subtlety. Because the infant’s activities were now more elaborate, multi-part sequences, mothers could shape and time their contributions to particular aspects of an activity. For example, upon seeing her child’s difficulty in squeezing a block, one mother demonstrated a two-handed squeeze. When she was had gotten the infant’s attention, she briefly froze her hands in an exaggerated fingers-spread position, closed them slowly on the object, and then slowly re-opened them. She then asked, “Can you do that?” and allowed the infant to take it from her.

Mothers also used more directive language, gestures and symbols with their infants. Patterns of maternal speech to 1-year-old infants have been documented in many studies (e.g., L. B. Adamson & Bakeman, 1984; Bates, 1979; Goldfield, 1993). We are preparing a report of speech and gesture content during these interactions, and their contextual embedding in dyadic interaction. For the present, we note that mothers used speech in more in more elaborate ways than previous sessions: they drew infants’ attention to objects and actions, named specific objects, and “shaped” their verbalizations to fit the current activity. Also, mothers more often pointed to objects to direct infants’ attention. Notably, they pointed to novel objects or configurations not currently being handled—sometimes not even within the infant’s reach—and, for the first time, pointed to details or parts of objects to which the infant was attending, as if to show the infant particular features (see Zukow-Goldring & Arbib, 2007).
Finally, mothers now began to use other manual gestures: for example, extending the hand palm up, in an “requesting” gesture, or waving while saying “bye-bye.” These gestures were coordinated with speech and with the current activity (e.g., holding up two fingers while saying “two,” clapping while singing “If you’re happy and you know it…”). Certainly mothers were driving all of these symbolic embeddings, but infants occasionally made efforts to copy a gesture, or to vocalize in synchrony or alternation. Thus around 12 months we saw dramatic increases in what Adamson (2004) has called “symbol infused joint attention.”

Altogether, these changes show a mutual adaptation: of scaffolding, on the mother’s part, to the infant’s growing abilities, and of attention and action, on the infant’s part, to the opportunities afforded by the mother.

In sum, at 12 months, the infant’s attention was neither “disrupted” nor “distracted” by the mother’s bids but, rather, often incorporated the mother’s activity, during the toy bid and otherwise, into his or own object activities. Both participants in the dyad showed more adaptive readiness, and behavioral range, to engage in and promote collaborative activity, The conjunction of mothers use of speech and gestures, their refined tactics for scaffolding, and their engagement in ritualized activities, produced distinctive triadic interactions. As a result, we characterize these interactions as “Well Coordinated.” Note that we also characterized interactions at four months as “Well Coordinated”, due to how highly
labile infants’ attention was to maternal bids. Now, coordination shows a very different level of complexity.

Discussion: Developmental Trajectories

Based on our observations, we propose a set of developmental trajectories that characterize the increasing sophistication of mother-infant sensory-motor coordination over the first year (see Table 1). These trajectories include an increasing number of loci of attention, refining haptic articulation, the emergence of routines, and increasing continuity between episodes of activity. By contrast, we propose, some elements were present from the youngest age: shared interest in dyadic activities, and the dyads simple-yet-coordinated attention to objects. Changes in infant social interaction parallel changes in attending to objects across the first year: as infants show more complex behaviors with objects, they produce ever more complex responses to their mother. Also, mothers’ behaviors towards the infant and the objects also changed and diversified. By altering the infant’s environment, and the affective expressions she associates with it, she too helps shape each interaction and thus contributes to how triadic attention develops. Our analyses lead us to argue that 12-month complexity is not sudden or discontinuous but a culmination of continuous changes along these trajectories.
Consistent Properties of Dyadic and Triadic Interactions

Coordinated Arousal and Attention.

From the earliest months, infants show positive affect, including smiling and laughing, during play with their mother. At four months, mothers often initiated reciprocal affect (Hsu & Fogel, 1993). The joint positive affect of interactions at the earliest months foreshadows the more complex interweaving of affective exchanges and shared activities in later months.

Dyads also coordinated their actions and attention to objects from the first months. Four-month-old infants found their mother’s object-in-hand highly salient, and mothers used objects to deliberately draw infants’ haptic and visual attention. While such “well-coordinated” attention does not reappear until 12 month session, this “you touch/I touch” might be a rudimentary pattern that leads to mirroring a partner’s actions at the end of the first year.

Changing Properties of Dyadic and Triadic Interactions

Multiple Loci – Decoupling Attentional Streams.

Over the first year infants gradually came to attend to several objects simultaneously (For a similar pattern of results, see Belsky & Most, 1981; Bruner, 1973; Fenson, Kagan, Kearsley, & Zelazo, 1976; Kotwica, Ferre, & Michel, 2008.) At four months, attention was “all or none” towards a single object; whereas at six
month infants could grasp one object while gazing and reaching for another. Yet when mothers presented another object, six-month-olds looked like an unstable dynamical system, teetering on a ridge between one object and the other. Their difficulty seemed to be reflected in over-arousal (e.g., gaze aversions) and negative affect. By nine months, all infants easily alternated activity between two targets of attention, without negative affect. However, they held objects apart and did not gaze at them simultaneously.

The transition to attending to multiple loci also involved a gradual decoupling of the visual and haptic modalities. At four months, when exploring an object, gaze and touch were virtually inseparable. At six months, the infant visually mediated its reach until the object was collected, but could then direct gaze elsewhere. At nine months, they could shift gaze rapidly between objects, and even look and reach to the next object while still examining the current one. Thus, changes in the timing of coordinated actions allowed the incorporation of an increasing number of loci within a behavioral routine. The mother’s face – a focus of interest from the beginning – also shifted from being a separate, momentarily all-consuming focus of attention to one among several attended targets.

Haptic Articulation.

Within episodes of object exploration, several changes were observed. In terms of hand shapes and adaptation of actions to object properties, there was
predictable refinement. At four months object handling was bilateral and symmetrical; at six months infants often ‘meandered’ through a variety of simple, exploratory manipulations; at nine months they produced small elaborations and precise, affordance-dependent touches.

Emergence of Routines.

Between the fourth and ninth sessions we saw a slow emergence of repeated multipart sequences of actions within the infants’ object play, what we called “routines”. At four months, the undifferentiated fingering provides little evidence for organization of actions. At six months manual actions are more differentiated but organization is still simple: infants appear to ‘meander’ through a variety of simple, exploratory manipulations on an object. At nine months, infants were able to chain together a series of simple actions into sequences which they fluently repeated. The repetition of the sequence of behaviors indicates that the infants have learned that each action is preceded by, and follows, another. They also began to show “follow up” attention to objects upon which their own actions had had a visible impact, for example, gazing to objects that had fallen off of the tray. Again, here, the infant has chained together a sequence of activities wherein falling is followed by the outcome of being on the ground. By 12 months, these multipart sequences evolve to be performed across the dyad, such that some elements of the chain are enacted by the mother, and some are enacted by the infant.
Extending Continuity of Action

At four months, infants’ convergence of attention on one object, and utter distractibility by the mother’s face, suggest that there is little continuity between his one action and the next. At six months, infants did not manipulate objects in a consistent way over time. In contrast, at nine months, upon returning to a previously attended object, infants tended to repeat their previous actions on that object, even after a delay or intervening activity. At twelve months, infants not only repeated their routines, but also incorporated their mother’s actions. Thus 12-month continuity entails not only attending to multiple loci but also parsing, or recognizing, the mother’s activity on an object and reproducing this activity on another object.

Parsing Action Trajectories

In order to imitate the mother’s actions the infant must take what is available from the visual input of the mother’s actions and translate that into motor commands that he can perform with his own body. We conceive of actions as complex dynamic input pattern, that must be recognized and segmented, or “parsed,” into a sequence of simpler actions (e.g., reaching for an object, shaping the hand, grasping it, bringing it towards oneself, and manipulating it). Typically each step has a nested sub-structure of very specific moves (e.g., reorienting blocks in a series of small adjustments, in order to align them). The complexity of this structure, which usually looks smooth and effortless, becomes apparent when it is
compromised by neurodegenerative disorders (Sainburg, Poizner, & Ghez, 1993). Moreover, recognizing motor sequences from visual input is a challenging task, indeed, it is considered one of the most difficult problems in machine vision (Jhuang, Serre, Wolf, & Poggio, 2007). In order to imitate, infants must recognize these actions despite differences in orientation, body size and available motor skills (Wohlschläger, Gattis, & Bekkering, 2003). The infants’ experiences over the first year might contribute to progress in infants’ recognizing their mothers’ actions from visual input. Over the first year, the infant has repeated experiences of reach, grasp, manipulate. Eventually, he combines these actions into repeated multi-part sequences of actions. Additionally, as his articulation becomes more refined, he increasingly modifies his behavior to the affordances of the objects. We suggest that the infant’s observation of self generated actions may contribute to the infant’s attunement to the refined, multipart actions of the mother, and the infant’s recognition differentiation and organization of this activity into parts that they can replicate.

**Changing Maternal Scaffolding**

Mothers also showed a developmental trajectory in the scaffolding that they provided. All mothers played a dominant role with four-month-olds: they delivered objects to infants’ hands, waggled them or made them loom, etc. At six months, mothers’ bids were often rebuffed by their infants, and they shifted to following the infant’s attention. By nine months, when infants had established solo play routines,
the mothers began to imitate and temporally coordinate these, often with exaggerated motion and positive affect. Indeed, the mothers used positive affective displays in temporal synchrony with turn-taking actions, as if to punctuate the sub-events of the game. Moreover, as mothers began repeating infants' actions, they provided a new selective environment: they showed infants repeated examples of their own manipulative actions, performed by another.

At 12 months, the mothers’ actions changed further, to fit the infants’ refined manual skills and capacity to organize action. For example, because infants could attend to multiple loci, mothers no longer needed to loom or statically present objects. Instead they could scaffold infants’ activity by demonstrating it—in an exaggerated manner (Zukow-Goldring & Arbib, 2007)—while the infant watched.

_The Emergence Of Twelve Month Triadics_

At 12 months, we see a number of novel high-level configurations of mother infant object engagements. We focus on three main types of triadic engagement at 12 months: projects, imitation, and games. We also note developments in infant gesturing. Our descriptions acknowledge a significant jump in complexity in our dyads between the nine and twelve month videos: the high-level configurations observed at 12 months are simply not present in earlier months. However, rather than resorting to postulating a novel representational ability as the cause of this
activity, we can view their novel presence as a continuous developmental outcome of a number of developmental trajectories.

In both imitation and games, infants divided their attentional modalities between two independent loci: objects in their possession, and activities/objects of the mother. Infants incorporated elements of their mother’s action and objects to their own ongoing object activities. As the infant’s haptic differentiation becomes further refined, and he becomes more proficient with performing his own multipart sequences, we suggest he becomes more attuned to the actions of his mother’s hands, and is able to parse the manipulations she performs. The attentional continuity he now maintains across multiple loci may also facilitate his translating the activity he has parsed from objects within his mother’s possession to the objects within his possession.

Our infants showed a slowly developing propensity to continue sequences of manipulations across spatially distributed targets. At 12 months, this develops to allow the infant to incorporate elements of the action the mother performs at one locus, objects in her possession, to his own ongoing manipulations with objects in his possession.

We assert that the bridging of these sensory-motor ontogenies with interest in shared activities with the mother transforms the early “you-touch /I touch” into a
replication of the particulars of mother’s actions. This manifests as either imitation or experimentation.

The changing social ontogeny also plays a role in this transition. For some time before he became so proficient, the infant had been immersed in an imitative environment with repeated opportunities to observe his own acts (and his vocalizations) being reproduced, and envalued with positive affect, by his mother. The mothers’ excited expressions might have cemented infants’ attentiveness to this action mirroring, and enhanced their enjoyment of the reciprocity of the interaction.

Games share the above elements of imitation; additionally, they have accrued additional structure –repeated patterns of alternating activity between the mother and infant, marked by negative affect at moments of disruption- that enables us to identify participant-specific actions, or “roles”, in the joint activity. Above we describe how by nine months, the infant establishes and performs routines—repeated sequences of actions—on objects in his own possession. Experience with these solo routines likely provides additional structure to the imitated or complementary actions that have emerged by 12 months, leading to the emergence of co-dependent routines. The mother’s actions at nine months also likely help. By performing her part of the routines on cue, readily reciprocating the infant’s positive affect, and acting surprised when something went awry, the nine month mothers helped establish regularity in the dyad’s turn-taking cycles, and helped highlight their
disruption. This would lead, at 12 months, to the further elaboration of “games” and to promoting the infant’s parsing of his own, and his mother’s, activity.

Triadic activities are also characterized by face-to-face expressions of affect and mutual gaze which occur at action-dependent moments during prolonged routines. This is an elaboration of a pattern of co-arousal that has existed since the earliest months. What has developed is the integration of affective exchanges into increasingly complex exchanges. For example, in the case of games, the regularities produced by the co-repetition of mother-infant routines can also make a break in that pattern salient. Thus, a repeated, fun-filled exchange, disrupted by the mother’s failing to play her role, provokes a display of negative affect in the 12 month old infant. While the timing and conditions of such affective displays has changed, their occurrence at 12 months is interpretable as the outcome of mother-infant co-arousal embedded in the elaboration of routine.

Finally, gesture became an important component of 12-month dyadic activity. Infants’ first gestures were re-creations of portions of their imitative activity or game that the dyad had recently played. Now, however, the action was repeated without the mother’s participation. Thus, an action that had been organized around a material outcome seemed to be replicated with a social motive: to influence the mother. Infants had already, for many months, provoked distal effects on their mothers with facial expressions. At 12 months, they extended that by producing
manual and postural actions that evoke responses. When mothers responded, infants showed gleeful satisfaction. In this way, mothers’ scaffolding might contribute to motivating further adaptive selection for communicative actions that are “once removed”—that is, taken outside of their immediate object or game context.

What about conventionally symbolic gestures? Most of these, even at 12 months, were produced by mothers. Mothers’ pointing (deictic) gestures compelled infants to shift their gaze to a target, and thus focus—at least partly—on particular aspects of her target object or activity. Mothers’ iconic gestures, for example reaching out to receive a toy, scaffolded infants’ participation, even to initiate a game. By making their gestures slow, simple, and visible, parents might further have scaffolded infants’ own budding use of non-verbal symbolic gestures (Bates, 1979). Thus, although still a “new frontier” for 12 month olds, the production of scaffolded symbolic gestures by collaborative mothers suggested an impending foray into new social-communicative territory.

**Complexity and Continuous vs. Discontinuous Development**

As in many previous studies of triadic interaction, we too saw a significant jump in complexity in our dyads between the nine and twelve month videos. This radical shift in the complexity of behavior has lead many researchers to propose a discontinuous mechanism to explain that complexity, namely the appearance of a novel cognitive ability.
Discontinuous accounts are characterized by their supposition that it is a novel representational capacity that is responsible for the apparently unprecedented behavior. A common such supposition in the 12 month literature is the capacity to represent “goals”. It is this new capacity that is required, for example, for the infant to understand what the mother is “trying to do”, and further underlies his ability to produce a similar “attempt” himself. Inter-subjectivity would be held responsible, for example, for the infant smiling at his mother after imitating her, or succeeding at a task that she had just demonstrated. In such situations, the participants’ affect is taken as their mutual acknowledgment of the fact of their joint engagement. That is, it is the infant’s representation of the mother’s violation of their tacit agreement to co-engage in the routine that is invoked to account for the infant’s displeasure. A final supposition of many discontinuous accounts of 12 month triadics is the notion that communication is “intentional” (Grice, 1975). It is the participants’ representation of gestures and vocalizations as “intended” that is said, for example, to enable the infant to understand the mother’s pointing at an object.

These accounts provide an intuitive scheme for organizing the interactions were observed in this study, fitting our folk-theoretic accounts of such interactions. However, positing the appearance of these novel representational abilities still leaves the cognitive origins of these activities a mystery. Many of the papers on triadic attention that have postulated mental representations like “goals”, “intentions” or
“inter-subjectivity” have been insightful and informative. However, they run the risk of committing the “nominal fallacy” in which the description of an event is taken as isomorphic with its mechanism. They also keep the individual at the center of the account, which causes problems for any explanation of social phenomena. An alternative approach is to focus on how bodies engage with the world and with each other. By remaining focused on activity, and by viewing the dyad as a system, our account can identify changes in the organization of that activity across developmental time. Over the four-12 month period, we observed our infants demonstrate increasingly complex routines, involving increasing haptic differentiation, more loci of attention, as well as multipart sequences and embedding. Infants changing behaviors elicit novel behaviors from the mom which in turn provide novel support and reinforcement towards twelve month triadics.

Our description acknowledges the emergence of novel collaborative triadic activities at 12 months. However, when we track the interaction as multimodal multiparty configuration, we no longer need to resort to postulating a novel representational ability to explain the emergence of these new activities. Instead we observe that the actions at each session build off of those viewed at the earlier sessions. In this way, an embodied and distributed approach provides an alternative account of triadic attention that is grounded in continuous changes in the activity of our participants rather than a simple shift in internal structures.
<table>
<thead>
<tr>
<th></th>
<th>Co- Arousal</th>
<th>Loci of Attention</th>
<th>Haptic Articulation</th>
<th>Emergence of Routines</th>
<th>Continuity of Actions</th>
<th>Maternal Scaffolding</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 Mo</td>
<td>Eye contact pleasantly arousing for both.</td>
<td>All modalities converge on one target, abandon previous.</td>
<td>Bimanual, symmetrical clapping, bring to mouth; Undifferentiated fingering.</td>
<td>One simple activity at a time.</td>
<td>One simple activity at a time.</td>
<td>Mom loom own face, initiate eye contact &amp; positive affect; Waggle object-in-hand in line-of-sight, touch to infant’s hand (give).</td>
</tr>
<tr>
<td>6 Mo</td>
<td>Infant directs some negative affect, esp. when mother removes attended object.</td>
<td>Can passively grasp one target, while look to/touch another; Returns to grasped object; Too much active attention shifts all to new target.</td>
<td>Extend arm, <strong>grasp</strong>, retrieve object; Simple exploratory manipulations (hit, rub, grasp, push away, etc.)</td>
<td>“Meander” through a series of simple exploratory manipulations on a single object</td>
<td>Track object over delay, but at return begin a different type of manipulation rather than return to prior</td>
<td>Mom follow infant’s attention to object &amp; facilitates its further access.</td>
</tr>
</tbody>
</table>
### Table 4 Developmental Trajectories, Continued

<table>
<thead>
<tr>
<th></th>
<th>Co-Arousal</th>
<th>Loci of Attention</th>
<th>Haptic Articulation</th>
<th>Emergence of Routines</th>
<th>Continuity of Actions</th>
<th>Maternal Scaffolding</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>9 Mo</strong></td>
<td>Mom initiate positive co-arousal over “<strong>co-same</strong>” (when imitate infant’s action);</td>
<td>Fluently alternate loci; Easy transitions between active/passive manip</td>
<td>More elaborate sequences and more precise, affordance-dependent handling of obj; Simple embedding (e.g. hit with object-in-hand)</td>
<td>Multi-part sequences; Iterate rapidly with overlapping timing; Follow-thru, actively observe consequences.</td>
<td>Routines retained over time, including overhandling of a different object</td>
<td><strong>Mom imitate</strong> infant sequences, w/exaggeration and pos. affect; Mom complement familiar routines</td>
</tr>
<tr>
<td><strong>12 Mo</strong></td>
<td>Infant initiate positive affect upon his own imitation of mom’s action; Infant look to mom’s hands in frustration, to mom’s face in success; Infant initiate w/give.</td>
<td>Multiple loci: e.g. Infant gazes to mom’s object-in-hands, then similarly manipulates own obj.</td>
<td>Refined embedding (e.g. adjust object-in-hand to suit act with other object-in-hand); Variable actions that have same effect on object; Extend object-in-hand (give).</td>
<td><strong>Games</strong> involve repeated sequences of actions across dyad performed in alternation; longest persistence, esp. at outcome-based “<strong>projects</strong>”;</td>
<td>Repeat not only own actions, but also mom’s (i.e. <strong>Infant imitate</strong> mom).</td>
<td>Mom precise demonstrations, some performed in viewing but not touching distance from infant, consistent with infant’s “project”; Mom point &amp; gesture</td>
</tr>
</tbody>
</table>
References


Chapter 2, in part is currently being prepared for submission for publication of the material. De Barbaro, Kaya; Johnson, Christine M.; Forster, Deborah; Deák, Gedeon O. The dissertation author was the primary investigator and author of this material.
CHAPTER 3: SENSORY MOTOR DECOUPLING CONTRIBUTES TO 12 MONTH SOCIAL “REVOLUTION”: A LONGITUDINAL INVESTIGATION OF MOTHER-INFANT-OBJECT INTERACTIONS

de Barbaro, K., Johnson, C.M, Forster, D. & Deák, G.

Introduction

As humans we differentiate ourselves from other species by our interest and willingness to teach and learn from others (Tomasello, 1999; Vygotsky, 1978). We share interest in what others attend to, and what they are doing. The skills for sharing develop in contexts of joint, or triadic, attention. In its simplest definition, triadic engagement refers to interactions between two partners that are mutually constrained both by each other and a shared focus of attention. In infant-parent dyads, these episodes provide an opportunity for the infant to access and participate in the practices of the parent (Lave & Wenger, 1991). In this regard triadic attention is the original learning environment. Empirical studies have indicated that triadic interactions support a wide range of developmental practices, including language learning (Bruner, 1975; Smith, Adamson, & Bakeman, 1988; Tomasello & Farrar, 1986) social skills (Bornstein & Tamis LeMonda, 1989), interactions with objects (Lockman & McHale, 1989; Zukow-Goldring & Arbib, 2007) and culture (Bruner &
Haste, 1987). Looking beyond this broad characterization, the literature offers many operational definitions of triadic attention. Some involve coordination of partners’ gaze to objects, as in gaze following (Flom & Pick, 2005; Scaife & Bruner, 1975) and gaze alternation (Newson & Newson, 1975; Striano & Stahl, 2005). Others involve coordination of actions on objects, as in certain types of games and imitation (Bakeman & Adamson, 1984; Ross & Lollis, 1987). Pointing and gesturing may also be considered triadic activities, as they reference objects or events external to the dyad (Butterworth & Grover, 1990; Leung & Rheingold, 1981). In social referencing, infants change their behavior towards an object based on their observation of an adult’s behavior or affect towards that object (Carver & Cornew, 2009; Slaughter & McConnell, 2003; Walden & Ogan, 1988). Finally, some kinds of blocking and “teasing” episodes, wherein an adult forcibly affects the infant’s current activity (e.g., by covering an object of the infant’s current attention), and the infant responds by looking up at the adult, are also considered triadic interactions (Carpenter, Nagell, Tomasello, Butterworth, & Moore, 1998; Phillips, Baron-Cohen, & Rutter, 1992).

The current conventional model is that these triadic phenomena all are the consequences of a single underlying mechanism, namely, infants’ emerging ability to infer the intentions of a social partner (Striano & Reid, 2006; Tomasello, Carpenter, Call, Behne, & Moll, 2005). The evidence for this theory is that a small number of triadic behaviors emerge in a semi-predictable age sequence (Carpenter et al., 1998).
Certainly there is converging evidence that triadic behaviors become increasingly prevalent from 8 to 18 months of age. Thus some scholars describe this period as a “social revolution” (Adamson & Bakeman, 1991), or the birth of “secondary intersubjectivity” (Trevarthen & Hubley, 1978). This latter term refers to the notion that what underlies the changes at the end of the first year is an awareness of the subjective experience of others thought to follow primary intersubjectivity. In any case, the causal mechanism of this model is thought to be a shift to a novel conceptual-representational capacity. This proposed representational shift is internal to the infants’ CNS; thus as researchers we do not have direct access to it. Hereafter, then, we will refer to this model as the “conceptual/intersubjectivity,” or C/I, theory of triadic skill development.

The C/I theory, for all its appeal and its parsimony, raises several concerns. One is the status of evidence that many triadic behaviors regress to a single underlying factor. Unfortunately, regularity in the ages at which infants “pass” several specific tasks (Carpenter et al., 1998; Slaughter & McConnell, 2003; Striano & Bertin, 2005) is merely circumstantial evidence. If multiple cognitive factors change between 8 and 18 months, and meanwhile infants gain different experience with various triadic skills, we could see a pattern of change over time due to any number of factors. Put differently, whereas a heterogeneous timeline of development can imply multiple underlying factors, the opposite is not true: even if triadic skills
develop in a very organized manner (a claim that itself would have little support), is not sufficient evidence for a singular underlying factor.

Recent projects have recently explored another class of model, wherein well-specified changes in lower-level factors might explain discontinuous changes in high-level behavioral skills (e.g., Grossberg & Vladusich, 2010; Simmering, Schutte, & Spencer, 2008). For example, Triesch, Teuscher, Deák & Carlson (Deák, in press; 2006) modeled the development of gaze following as the product of interacting factors, including infants’ reward values (e.g., seeing their parent’s face), the “structure” of the social environment (e.g., parents’ looking patterns during face-to-face play), the infant’s learning mechanisms, including reinforcement learning and habituation, and other perceptual and affective variables (e.g., attention-shifting abilities). This model is informed by neuroscience and behavioral evidence. Notably, in computational simulations, triadic behaviors (e.g., gaze-following to invisible objects, social referencing, and gaze alternation), emerged without any higher-level cognitive “engine” such as a C/I module.

Such results show that multiple lower-order, biologically and ecologically plausible factors can systematically lead to the emergence of higher-level behaviors. By labeling some behaviors as triadic, we implicitly hint at a single underlying cause. Yet patterns of emergence can be due to the rate of maturation of different factors, to heterogeneity of experience in the social environment, and/or the differential
reward/cost of different behaviors. In fact, by varying almost any factor in simulations, we can change the emergence of various behaviors (Teuscher & Triesch, 2007; Triesch et al., 2006). Clearly, then, findings that infants pass tests of some triadic skills before others do not support one-factor high-level (C/I) models of triadic development.

A related concern is that C/I and similar models of triadic development privilege maturational accounts of internal or individual conceptual “capacities” over changes in sensorimotor development, and over contributions from social partners’ behaviors. For example, development of the motor system has been described as limiting infants’ responsiveness to certain social distinctions (e.g., Trevarthen, 1979), as though a mature cognitive capacity was being masked by a “superficial” motor one. Insights from theories of embodied and distributed cognition suggest that the organization of sensorimotor systems in real time is cognitive activity (Clark, 2008; Hutchins, 1995). An important implication of the distributed frameworks is that one cannot exclude the development of sensorimotor systems from any explanatory account, any more than one could explain walking without considering the biophysics of muscles and joints. The tendency in the psychology literature to “privilege” internal representations as the drivers of cognitive activity has been criticized

---

2 The problem is confounded because in studies such as Carpenter et al (1998), where each triadic behavior was measured by only one task, each test is subject to multiple unexplored sources of methods variance (e.g., action-complexity of “passing” actions; uncertainty-conditions in each task; uneven scoring criteria), as well as differential practice and fatigue effects.
elsewhere (Oyama, 1985; Thelen & Smith, 1994). The alternative view is that complex behaviors emerge from the interplay of a dynamic system of heterogeneous factors, including factors both internal and external to the central nervous system of the infant. This view does not preclude that higher-order conceptual representations play a role in triadic social behaviors. However, neither does it arbitrarily assume that those representations are of principle importance.

Previous longitudinal studies of triadic interaction carefully operationalize particular configurations in infants’ responses to the adult and a third object. The infant’s behavior at each age is coded only as “passing” or “failing” some discrete high level behavioral criterion. This paucity of data collected provides no information about how the changes in incidence of the phenomenon of interest come about. In this way, the prior methodologies almost force researchers to attribute the sudden appearance of triadic behaviors to a change in an invisible cognitive capacity of the infant (e.g. Stern, 1985; Tomasello et al., 2005; Trevarthen & Hubley, 1978).

By contrast, a distributed cognitive account focuses on each partners’ access to the others’ activities: their motions, words, and affect through space and time (Forster & Rodriguez, 2006; Hutchins, 1995; Johnson, 2001). Under this framework, the increasing complexity of triadic interactions is made visible via changes in the way that partners attend and respond to the emergent actions of one another. This shift is a boon to developmentalists. Specifically, while higher-order representations
are hypothetical and assumed, the sensorimotor activities between partners, such as parents’ bids for infant attention, or the distribution of infants’ haptic and visual attention in response to those bids, are overtly measurable. By grounding our developmental account in visible changes in the interaction across the first year, we aim to provide a more coherent developmental account of how triadic behaviors emerge.

The current study

In the current study we focus on the development of shared actions on objects, such as imitation and turn-taking games. Both imitation and games involve the mother and infant engaging with objects via visual and haptic modalities. Our analysis of moment-by-moment changes in these independent perception and action channels (e.g., mother’s gaze; infant’s hands) across the first year show that from the earliest months infants attend to objects being manipulated by their mothers. It is the organization of actions; the timing and sequence of the activity both participants, as well as the particulars of the use of these channels that changes through longitudinal time.

Thus, we characterize states of triadic interaction across the first year by examination of moment-to-moment changes in attentional activity in infant-mother face-to-face interactions with a set of objects (see also de Barbaro, Johnson, & Deak, in prep). Behavioral coding yields a detailed, dense record of events, embedded in an
ongoing stream of activity that can be examined for contingencies at different time-scales, from 10s of msec to minutes. Also, we track the same infant-parent dyads longitudinally, observing them when the infant is four months, six months, nine months, and 12 months of age. Thus, dense behavioral records are compared at several key points in development, so that developmental trends can be imputed across months.

By utilizing multiple quantitative methods to describe the changing patterns across these channels, we observed a coherent set of developing behavioral trajectories that, we suggest, reveal a *continuous* pattern of development of shared actions on objects. Moreover, our approach provides a novel *mechanistic* developmental account of the changes in infancy that account for the emergence of triadic behaviors. The analyses we undertake in this paper stem from our previous observational work (de Barbaro et al., in prep), which suggested a novel feature of imitation and games episodes. Specifically, during periods of imitation, infants often alternated gaze between toys manipulated by mother and toys they were manipulating themselves (see Figure 2). We believe that this distribution of attention is important for early imitation.
The y-axis indicates a specific partner (infant or mom) and sensorimotor modality (gaze, hands). Colored marks indicate, frame-by-frame, moments of sensorimotor contact available objects. Contact with partner is always indicated in yellow; remaining colors indicate contact with objects. Widths and additional markings on the colored bands are not relevant for this study. The thick blue line with darker inner beads in the row “Dyadic Attention” indicates moments of imitation or games. Note that during the imitation episode, the infant is alternating his gaze between the toys he is manipulating (teal and orange toys) and the toys the mother is manipulating (green and grey). While infant attends to both sets of toys over the duration of imitation, proportionally he attends more to the toys he is manipulating.

Mothers’ manipulations provide the form for imitation, and infants often repeatedly gazed back to observe their mother’s productions of a particular manual action during the course of apparent reproductions of that action (Zukow-Goldring & Arbib, 2007). Additionally though, recreating the actions also required intense
focused attending (Ruff & Saltarelli, 1993) with coordinated manual and gaze over
the course of multiple attempts to reproduce the actions (not always successfully) on
their own toys (de Barbaro et al., in prep).

Importantly, we noted a developmental trajectory in infants’ tendencies of
distributing attention. Younger infants did not appear as adept in distributing their
attention across multiple objects in the same way but rather appeared to attend to
only one toy at a time. At six months, infants began to distribute their attention
between toys manipulated by the mother and previously attended toys, but this
appeared relatively unstable. By nine months some infants, and by 12 months, all
infants, were very adept at attending to two loci (See Figure 3; de Barbaro et al., in
prep; see also Belsky & Most, 1981; Bruner, 1973; Fenson, Kagan, Kearsley, & Zelazo,
1976; Kotwica, Ferre, & Michel, 2008.). Attending to only a single locus of attention
appeared to curtail the complexity of possible interactions with mother: it allows
infants to engage with the toy that the mom presents, but they are not able to do the
kind of alternating between the two loci that appears required to repeatedly observe
the mother’s actions and then perform those actions on objects in their own
possession. This suggested to us that distributing attention across multiple loci was
an important mechanism contributing to the emergence of shared actions on toys
(de Barbaro et al., in prep).
Figure 3 Typical Four, Six, and Nine Month Sessions with Maternal Bids Indicated
Figure 3 Typical Four, Six, and Nine Month Sessions with Maternal Bids Indicated, Continued

As above, the y-axis indicates a specific partner (infant or mom) and sensorimotor modality (gaze, hands). Colored marks indicate, frame-by-frame, moments of sensorimotor contact with available targets. Contact with partner is always indicated in yellow; remaining colors indicate contact with objects. Widths and additional markings on the colored bands are not relevant for this study. The start and stop boundaries of all visible maternal bid negotiations are indicated, with the proportion of infant attending to the maternal bid toy marked proximally to each bid (see Methods Section). Note that modalities are increasingly distributed across more objects, either simultaneously or via rapid alternations, across first year, as described in text.

In this paper we systematically test this hypothesis by organizing continuous records of dyadic interactions around a specific kind of attention-sharing opportunity: the mothers’ presentation of a novel toy, which we call a “maternal bid.” Maternal bids occur at all ages and at all ages infants respond to these by attending via gaze and manipulations. Although these bids were unscripted, they
occurred multiple times at each age for every infant. As prototypical “openers” for various kinds of triadic interactions, maternal provide a naturalistic unit though which to assess changes using continuous variables that capture the response to maternal bid across the whole first year (for similar methods in a different population, see Forster & Rodriguez, 2006).

Methods

Participants

26 mother-infant dyads participated when infants were within one week of four, six, nine and 12 months. All families were participants in a larger longitudinal project (Deák, PI, NSF SES-0527756). The MESA project included 40 mother infant pairs with two monthly assessments, once at home and once in the lab, from three months to nine months, and then at 12 months. The current work discusses findings from a subset of this project. Mothers were initially contacted for participation in MESA through play groups and later by word of mouth.

Materials and setting

Dyadic free play interactions were recorded in monthly sessions at the participants’ homes. The sessions were designed as controlled but naturalistic social situations, where the mother and the environment would induce a spontaneous and dynamic flow of events that would sometimes compete for infants’ attention. Families were visited monthly from four to nine months, and again at 12 months;
however, the coding for the current analysis is focused on four, six, nine and 12 month sessions.

At four, six, and nine months of age, infants were placed in modified Exersaucer© walkers, in order to control for differences in infants’ height, and postural stability (see Fogel, Messinger, Dickson, & Hsu, 1999). At each session, the dyad was provided with a set of three objects. The four month session included a colorful twisting caterpillar, a hard plastic musical toy with buttons, and a weeble. The six month session included a foam squishy soccer ball, a weeble, and a hard plastic musical triangle-ring. The nine month session included a foam squishy football, a weeble, and a hard plastic musical rattle. At the four through nine month sessions, there we fashioned wells on either side of the walker, within reach of both the mother and the infant, to “store” toys not “in play” (see procedure below).

At 12 months of age, the session was conducted somewhat differently in order to accommodate the infants’ developmental changes, and thus maintain more natural interactions. At 12 months, infants were too large and independent to tolerate the walker: in pilot testing they frequently become fussy or distressed in the walker, whereas they were content to sit freely on the floor and play. Thus, mother and infant were seated on the floor, in the same location and at the same distance as in the previous session. This age-appropriate transition in seating maintained constancy of the emotional context and naturalness of the interaction. Moreover,
infants at 12 months no longer require postural support and could control their reaches more precisely, so the “motor scaffolding” provided by the walker was no longer appropriate. Second, because infants at this age more quickly lost interest in very simple toys, the dyad was given somewhat more complex toys, including a set of four soft, colorful silicone blocks, two wooden ladybugs, and colorful plaster rings linked to form a chain. The toys were distributed on the floor between the dyads, roughly as they had been on the tray in earlier sessions. Importantly, the infants’ longer arms and greater motor control now made the objects as reachable as the toys had been in earlier sessions. The infant could also easily look at and between any of the toys or the mother, as in earlier sessions.

At each session, three cameras recorded the interaction: one pointed at each participant’s face and upper body, and one further away and to the side, for a broader contextual view of the dyad. The angles and distances between the participants and the cameras were carefully matched across subjects and sessions in order to maintain video quality. Photographs and measurements of each family’s play room, taken at an earlier session, were used to duplicate the lay-out in future sessions.

The videos from each camera were later captured and synchronized in post-production. A screen shot example of the resulting synchronized three-view image used for coding is shown in Figure 4.
Procedure

At each session, one experimenter set up cameras while the other obtained informed consent. After the mother and infant were seated, they engaged in a six to seven minute free play interaction. Other measures were carried out following free-play, and will be described elsewhere.

All mothers were instructed to “play as they normally would” with their infants, using the objects as they felt fit. At the four through nine month sessions, mothers were asked to play with one object at a time, and told that the wells at the
side of the tray could be used to keep the other objects accessible. At 12 months the toys were left on the ground and mothers were not told to focus on single toys (however, the mothers had been so instructed at seven previous sessions). Despite the differences in the setup and instructions at the 12 month session compared to the earlier sessions, at all ages all objects were simultaneously accessible, both visually and manually, to both mother and infant.

Observational analysis

In order to identify behavioral differences in younger and older infant-mother dyads, the first two authors engaged in extensive analysis of videos from a subset of the infants, as detailed in de Barbaro, Johnson, and Deák (in preparation). This prolonged analysis during initial quantitative coding efforts eventually lead to the coding schemes described below, and provided hypotheses for analyses (Gnisci, Bakeman, & Quera, 2008) that were not well specified in the relatively impoverished behavioral literature on the emergence of triadic social interactions.

Systematic coding

**Coding Procedures.** Videos were coded by trained undergraduate students who were unaware of the hypotheses of the project. For these analyses the coders systematically coded the first 2-4 minutes of free play interaction for each dyad at each session. Coders viewed the three synchronized video angles simultaneously, frame by frame (10/sec), and marked events using Mangold Interact©. Coders were
trained to reliability (using Cohen’s Kappa criteria) for each variable. During this process we referred extensively to visualizations that allowed us to identify moments of match and mismatch between coders. 15-20% of the sample was coded by a second coder for reliability.

Coding Schemes. All sessions were coded for the targets of infant’s visual and haptic attention, as well as mother’s haptic attention. We also coded episodes of imitation and games.

Infant Gaze. Infant gaze was coded for the following targets categories. Visual contact was coded no matter how brief, and a code was terminated when contact between gaze direction and the target lapsed. Because infants could have multiple targets in view if they were very close together, targets were not mutually exclusive. (Kappa for this measure was K= .77)

A. Objects: Each individual object was specified.

B. Partner’s Face: All direct looks to the head area were included.

C. Partner’s Hand: gaze to the left and/or right hand below the wrist (looks to the arms were included in the category “Partner’s Body,” below). Also, when the partner was holding an object, that object was also included as a gaze target.

D. Partner’s Body included any part of the body except hands or face/head.
E. Previous location: This was coded when infant continued to gaze at the location where an object had just been, typically right after the mother removed it.

F. Other: All other locations, including the tray, the floor (and objects that had fallen there), and additional objects and locations in the environment behind and to the sides of the participants were included.

Manual Contact. Manual actions on objects were coded no matter how brief, and a code was terminated when contact between a hand surface and the object had lapsed for more than 1 second (so that repetitive stroking or tapping actions were not coded as many single instances of contact). We did not specify contact between right and left hands. We also did not differentiate between occasions when a single object was contacted with one hand versus two hands. More specifically, each frame for each participant was classified using one of the following codes. (Infant manual contact Kappa = .83; Maternal manual contact Kappa = .88)

A. None: Absence of activity of either hand to any object

B. Object: This included any kind of touch to an object by any part of the hand, not just grasping. This was because we observed a wide range of seemingly exploratory movements when in contact with objects, and we made the simplifying assumption that all touching actions would all yield information about the objects.
C. Multi-object contact. This described any time when in the participant had manual contact with more than one object (typically—two objects, one in each hand). The identity of each object was specified.

Imitation and games. A high level code of “imitation/games” was coded whenever the infants’ actions replicated or extended/expanded the mothers’ action. Coders noted the presence or absence of imitation and games for predefined event boundaries of three seconds each, following previous studies on macro level communication patterns (e.g., Hsu & Fogel, 2003). (Kappa= .88)

Results

We will first look at attention distribution between during imitation/games only (“Imitation and games analysis”) that occur at nine and mainly 12 months. Next, we will look at the patterns of attention distribution at all ages, as the infant responds to maternal bids (“Maternal bids analysis”). Finally we will explore the contributions of one particular developmental trajectory – sensory motor decoupling- as it contributes to the changing patterns of distribution to maternal bids toy and other toys.

All results were calculated via algorithms that analyzed the various dimensions of coded data described above. These algorithms were used to derive behavioral codes, such as the start and end of “bid” episodes, as well as and also to analyze specific features of interest of the interaction. All algorithms were written in
Matlab by the first author. Due to the novelty of analyses used, validity of the algorithms was investigated by visualizations of the original sample of five infants.

Imitation and Games analysis

Count. We observed sixteen episodes of imitation/games at nine months (distributed over nine dyads) and 85 episodes at 12 months (distributed over 20 dyads, from 1 to 6 episodes per dyad). Of these, four episodes at nine months and 67 at 12 months fit the criteria that we used for maternal bids, that is, that the objects manipulated by mom were not manually contacted by either participant for five seconds prior to start of the imitation episode. This subset of imitation and games was used in the following analyses.

Duration. The median of duration of imitation and games episodes was 8.1 seconds. The range of durations was 3-23 seconds and the distribution was skewed to the left, such that there were relatively more short episodes.

Proportion of Attending to Toy manipulated by mother during imitation. Using the subset of imitation/games identified we calculated the proportion of infant attending to the toys manipulated by mother —“Maternal-Manipulated Toys” (MMT)—relative to the other toys engaged in the imitation episode (i.e. those solely manipulated by the infant). This analysis was performed on each individual episode of imitation for each dyad at each session. For each episode, we used the
micro level coding to identify those toys during the imitation episode that were manipulated by the mother. For all MMT and non MMT objects that had any sensory motor contact by the infant over the course of an imitation/ game episode, we created a timeseries that specified how many modalities the infant had on that object at every frame. This timeseries was a continuous range of values from zero (no modalities in contact with the object) to three (gaze, right hand, and left hand simultaneously in contact with object). Thus the timeseries indicated the amount of sensory contact that an infant directed towards an object at each moment of the imitation/games episode. Summing over the values (or, identically, taking the area below the curve) of each timeseries for each object during the course of the episode of imitation gave us a measure of how much attending occurred to each object over the period of the imitation/games episode. Using these values, for each bid toy we calculated the amount of attention to the MMTs relative to the sum of attending to all other toys contacted during the bid negotiation. These proportions allowed us to see how the modalities were distributed during the episode of imitation/games. Again, these values were calculated for each MTT for each episode of imitation and games, occurring in each session, of each dyad.

From the proportion of attending to the MMTs, we created a histogram of all imitation episodes (See Figure 5a). This histogram indicates that in the majority of identified episodes of imitation/games, infants attended to the MMT 20% or less
(and all other toys 80% or more) of sensory motor contact for that episode. In 26% of the episodes the infants attended to the MMT 10% or less. The remaining episodes were distributed relatively evenly across the remainder of the bins. Thus for the majority of imitation episodes, infants attended to the maternal bid toys with relatively low proportions relative to their attending to other toys. Qualitative examination of such low proportions indicates that they correspond to infants using visual contact to occasionally “check in” on the actions their mothers were producing as they manipulate the toys in their own hands.

a.

![Figure 5 Maternal Bids Histograms, Full Sample](image)

Figure 5 Maternal Bids Histograms, Full Sample
The histogram in Figure 5(a) shows proportion of all imitation episodes (across all dyads and all sessions) with given amount of attending to MTT relative to other engaged toys during imitation episode. Histograms in Figure 5(b) show proportion of total session bids (all dyads, all bids) with given amount of attending to bid toy during bid negotiation. Session four, six, nine and 12 are represented in order from left to right, top to bottom. For each histogram, the x-axis delineates the boundaries for the histogram bins. From left to right, the bins represent increasingly high proportions of attending to the bid toy relative to the previous toy(s). The y-axis identifies what proportion of the total bids made during each session (summing over all dyads) fall within that bid. For example the two rightmost bins identify the proportion of total bids where infants put 100% of attentional streams on the novel toy during the bid negotiation, and the proportion of total bids where infants put between [90-100] percent of attentional streams on the novel toy during the bid negotiation. Again, each bid is represented once, such that each dyad contributes anywhere between 0-10 bids to the total for that session. Amount of attention is a summation of gaze and manual contact to given toy over the course of the bid negotiation.
Maternal Bids analysis

While imitation and games occur only at the end of the first year, mothers manipulate toys from the earliest sessions, and infants attend to these manipulations from the earliest sessions. In our next analysis we capture changes in the infants’ responses to the maternal presentation of a novel toy, what we term “maternal bids”. More specifically, we tracked the amount of infant attention to the “maternal bid toy” relative to “other engaged toys” toys in the period of negotiation. In doing so we create a continuous measure of mother-infant-object coordination that spans the first year. Instead of “apples to oranges” comparisons between macro-level states, a continuous measure provides a way of tracking progress in a single feature of changing attention relevant to imitation and games.

To conduct this analysis we first identified “maternal bids”. A bid began when the mother manually contacted an object that neither mother nor infant has manually contacted for the previous five seconds. This object will henceforth be referred to as the “Bid Toy”. Next, we identified the set of objects that the infant had been in contact with (via gaze or manual contact) anytime from 5 sec before the mother touched the new bid object to 0.5 sec after the mother touched it. Using these we calculated a window of time in which the infant distributed its attention to both the bid object and the previous objects: the “negotiation” window, as the infant negotiated contact between the mother’s bid and the other objects of ongoing
attention. Thus the duration of these windows was a function of infants’ attention in response to each individual bid.

The negotiation window began with the first frame that the infant contacted the bid object via any modality. Thus only bids to which the infant responded with at least one frame of sensory contact to the bid toy were considered for the analyses. The window continued as long as the infant maintained “active engagement” with the novel object as well as at least one of the previous objects. We operationalized active engagement as any contact via any sensory stream (gaze, manual contact) for at least one frame every 5 sec. The negotiation window was defined as ending 1 sec after the last frame that the infant contacted the last previous toy or the novel toy, whichever occurred first. (See Figure 6). Each parameter used for the calculations was adjustable and we tested at various combinations. All sets of parameters showed the same basic shift from earlier to later sessions, though some of the other sets showed additional significant differences between the consecutive pairs of months. The results we describe here are using a set of conservative parameter estimates that we think are empirically relevant.
Figure 6 Maternal Bids Identification Parameters

Figure 6 represents approximately 30 seconds of interaction between mother infant and toys in order to illustrate the parameters used in identifying bids and their start and stop times. X-axis represents time in the interaction. Areas on the y-axis correspond to sensory-motor modalities of attending, presence of a colored bar at a given time represents contact between given sensory motor modality and a specific target. The yellow target represents the partner (mother or infant), the remaining colors represent different toys. Descriptions of each of the parameters are given in main text. The square shaped box indicates the start and stop boundaries for the maternal bid to the pink toy, beginning approximately at 54 seconds and ending at approximately 65 seconds.

*Count Bids.* The number of bids for an individual dyad at each session ranged from 0 to 11 in the four through nine month sessions. The max at 12 month session was 30. Mean number of bids were 2.7, 2.6, 3, and 11.2 in each of the four sessions respectively.
Duration of Bid Negotiation. The duration of bids ranged from 1 to 45sec. Descriptive statistics are reported in Table 5. As durations were overall skewed to the left, we calculated the median bid duration for each infant for each session. We removed one outlier at 4months whose median value was 4.14 standard deviations over the mean of the median values (29.8sec median duration bid). Repeated measures analysis of variance (ANOVA) with a Greenhouse-Guisser correction identified a significant main effect of session on the duration of bid negotiations, (F(1.96, 37.33) =18.55, p =0). Post hoc pairwise testing with Bonferroni corrections revealed that there was no difference in durations between the four and 6 month sessions (p=1) or between nine and twelve month sessions (p=1). Results showed trending differences between four and 6 month sessions and nine and twelve month sessions (p= [.065, .13]). 12 month bids were significantly longer than four month sessions.

Table 5 Average Median Bid Durations

<table>
<thead>
<tr>
<th>Median Bid Durations</th>
<th>4mo</th>
<th>6mo</th>
<th>9mo</th>
<th>12mo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>2.52(2.45)</td>
<td>2.86(1.98)</td>
<td>8.68(7.53)</td>
<td>6.30(4.65)</td>
</tr>
</tbody>
</table>

Proportion of Attending to Maternal Bid Toy. For each bid, analysis of the distribution of attention was carried out as described in imitation section. For each session, we created a histogram of the proportions of time attending to the bid
object (over all modalities), for every bid in every session by every dyad. The histograms allow us to see overall changes across sessions in the distribution of the types of responses that the infants would make in response to a maternal bid (See Figure 5b).

Comparing across the sessions, results indicate that responses to maternal bids progressed with increasing age from a majority of bids receiving a high proportion of time attending to the novel toy relative to other toys (i.e., 99%+ time on at least 50% of bids), to a relatively flat distribution (at intermediate ages) where all proportions were equally likely, to a majority of bids, at older ages, being met with persistent attention to previous toys (i.e., 75-99% of time to the previous object on 50%+ of bids). Due to various distributions of data captured by the histograms we felt that classic descriptive measures (e.g. mean, median) did not adequately capture the change in time over sessions. For this reason we modeled the bid outcomes using a beta distribution, which is a two parameter distribution across (0, 1) which can take a variety of shapes (See Figure 7)\(^3\). These features make it advantageous over other, more common distributions (Cross, Herman, & Gaunt, 2006). Bestfit beta functions are indicated on Figure 5b. Statistically, we can assess longitudinal changes by

\(^3\) Note1. The beta distribution is calculated according to the raw values, rather than via the histogram bins. Note 2. Beta distributions are defined over the range (0,1) which does not include the values 0 and 1. Due to our criterion that infants attend to the bid toy via any modality at for at least one frame, there were no bids with 0% to the bid toy, however, there were two values (out of 554 total) of 100% to the bid toy (in particular at 4 months). We adjusted these values by 0.0135. Otherwise the largest non-“one” value was .978.
comparing the 95% confidence intervals for the alpha and beta values of the best-fit models for each session (See Table 2). Across all bids of infants, confidence intervals of betas between 4 and 9 months do not overlap. The confidence interval for 6 month betas spans the values between 4 and 9 months, overlapping approximately 50% with each, indicating a transitional period. Beta values are also significantly different between 9 and 12 months. These data indicate that in response to maternal bids, older infants show increasingly lower attending to bid toy relative to other engaged toys. Alpha values overlap across all ages. In the next analyses we see how these patterns of responding to the bid change based on individual differences in infant sensory motor coupling.

![Image of Beta Distribution Range](image_url)

**Figure 7 Beta Distribution Range**

Figure 7 indicates a number of different shapes taken on by the beta distribution as a function of varying alpha and beta parameters (each color). X-axis shows the range of values over which beta distribution is valid, the y-axis shows the probability density function at each point on the x axis. There are two parameters which describe a beta distribution, alpha, and beta. As alpha nears beta, the distribution becomes more flat, (though note changes at higher or lower values) as one parameter becomes small relative to the other, the distribution becomes increasingly skewed to the right or left.
Table 6 Best-fit Beta Parameters for Response to Maternal Bid

<table>
<thead>
<tr>
<th></th>
<th>Four Mo</th>
<th>Six Mo</th>
<th>Nine Mo</th>
<th>Twelve Mo</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>All bids</strong></td>
<td>1.06(0.68, 1.66)</td>
<td>0.95 (0.56, 1.61)</td>
<td>1.03 (0.57, 1.86)</td>
<td>0.82 (0.67, 1)</td>
</tr>
<tr>
<td><strong>All dyads</strong></td>
<td>0.37(0.23, 0.60)</td>
<td>0.61 (0.42, 0.88)</td>
<td>0.89 (0.67, 1.19)</td>
<td>1.73(1.58, 1.9)</td>
</tr>
<tr>
<td><strong>6 Mo Im G-H Coup</strong></td>
<td>0.95 (0.53, 1.69)</td>
<td>1.2 (0.52, 2.76)</td>
<td>0.95 (0.39, 2.29)</td>
<td>0.78 (0.6, 1.01)</td>
</tr>
<tr>
<td></td>
<td>0.39(0.22, 0.68)</td>
<td>0.5 (0.25, 0.98)</td>
<td>0.76 (0.48, 1.19)</td>
<td>1.8 (1.6, 2.02)</td>
</tr>
<tr>
<td><strong>6 Mo Mat G-H Coup</strong></td>
<td>1.36(0.65, 2.84)</td>
<td>0.94 (0.49, 1.79)</td>
<td>1.14 (0.52, 2.49)</td>
<td>0.9 (0.64, 1.25)</td>
</tr>
<tr>
<td></td>
<td>0.36(0.14, 0.91)</td>
<td>0.88 (0.60, 1.3)</td>
<td>1.08 (0.75, 1.55)</td>
<td>1.67 (1.44, 1.95)</td>
</tr>
<tr>
<td><strong>9 Mo Im H-H Coup</strong></td>
<td>0.99(0.58, 1.67)</td>
<td>0.92 (0.51, 1.66)</td>
<td>0.99 (0.49, 2.01)</td>
<td>0.77 (0.59, 1)</td>
</tr>
<tr>
<td></td>
<td>0.37(0.21, 0.66)</td>
<td>0.73 (0.5, 1.06)</td>
<td>0.71 (0.47, 1.06)</td>
<td>1.43 (1.25, 1.62)</td>
</tr>
<tr>
<td><strong>9 Mo Mat H-H Coup</strong></td>
<td>1.23(0.43, 3.49)</td>
<td>1.47 (0.41, 5.21)</td>
<td>1.87 (0.97, 3.61)</td>
<td>1.06 (0.81, 1.39)</td>
</tr>
<tr>
<td></td>
<td>0.37 (0.13, 1.1)</td>
<td>0.49 (0.18, 1.35)</td>
<td>2.91 (2.29, 3.7)</td>
<td>3.17 (2.8, 3.59)</td>
</tr>
</tbody>
</table>

Best-fit parameters for beta distribution of Response to Maternal Bid of full sample and split groups based on coupling values. Alpha and beta values are provided as well as 95% confidence interval following in parentheses. See also Figure 11 for graphical representation of the intervals. Coup= Coupling; Im = Immature; Mat = Mature.
Sensory motor coupling

Coupling indicates the degree to which infant’s sensory modalities functioned in unison, that is, the degree to which they were directed towards a single target vs. multiple targets (i.e. “decoupled” modalities). We assessed the degree of coupling between two pairs of infant sensory-motor modalities, 1) gaze and hands (GH) 2) the infants’ two hands (HH).

Coupling between the HH was calculated by first summing the number of frames in which hands were active, that is, when at least one toy was being manually contacted. Next, we calculated the number of frames that two different toys were being simultaneously manually contacted.

GH coupling was slightly more complex given that there were up to three streams (gaze, right hand, and left hand) that could be coupled or decoupled at any moment. Thus for each object we first counted the number of eligible frames, that is, frames in which 1) both gaze and hands were “engaged”, i.e. in contact with any of the objects available during the session and 2) at least one of the two modalities was engaged specifically with the current object. For example, while calculating proportion match for right and left hand for the animate object, eligible frames included, for example, frames where infant had both hands on the object of interest, one hand on the object of interest and one hand on another, secondary, object, or
both hands on secondary objects. We did not include contact with maternal hands as a possible target within this analysis.

From these eligible frames we calculated the proportion of “coupled” frames, in which gaze and at least one hand were simultaneously in contact with the object of interest. Decoupled frames were frames in which gaze was in contact with the object of interest but hands were not, or vice versa. As infants might be holding two different toys, each in a separate hand (e.g. blue in the right hand and red in the left hand) and gazing to only one of them (e.g. to the blue toy), a single frame could contribute both “coupled” and “decoupled” data points. The denominator was calculated by summing the eligible frames for each object. The numerator was calculated by summing the coupled frames for each object. High values indicated larger proportion of multimodal attending match, where both modalities were on the same object; low values indicate low proportion match and high proportion mismatch. Figure 8 indicates the proportion coupling for both GH and HH pairs.
X-axis indicates four sessions; y-axis indicates proportion of session that these modalities were directed towards the same target. High values indicate increased proportion of session that modalities were coupled, low values indicate modalities were decoupled or on different toys for increasing proportions of the session. Significance ratings are at p<.05 and were obtained via repeated measures ANOVA (see text).

We used repeated measures ANOVA to assess the change across sessions in coupling. For GH coupling data, all but the nine month data had skewness and kurtosis values between 1 and -1. At nine months, gaze hands showed skewness of 1.55(SE= .46) and kurtosis of 2.5 (SE =.90). Given a greenhouse-Geisser correction for violation of sphericity, the main effect of the sessions was statistically significant (F (1.96, 37.33) =18.55, p =0), indicating that infants showed decreased GH coupling over the course of the year. Post hoc pairwise testing using a Bonferroni correction revealed that there was no difference in coupling between the four and 6 month sessions (p=.65) or between nine and twelve month sessions (p=.77). However, four
and 6 month sessions were significantly different from nine and twelve month sessions (p<.003).

Repeated measures ANOVA of HH coupling data also indicated a main effect of session was statistically significant (F (2.3, 53.5) =21.98, p =0). Post hoc testing revealed that there was no difference in hands coupling between the four and 6 month sessions (p=1) but that infants showed more hands decoupling between 6 and nine months (p=.03) and between nine and twelve months (p=.04).

We also assessed the relationship between GH coupling and HH coupling from month to month. GH at four months was positively correlated with GH at six months (r= .47, p=.01). There were no significant correlations between HH coupling across sessions. There were relatively small but positive correlations between gaze hands coupling and HH coupling values from month to month, reaching significance only at the nine month session.

Relations between response to bids and decoupling

Our hypothesis is that the degree of decoupling between pairs of sensory motor modalities contributes to infants’ responses in certain kinds of dyadic interactions, such as their responses to maternal bids. Specifically, as the infant decouples more, the more his response to the bid begins to look like those seen during episodes of imitation and games, where the infant distributes attentional contact between maternal bid toys as well as another set of toys. Moreover, as he
decouples more, the more he attends to bid toys with a relatively low proportion allowing him to visually “check in” to his mother’s manipulations while still devoting much of his attentional resources to coordinating gaze and hands to toys in his own possession as he reproduces those manipulations on his own toys. The analyses above show that overall, decoupling increases through the months as infants’ response to maternal bids becomes increasingly mature. The following analyses test for a more stringent criteria: that decoupling is related to response to bids at a more disaggregate level. As the typical number of bids per session was relatively low, it was not possible to model individual data reliably, as we did with the beta distribution above. However, we pursued disaggregate analyses to investigate this relations between decoupling and the response to maternal bids using two methods: a “split-halves” analysis, and a regression model.

*Split group analyses.* We identified subgroups of babies based on their decoupling scores and compared the beta distributions for the response to maternal bids between the groups. This allows us to assess whether differences in relative amounts of decoupling at a given month lead to differences in response to maternal bids.4

4 As a more stringent test of the hypothesis in the following analyses we also tested measures of decoupling that only considered moments of the interaction outside of the maternal bids. That is, activity within the maternal bids did not contribute to these decoupling scores. Differences between two versions were not significant.
**GH decoupling.** We divided GH coupling data at the 50th percentile of infants at six months, creating two groups, those infants that decoupled more, or “Mature” GH decouplers and those infants that decoupled less, or “Immature” GH decouplers. Next we used best-fit beta distributions to model the distribution of novel-toy attention per bid, for mature versus immature GH decouplers (see Figure 9). Note that MB responses at each session have been split according to the infant’s 6 months decoupling scores, rather than the decoupling score from each session. Confidence intervals indicate that betas of 6mo immature GH decouplers are statistically indistinguishable from all dyads’ 4 mo betas, and that those betas of 6mo mature gaze hands decouplers are non overlapping with 4mo dyads betas but instead statistically more similar to all dyads 9 mo dyad betas. Alphas are not significantly different across any of the groups between any ages. Thus infants who are less proficient at GH decoupling at 6 months respond to maternal bids like the infants as a group at 4months, whereas infants who did more decoupling at 6 months responded to maternal bids more like the infants as a whole at 9 months. Moreover, by 9 months, the response of these proficient 6 month decouplers bids become statistically indistinguishable from all babies responses at 12 months, while the more immature babies responses confidence intervals still overlap with the full sample at four months (see also Figure 11).
Axes are included only as HH decouplers. As in the GH analysis, we used best-fit beta distributions to model the distribution of proportion attending to the novel toy at each session for the two groups (see Figure 10). Confidence intervals indicate that betas of nine mo immature HH decouplers are statistically indistinguishable from both overall four and six mo dyad betas, and that those of nine mo mature hands decouplers are statistically indistinguishable from overall 12 mo dyad betas, but highly different from the four, six and nine month beta parameter of the overall group. Again, alphas are not

Figure 9 Maternal Bids Histograms, Immature and Mature GH Couplers

Axess are identical to those described in Figure 5. Four leftmost Figure 9 histograms include only data from those infants identified as immature decouplers based on their six month data. Four rightmost Figure 9 histograms are based on data from infants identified as mature decouplers based on their six month data.
significantly different across any of the groups between any ages. Thus infants who are less proficient at HH decoupling at nine months respond to maternal bids in a manner similar to infants in general at six months. Infants who are more proficient at HH decoupling at nine months respond to maternal bids more like infants as a group at 12 months.

![Immature 9mo H-H Couplers vs Mature 9mo H-H Couplers](image)

Figure 10 Maternal Bids Histograms, Immature and Mature HH Couplers

Axes are identical to those described in Figure 5. Four leftmost Figure 10 histograms include only data from those infants identified as immature hand-hand decouplers based on their nine month data. Four rightmost Figure 10 histograms are based on data from infants identified as mature hand-hand decouplers based on their nine month data.
Confidence intervals for beta parameter of beta distribution for each month, for mature infants (top), whole sample (middle), and immature infants (bottom) babies. Months are indicated by color, with 4 months at the top and 12 months at the bottom. The left hand side of Figure 11 shows confidence intervals for babies based on GH coupling at 6 mo, the righthand side shows confidence intervals for babies based on HH decoupling at 9 mo. Note that the full sample is represented twice for each month, once on the right, and once on the left. Alpha parameters are not represented as there was no significant variability across any pair of months given any sample or subsample.

**Regression analyses.** We further tested whether disaggregated changes in infant decoupling predict changes in response to the maternal bid by regression analyses. For the purposes of the regression, the bid is the unit of analysis. Each dyad at each session contributes as many bids as occurred in that session. Our model uses GH decoupling values and HH decoupling values to predict proportion of attending to the bid toy. Because decoupling values were calculated as a single index of decoupling over the entire session, each of the bids that were contributed by a given
dyad at a given session had the same GH decoupling value, and the same HH decoupling value. The proportion of attention to the novel toy was calculated for each bid, as described above. Coupling values explained a significant proportion of variance in the proportion attending to the bid toy, $R^2 = .35$, $F (2, 547) = 148.04$, $p = .00$. Standardized coefficients indicate that increases in GH coupling translated to increasingly higher attending to the novel toy, $b = .39$, $p < .001$. Likewise, increases in HH coupling translated to increasingly higher attending to the novel toy, $b = .28$, $p < .001$. Thus, changes in both types of decoupling are linearly related to changes in proportion attending to the novel toy, such that as coupling decreases (or decoupling increases), infants put relatively less and less attention to the maternal bid toy.

Discussion

Previous theories on the development of triadic attention are based on high-level or macro-level codes that specify only the presence of absence of games or imitation. Instead, in this study we create a continuous measure of one component of triadic attention: the infants’ response to the mothers’ presentation of a novel toy, what we termed a “maternal bid.” Our histograms can be considered probability estimates of how likely the infant is to distribute his attention to the maternal bid toy relative to the other toys. Analysis of changes in infants’ response to maternal bids over the first year reveals a previously unreported trajectory that, we argue, provides a new window into the development of shared actions on toys.
At the youngest age, 4 months, infants’ modalities were highly coupled, with 50% of infants maintaining coupled gaze and manual activity throughout more than 80% of the session. At the dyadic level, this manifests as the infant transitioning to whatever toy the mother manipulates in front of the infant. This in combination with the short duration of the majority of their bids indicates they do not distribute their attention between both objects but rather get “stuck” on a single toy, causing the period of negotiation to resolve quickly. This recalls to us a very unstable dynamical system which where a single step towards one toy leads to a tumbling effect “down the rabbit hole” to full multimodal contact with the toy.

By six months, the dyadic outcomes were solidly between what they were at four months and what they would be at nine months. Instead of switching over all attention to the mother’s bid, six month infants increasingly distributed their attention between the maternal bid toy and the previous toys. Analyses indicated that only 45% ended with the previous toys being “dropped”, nearly 50% less than the four month babies. Moreover, those babies who spent high proportions of the session with gaze and hands on the same target were twice as likely to “fall down the rabbit hole” than their peers. However six month dyads still showed some important differences from older dyads. As most of the 6month infants (21 of 26) showed high proportions of the session with two hands coupled it follows that attention was distributed mostly via simultaneously gazing to one toy while holding another, or by
alternating (sequentially) their hands between the bid toy and the other toy(s). In later sessions, infants would also distribute attention via two hands. Regression models indicate that this allowed them to maintain even smaller contact attending to their mother’s toys. The transitioning status of six month infants was also reflected in the beta models, which indicate that the distribution of responses at six months overlapped with both four month and nine month sessions. Additionally, the duration of bids tended to be shorter at six months than at the older months, with the median duration of six month bids less than half of the duration of those at nine and 12 months. These differences in duration trended significant. While six month infants are beginning to distribute their attention between their mothers' bid and their other toys, this pattern is unstable relative to that seen in older dyads.

The nine month group’s responses to the maternal bid were further shifted towards the 12 month distribution of attending between the maternal bid toy and the other toys. In the majority (60%) of bids made at nine months, infants attended to the mother’s objects only 10-30% of the negotiation period; the rest of bids at this session were relatively equally distributed. Such low proportions of attending to the maternal bid toy were also the most prevalent at 12 months, as well as during episodes of imitation. This change was reflected in our statistical models, as the beta parameters of the distribution of responses to the maternal bid toy of nine month dyads overall no longer showed overlap with the four month dyads. Moreover, dyads
with infants who showed high hands decoupling at this session showed a distribution of attending to the bid toy statistically equivalent to that of the 12 month dyads. Finally, the trend toward longer durations of the bids at 9 months indicate that distributing attention between mothers’ bid toy and ongoing objects was a more stable configuration of attending than it was in earlier sessions.

Regression analyses confirm the role of changing gaze hands and hands decoupling in bid outcomes. Across all bids, as infant decoupling increased in a session, we also observed increasingly low contact attending to novel toy during bid negotiations. We suggest that as the infants decouple more they become better able to dance on the ridge between the basins of two toys and not fall into either– something that seems important for shared actions on toys. In particular, we suggest that this behavior allows the infant to distribute attention between a bid toy and the objects of ongoing attention in such a way that the infant can incorporate some elements of the mother’s manipulation into his own play, as in an episode of imitation or games.

We had a number of other findings. First, we did not find strong correlations between gaze hands coupling and hands coupling. This suggests that these are really independent motor systems with their own developmental trajectories for organizing action. Next, correlations indicate that the relative amount of decoupling between one session and the next are not related. This pattern could occur if shifts in rates of
decoupling occurred at different times longitudinally across infants. The one exception was gaze hands coupling between four and six months showed moderate (.47) positive correlations.

We identified a number of limitations to our study. First, the time consuming nature of the coding limited our sample size as well as the duration of interaction we coded. This led to challenges in studying individual differences. Given an increased number of bids per session per dyad we could model individual level effects to control for the possibility that a small group of infants is driving the effects we observed. We attempted to minimize false results by using multiple converging methods for looking at the relationship between decoupling and response to maternal bid. Next, there were a number of differences between the four through nine month sessions and the 12 month session. While these adjustments were necessary for the infants’ comfort and participation, there were important changes that need to be considered when interpreting changes in data between nine and 12 month data.

Importantly, there are a number of ways in which the differences might account for the shift to lower proportions of attending to the bid toy between nine and 12 months. First, there were seven toys to attend to at 12 months and only three in the earlier sessions. Occasionally the infant would gaze to a pile of many toys at once, which could have greatly affected the proportion of attending to the maternal
bid toy. Next, at the four through nine month sessions but not at the 12 month session, the mothers were encouraged to allow only “one object on the tray”. We often observed mothers at 12 month sessions make bids for infants’ attention using two toys simultaneously (one in each hand), whereas this was rarely seen in the earlier sessions. This could further lead to increased likelihood of 12 month infants gazing to a higher number of toys simultaneously, in turn reducing the amount of attending to the maternal bid toys. We intentionally operationalized the measures so as to minimize the impact of this difference. Specifically, we set the value of gaze to maximize at 1 “point” per frame, such that gazing to additional toys would not affect the relative proportions between the gaze and hands. However, as there were only maximally three toys to gaze at in the earlier sessions, this could have reduced attending to the novel toy between nine and 12 months.

Next, notwithstanding our instructions to keep only one toy on the tray, nine month infants often held two toys, one in each hand. However, 12 month infants did so more often. This may have been due to mothers “removing” additional toys at 9 months in attempt to obey the instructions of the experimenter. Given that we operationalized each hand was always worth one, it is not likely that this would contribute to changes in proportion attending to the novel toy between nine and 12 months. It could however contribute to increases in decoupling values between nine and 12 months.
Fortuitously, many of the developmental trends were well established across the period four through nine months where there were no differences between sessions, including decreases in GH and HH coupling as well as the shift to increasingly low MB-contact bids (when considering the mature 9mo HH decouplers). Thus although the shift to a different setting at 12 months may have affected the raw values of some of our measures, this leads us to be confident that pattern of results we observed is valid.

Finally, while we focus in these analyses on the contributions of sensory motor behaviors to changes in mother-infant-object attending, many other factors contribute to the response to the maternal bid. Using our high density coding we are able to identify many factors of import. For example, the mother physically removing the previous toy from the workspace can increase the likelihood of looking to the novel toy, as can the amount of time infant spent looking to the toy before the mother made a bid upon it. When we include such additional factors as predictors of the proportion of attending to the bid toy, we can explain twice as much variance as when we consider sensory motor coupling alone. Other factors we cannot currently account for given our coding scheme. For example in our observational analyses we identified changes in the maternal bid style that were systematically related to age, such as the proximity with which the mother manipulated the toy to the infant’s face and hands (closer at four month sessions) (de Barbaro et al., in prep).
Additionally, the current work focuses on the dynamics of infant’s engagement in response to a single action by mother. In future work we plan to explore more temporally precise dyadic sequential patterns.

Conclusions

Overall, our pattern of results suggests that development in relatively low level sensory motor coordination is an important factor in the dyadic level engagement between infant and mom. This provides a novel picture of the developmental transition to imitation and games. Where previous accounts identify a surprising shift to a novel dyadic behavior, our methods reveal continuous changes in an important feature of mother-infant-object coordination.

Our analyses indicate changes in the infants’ attention distribution to actions, or bids, made by the mother. We observed that the infant transitions from taking up the mother’s bids with very high proportion of attentional streams to taking them up with a relatively small proportion of his attentional streams, allowing the infant maintain ongoing attention to a second set of toys in his own possession. It is this distribution of attention that allows the infant to visually check in on the actions of his mother while simultaneously coordinating gaze and manual actions necessary to reproduce those observed actions. Further, we observed that changes across the first year in mother-infant-object coordination are related to infant sensory motor decoupling. A dynamical systems metaphor is a useful tool for summarizing this
developmental trajectory. We consider each toy as a potential basin for attention, and that the maternal bid increases the likelihood that an infant will descend into the valley of a particular toy. At the earliest sessions, infants attended to one toy at a time. Thus, each toy can be represented as a deep basin with steep sides. As the infant attends to a novel object manipulated by his mother, this drives the infant’s exploratory energy out of the current basin into a new local minimum. At four months, when this happens, the prior basin (represented by some residual activation of a conjunction of location and feature information) quickly returns to baseline level. By 9 months, when infants can distribute attention between multiple toys, infants are not in the same danger of falling down one basin or the other but can rather effortlessly walk the line (or the “saddle”) between multiple basins, attending to both toys but not succumbing to one or the other. In future work, we can model these processes using dynamic field theory models (Simmering et al., 2008).

At the least, the work we describe here indicates that an internal shift in representing intentions of others is not sufficient for shared actions to develop. At the most, it is one piece of an alternative account of triadic attention that is grounded in continuous changes in the activity of our participants rather than a simple shift in internal structures. By focusing on changing processes of attentional behaviors across the first year, we identify a factor which, in combination with other
changes in the first year (see de Barbaro et al, in prep) leads to the emergence of shared actions on toys at the end of the first year.

Motor changes in coupling between gaze and hands have been described in developmental psychology literature (Bushnell, 1985; Corbetta & Thelen, 1996; Flament, 1975; White, Castle, & Held, 1964). However they have not been linked to changes in the dyadic interactions as we have done here. By considering the social interaction as a configuration of sensory motor attending between the mother and infant we identified an element that was critical to it. Simply put, in order to imitate the mother’s actions on a set of toys, the infant must distribute its attention between toys in his own possession and toys within his mother’s possession. Even more simply, in order to attend to the mother’s actions on toys during ongoing toy play, the infant must either drop the original toys or distribute his attention between the toys currently held and the mother’s toys. Thus it is not surprising that changes in one method by which the infant distributes attention – that is, decoupling – affect the manner in which the infant can engage with maternal bids.

Aside from the particular results that we present, the methods we pursue in this paper, inspired by distributed and embodied cognition (Hutchins, 1995; Johnson, 2001), suggest novel ways to gain traction on previously intractable problems. A distributed cognitive account focuses on each of the partners’ access to the emergent products of the others’ activities in the world: their motions, words, affect,
through space and time. It is via our attention and responses to those emergent products that coordination occurs between individuals and their environments. This is a boon to a developmental account in that it focuses on the changing processes of cognition that are visible rather than invisible, and that are present from the earliest ages. By focusing on processes of interaction, rather than simply the products of interaction (e.g. high level behaviors such as imitation), we can identify changes in the process by which infants attend to and engage with their caregivers. Moreover it makes visible the processes of rich multimodal feedback that drive learning (e.g., Goldstein & Schwade, 2008).
References


Flament, F. (1975). *Coordination et prévalence manuelle chez le nourrisson:* Centre national de la recherche scientifique.


Chapter 3, in part, is currently being prepared for submission for publication of the material. De Barbaro, Kaya; Johnson, Christine M.; Forster, Deborah; Deak, Gedeon O. The dissertation author was the primary investigator and author of this material.
CONCLUSION

I introduced this dissertation with the goal of supplementing traditional developmental research methods and models, to move beyond theories that privilege internal structures as the drivers of social-cognitive development. In each chapter I primarily or exclusively discuss the results in terms of visible behavioral outcomes. The operationalized outcome measures were temporally dense, generalizable, and objectively replicable measures. In my first chapter, I describe patterns among heterogeneous measures of looking behaviors distinct from speed of encoding or rate of habituation. The results show that these measures can systematically characterize infant looking along dimensions that are generally unacknowledged in infancy research, but are well-grounded in the broader scientific literature on systems neuroscience and neuromodulation. In the study of triadic attention, the results could be characterized in terms of behavioral trajectories that encompass both definable perceptual-motor developmental trends, and a coherent set of caregiver behavioral changes and dyadic dynamics. These trajectories could be characterized in terms of more extended, macro-level events within the interactions, and in more detailed sequences of micro-behavioral synchronies and sequences. In all of these findings, we can describe the patterns of change without resorting to hypothetical, infant-internal, representationally high-level faculties.
We consider this approach of beginning explanatory efforts with well-grounded biological, physical, and social-contextual factors to be a conservative strategy. We do not have access to infants’ internal representations. This is not to say that representations of social events or contingencies don’t exist. To the contrary, to the extent that infants show systematically contingent behavioral decisions in response to specifiable social cues, we can infer that there has been some entrainment of cortical networks to abstract patterns of information (both external and internal). In the discussion we will consider our results as they relate to what is known about the neurological underpinnings of the processes described in the thesis.

An important epistemological claim of this thesis work is that good developmental research depends on capturing rich processes of behavior on both short and long timescales. This applies to research at both behavioral and neurobiological levels. Although the findings in this dissertation are restricted to behavioral data, it is possible to relate them to findings from neuroscience research. That is, neuroscience research suggest novel frameworks for understanding our results, and our work can provide novel insights about the kinds of social-behavioral processes that are tied to biological changes.

In the first chapter, our research methods and constructs were drawn from a theory of the neural modulation of attentional behavior, what we termed the “Aston-Jones Model of Attention” (AJMA). This model provides detailed characterization of
acute brain dynamics of attention, linking changes in the neuromodulator norepinephrine with moment to moment changes in the animals’ ecology and behavior. By contrast, the developmental research on infant attention has focused on stable processes of attention.

In the discussion for this chapter, we will describe the implications of this theory for developmental research. Specifically, if this theory applies to infants as well as non human animals and human adults, it provides a set of novel hypotheses linking visual attention to stress. In the discussion we will describe what is known about this pathway in the brain and then consider a number of previous findings from the developmental literature that may be mediated by these relations.

In the second and third chapters, we begin from a traditional developmental story based in a representational account of the development of imitation. We use methods that more precisely track the temporal dynamics of activity in order to get at the changing processes by which infants organize the sensory motor modalities of attention to their mothers and a set of objects. Our findings allow us to account for developmental change in social development without referring to the traditional representational account.

In this discussion I will introduce two sets of theories from cognitive neuroscience that are relevant to our developmental account. The first concerns the role of the mirror neuron system in the development of imitation or intersubjectivity;
the second regards the role of working memory in the development of turn taking and routines. I will suggest that without careful examination of the processes of development in a rich socio-behavioral ecology, neuroscience models can commit the same mistakes we sought to avoid. That is, our epistemological claim (Introduction) was that without documenting the structure of situated behavior, psychology researchers tend to privilege internal mechanisms as the engines of complex behavior. I will use insights from our findings to critique current neurobiological theories of social development, and to suggest a more interactionist, process-oriented account of the development of emerging social behaviors at the end of the first year.

In chapter one, we introduce the Aston-Jones Model of Attention to provide a novel, biologically grounded framework with which to consider the micro dynamics of infant attention. Data from chapter one indicates that we can reliably collect a variety of micro measures of infant visual behavior in rich naturalistic settings. Moreover, we found some evidence in support of AJMA that contradicted the traditional IP model. Namely, repeated stimulation led our “slow looking” infants to reduce their vigilant behaviors towards our stimuli, whereas our “fast looking” infants continued to be vigilant till the end of the session. This is counterintuitive to the traditional information processing model.
While there is still much work to be done before the claims of AJMA are fully tested and verified in an infant population, there are a number of implications of this framework for infant research that deserve further attention.

Within the AJMA framework, changes in rates of firing of norepinephrine and attentional vigilance are described to fluctuate as a function of levels of uncertainty in the environment. This idea helps to link conceptually the types of situations that lead to changes in NE firing and to changes in attentional behaviors (i.e., as uncertainty rises, the organism tends to seek more information). However, it is also useful to consider these changes as functions of stress and arousal systems rather than the abstract notion of uncertainty. We now detail these functional relations.

Biologically, the effects of stress on the central and peripheral nervous systems are mediated by two main pathways. The more commonly known to developmentalists is the hypothalamo-pituitary-adrenal (HPA) axis. Following a stressful episode, the hypothalamus releases corticotrophin releasing factor (CRF), which stimulates the release of corticotrophins from the pituitary glands, which in turn stimulate the release of cortisol from the adrenal glands (McEwen, 1998). This is a relatively slow-acting process, requiring about 20 minutes for serum cortisol levels to measurably change (Sapolsky, Romero, & Munck, 2000).
In addition to activating the HPA axis, stress-related events activate a second, faster-acting brain system. Specifically, CRF produced in areas outside of the hypothalamus, most notably by the central nucleus of the amygdala and the bed nucleus of the stria terminalis, activates various nodes of the sympathetic nervous system (Koob & LeMoal, 2001). Activation of the SNS elicits peripheral effects of the “fight of flight” state, such as increasing heart rate and slowing digestion in order to prepare for intensive activity. Additionally, the SNS impacts various central nervous system structures.

Specifically, increased SNS activity increases the involvement of subcortical structures such as the amygdala, while attenuating the involvement of the higher cortical structures in the prefrontal cortex (PFC; Arnsten, 1998). These changes are mediated by increases in PFC of the catecholamine neurotransmitters dopamine and norepinephrine. The PFC is acutely sensitive to varying amounts of these neurotransmitters; animal models show that increased catecholamine levels compromise performance on PFC dependant tasks, including working memory and focused attention (e.g. Arnsten & Goldman-Rakic, 1998). As stress-induced SNS changes prepare the body for fast action, the corresponding central changes shift the brain towards producing fast, automated responses rather than subtler, slower, frontally mediated actions. Moreover, they modulate the animal’s attention to
become highly stimulus dependent, privileging a tendency to quickly orient towards changes in the environment rather than prolonged attention towards focal stimuli.

Linking these changes in PFC function to stress has important implications for developmental research. While researchers have documented relationships between measures of stress and later developmental outcomes (Creasey & Jarvis, 1994; Feldman et al., 2004), for the most part we lack knowledge of the precise mechanisms by which these relations exist. The connections between sympathetic nervous system and cognitive functions suggest a possible direct mechanism linking stress to cognitive outcomes. Again, while we focused in our chapter on potential changes to attention, these links may affect many domains of cognitive development. While the research described focuses on acute effects, continued efforts investigating the role of arousal on cognitive processes could explore long term outcomes associated with chronic patterns of stress.

There has not been much work explicitly examining the relations between arousal and cognitive processes in infancy. However, there are some important indications from the previous literature that arousal does mediate looking behaviors. The most direct work linking arousal and attention in infancy comes from microdynamics of heart rate during infant looking tasks. Previous work has indicated a characteristic pattern of heart rate during the course of look, where heart rate decreases during the middle portion of a look to a stimulus (Colombo, Richman,
Shaddy, Follmer Greenhoot, & Maikranz, 2001; Richards & Casey, 1991). Additional research indicates that the degree of reduction of infants’ heart rate during this phase of looking corresponds to behavioral indicators of focused attention. Specifically, larger reductions in heart rate correspond to behavioral indices of focused attending, including reduced distractibility by peripheral stimuli, and facial features such as brow furrowing. Relatively shallower reductions correspond to features of more casual or non-focused looking, including faster reorientations to peripheral stimuli (“distractibility”; Lansink & Richards, 1997; Ruff & Saltarelli, 1993). The observed relations between these behaviors and the changes in peripheral sympathetic activity are consistent with those described by Aston-Jones and colleagues.

Other studies exist where physiological measures have not been collected but the behavioral data are consistent with AJMA. These are paradigms that have good potential for exploring relations between arousal and attention with infants. In particular, a large body of data suggests that infants look duration is based on an “optimal discrepancy” (Kagan, 2002, 2008). Given an object of “optimal” complexity, infants will spend less time looking to either a more complex object or a less complex object. This is contradictory to the traditional IP model as the IP model would suggest that the more complex the stimulus, the more time it would take for the infant to build a representation of that stimulus, and therefore, the longer that the infant
would continue to look to the stimulus. Under the AJMA framework, we would hypothesize that as complexity or novelty increases, arousal also increases. To a point, this shifts the infant into an interested and engaged state of attention. Optimal discrepancy is surpassed when the threshold of novelty is too high, leading to SNS affects on infant arousal and more vigilant scanning patterns which reduce the filter of attentiveness towards other stimuli, decreasing the amount of looking to the central stimulus.

In these studies, we hypothesize that the stimuli themselves may lead to changes in arousal that feedback to changes in looking behaviors. In the next set of studies, it appears that social interaction first impacts arousal, which then affects performance in a subsequent a non-social attention task. Using an experimental manipulation, Dunham and colleagues (1989) compared the effects of socially contingent interaction vs. random interaction on the amount of time infants subsequently activated a multimodal, rhythmic stimulus by fixating on a target light. Relative to the contingent social interaction condition, infants in the random condition avoided activating the stimulus for longer periods of time, and when they did look to the target light to activate the stimulus, they looked at it for less time. We speculate that these results are mediated by changes in arousal. Specifically, increasing levels of arousal of infants in the random interaction condition could have
led to more vigilant patterns of attention which would include decreased focus towards a single central stimulus, and therefore less activation of the target.

Again, these are speculative hypotheses. Future research could test for relations between arousal and attention by using physiological measurements such as saliva samples or heart rate. Alternatively, mild social stress manipulations such as the still face (Tronick, Als, Adamson, Wise, & Brazelton, 1979) could be used assess causal role of increasing arousal on attention. By taking more precise measurements of concurrent behavior and physiology we can begin to specify how these factors help to determine infant cognitive processes on both short and long timescales. This will provide a richer, more complex understanding of cognitive development.

Chapter 2 & 3

Chapters two and three carefully track the changing sensory motor coordination between the mother, the infant and a set of toys over the first year. From this data we create a novel account of the emergence of shared actions on toys. In particular, we identify a number of continuous developmental trajectories that we suggest slowly culminate in the presence of imitation and games at the end of the first year. These trajectories are very different from the much of the current research on the development of imitation or other turn taking shared actions such as games. In the following sections we explore a few of these current models.
Perhaps the most generative and contentious of modern theoretical concepts of imitation is the mirror neuron system. The system encodes mappings between visual observations of other-generated actions, and proprioceptive or motor encodings of self generated actions. Although there is virtually no behavioral data on MNS activity during mother infant interactions (but see Liao, Danly, Vankov, Makeig, & Deák, 2011; Marshall & Meltzoff, 2011), it is widely believed that that this system plays a role in imitation (Meltzoff & Decety, 2003; Ray & Heyes, 2011; Wohlschläger & Bekkering, 2002).

Whereas historically imitation has been attributed to “reading minds” or intentionality, the novel role ascribed to the MNS highlights the importance of attention to the micro dynamics of own and others’ actions as they unfold in time. To us this is a welcome, more precise framing of the nature of the phenomenon of imitation. Notably, this framing does not imply that imitation is a simple cognitive “problem.” Indeed, it is generally agreed that the process of parsing and translating dynamic motions into reproducible actions is quite complex, in particular given differences in orientation and scales between infants and their social partners (Meltzoff, 1993; Wohlschläger, Gattis, & Bekkering, 2003).

The contention lies in the ontological role of the MNS system. One model suggests that the MNS may provide a solution to this difficult problem of parsing actions—specifically, that a functioning MNS is *ontologically prior* to imitative
behaviors, or that mirror activity drives the parsing that is necessary for imitation (Meltzoff & Moore, 1997). The alternative model suggests that a functioning MNS is the *ontological product* of the developmental processes that lead to both biological and behavioral changes important for the production of imitative behaviors.

A principle proponent of the first model, Andrew Meltzoff, proposes that infants are born with a mechanism for associating the actions of others directly and specifically to their own actions (Meltzoff & Decety, 2003; Meltzoff & Moore, 1997). This claim rests on studies of newborn imitation. It is argued that if newborns imitate, then the MNS must precede and drive imitative behaviors, as there simply is not enough time or exposure for the MNS to develop through social interaction (Meltzoff & Moore, 1977). However, the initial studies of newborn imitation have not been replicated or have been shown only for one action, tongue protrusion. Another problem is that this is a common action that is elicited by many stimuli (Anisfeld, Masters, Jacobson, & Kagan, 1979). Moreover, a recent cross sectional study of many different actions to be imitated identified that imitation of tongue protrusion has been found to develop relatively late in the first year (Jones, 2009). For these and other reasons (Anisfeld, 2005), claims of neonatal imitation and have been challenged on both theoretical and empirical grounds (Jones, 1996; Ray & Heyes, 2011).
Importantly, to suggest that a fully functional biological system drives the onset of imitative behaviors sidesteps the developmental process of the biological system itself (Thelen & Smith, 1994). Whether this system is in place at birth or thought to mature by a certain later age, a developmental account must describe how it becomes “wired up” to respond to the actions’ of others in this way.

The alternative model suggests that the mirroring activation becomes established via various types of stimulation and experience of interacting with objects and others (C. M. Heyes & Ray, 2000). Empirical studies indicate sensitivity of MNS to training (Ray & Heyes, 2011). Specifically, repeatedly pairing non-matching actions (e.g. subject opens hand while observing a closing hand) can reverse typical mirroring patterns, such that observing the incongruent action leads to faster performance than observing the congruent actions (C. Heyes, Bird, Johnson, & Haggard, 2005). These behavioral changes are paralleled by changes in MNS activity (Catmur, Walsh, & Heyes, 2007).

These studies suggest that it is through repeated pairing of matching or congruent observed and self-generated actions that classic mirror neuron system associations are formed. In our own study (Chapter 2) mothers provided such input to infants by imitating their actions for months before the infants did any imitation. Other theoretical and empirical work suggests that imitation by the mother is critical for the development of the infant’s imitation skill (Jones, 2006; Ray & Heyes, 2011).
Additionally in Chapter 2, we noted that infant’s observation of self generated actions could lead to developments in recognizing and parsing the mother’s actions. This is another source of congruent visual and haptic input that is theorized to help establish the emergence of mirroring activation (Ray & Heyes, 2011).

Our studies add to this type of epigenetic model by identifying a number of factors in addition to action-parsing that appear to be important to imitation (see Chapter 2). These factors have not been identified in previous research on imitation development. For example, changes in infants’ distribution of attention allow infants to visually “check in” on their moms’ actions that are useful as they struggle to reproduce their mothers’ actions. This suggests a threshold where the majority of infants would not be able to produce any imitations until after this behavior is established, sometime between six and nine months. This is congruent with previous accounts of imitation development (Jones, 2009).

We also propose two other important factors in imitation development: improvement of haptic articulation, and the emergence of dyadic routines. As noted above, practice with doing and observing own actions provides infants with simultaneous motor and visual access of the sort thought to help establish associations important for parsing the actions of others and potentially MNS development. Additionally, we observed that as infants’ actions became increasingly articulated and organized into routines, mothers increasingly imitated their infant’s
actions. In this way the increasing complexity of infant’s object manipulations is part of a positive feedback loop that reinforces further elaboration of manipulations.

Another neurobiological or cognitive construct that comes to mind when reviewing our theory of imitation is working memory. Working memory (WM) is described as the ability to hold in mind a representation of something that is not currently in sensory contact. It is commonly thought of as persistent traces of sensory motor activity. Infants make important gains from 10 to 12 months in behavioral tasks that are thought to tap infant WM performance (Diamond, 1998; Munakata, 1998). For example, the A-not-B task is a classic developmental task in which an adult repeatedly hides a toy in one well (“A”) the presence of an infant, allowing the infant to “find” the toy after each trial. In the critical trial the experimenter hides the toy in the second well (“B”). Around 10 months, the majority of the infants search in well A although they have just seen the toy placed in well B. By 12 months most infants begin to correctly search in the B well. Modeling work indicates that immature performance can be simulated as a change the strength of the traces. Specifically, the rapidly decaying trace of the location of the toy in the most recent B trial is below the activity that has been built up over many trials of the toy being hidden in A. As the trace persists longer this no longer occurs (Thelen, Schöner, Scheier, & Smith, 2001).

Smith et al’s dynamic systems analysis (2003) revealed that infants’ performance could be made to look more or less mature by a varying factors that might increase or decrease the presence of this
Within our model of imitation development, changes in WM are most likely related to the infant’s continuity of actions, as well as increasingly distributed attention. Continuity of actions involves chaining repeated action into a sequence, as when actions are repeated, either by the infant or, increasingly, by a caregiver. Repetition of actions over a delay implies that the infant is storing a trace of the action, thus, changes in WM might extend the duration of routines or action sequences, especially when there are temporal gaps imposed by turn-taking. WM might contribute to distributing attention between toys, particularly when infants lose and later re-establish sensory contact with an object. At 9 and 12 months, re-establishing contact with a previously attended object allowed individual actions to become linked into increasingly complex routines.

However, while WM appears to be relevant to these processes, there are questions about its exact role in social development and dyadic interactions. Computational models of developmental change in A–not-B tasks and other related tasks often “manually” adjust a parameter that dictates the duration of the trace (see, e.g. Simmering, Schutte, & Spencer, 2008; Thelen et al., 2001). The implicit theory is that internal (biological) maturation of WM is what drives the observed trace (e.g. increasing vs. decreasing the wait time between hiding the toy and allowing the infant to search). Again, performance is not determined by a single internal mechanism but by a number of intersecting heterogeneous factors.
changes in behavior. This kind of parameter setting begs a question of how the parameter itself shifts.

An alternative is that changing WM is itself a developmental process facilitated, in part, by dyadic activity. The observational work described in this thesis is a step towards identifying the rich ecology of the types of experience that may lead to WM development. Specifically, our work suggests a number of specific biological-behavioral loops that are created as the environment provides novel challenges for the infant, and the infant gains social experience. We suggest that changes in such loops drive WM development, rather than the other way around.

In our work we specified a number of ways in which experience with manipulating objects can lead to novel behaviors. Such experience will correspond to changes in neural processing and structure. For example, an infant’s practice decoupling her hands to reach for and explore a new object while grasping a older one might provide practice that stimulates changes in working memory networks. Such changes will be dependent on specific sensorimotor contingencies. For example, when she switches gaze away from the old object, the likelihood that she will look back to this object is higher than if she was not still holding the object and therefore maintaining some haptic “reminder.” If she returns her gaze to that toy before its visual trace completely decays, the renewed visual input will reinforce the indirectly (i.e., intermodally), weakly facilitated association. Over time, a haptic
“reminder” might become less important, because a location-dependent visual trace can more robustly maintain itself.

This description also makes aptly clear that that infants’ changing distribution of attention during the first year, and related changes in parent-infant object coordination, are as much a matter of embodied sensory coordination as they are a matter of more traditionally conceived internal representations, or spatial working memory. As an infant decouples its hands to reach for and explore a second object while grasping a first one, they are no longer required to store a representation of that first object to continue to have access to it. Namely, the object remains in a privileged position of physical access. The organization of infants body in the world, is enough to accomplish this achievement (Hutchins, 1995).

In summary, rich descriptive observations of naturalistic interactions can ground and enrich past models of the internal mechanisms that drive behavior and development. In this chapter I have explored how better accounts of behavioral processes are necessary: biological accounts alone face the same pitfalls as the more traditional representational accounts. Ecologically valid accounts of the rich heterogeneous processes by which behavioral outcomes development require us to start with temporally rich, high-dimensional descriptions of observable actions of infants and caregivers engaged in minimally constrained social interactions.
References


