THE ROLE OF SHARED PRACTICE IN THE ORIGINS OF JOINT ATTENTION AND POINTING

by

Timothy Racine

B.A. (Hons.), Simon Fraser University, 1998

M.A., Simon Fraser University, 2000

Dissertation Submitted in Partial Fulfilment

of the Requirements for the Degree of

Doctor of Philosophy

in the Department

of

Psychology

SIMON FRASER UNIVERSITY

Summer 2005

APPROVAL

Name: Timothy Racine

Degree: Doctor of Philosophy (Department of Psychology)

Title of Thesis: **The Role Of Shared Practice In The Origins Of Joint Attention And Pointing**

Chair:
Dr. Bryan Sokol
Assistant Professor, Psychology

Dr. Jeremy Carpendale
Senior Supervisor
Associate Professor, Psychology

Dr. William Turnbull
Supervisor
Professor, Psychology

Dr. Michael Maraun
Supervisor
Associate Professor, Psychology

Internal Examiner:
Dr. Jack Martin
Professor, Faculty of Education

External Examiner:
Dr. Vasudevi Reddy
Reader In Developmental And Cultural Psychology
University Of Portsmouth

Date Approved : Monday, June 6, 2005

SIMON FRASER UNIVERSITY

PARTIAL COPYRIGHT LICENCE

Simon Fraser University

Ethics Approval

The author, whose name appears on the title page of this work, has obtained human research ethics approval from the Simon Fraser University Office of Research Ethics for the research described in this work, or has conducted the research as a member of a project or course approved by the Ethics Office.

A copy of the approval letter has been filed at the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for ethics approval and letter of approval is filed with the Office of Research Ethics. Inquiries may be directed to that Office.

Bennett Library
Simon Fraser University
Burnaby, BC, Canada

ABSTRACT

The infant's participation in sequences of joint activity that require visual attention, or what is often called joint attention (e.g., gaze following, social referencing, pointing), is usually argued to be an outcome of and evidence for the existence of particular psychological competencies on the part of the infant. This convergence in opinion occurs, I contend, because social developmental theory is predicated in large part on a causal-psychological-representational picture of meaning and mind. In a review of the relevant literature I suggest that what is presupposed or at least underappreciated in most theories of joint attention is the role of social practice in the understanding of other minds. Rather than assuming that these joint activities requiring attention reveal psychological competencies, I argue that: (a) it is with recourse to such practices that researchers theorize about the infant's understanding of mind in the first instance, (b) the mind is not some entity that is separable from human activity, (c) knowledge of such practices is what the developing agent requires to come to an understanding of other minds, and (d) because non-language using agents do not understand mental states, young babies, a fortiori, do not understand their own attention or that of others. I then focus on a single infant activity typically thought to reveal a leg hold on other minds (i.e., pointing) and I report two empirical studies of its development. In the first I evaluate theories of the emergence of pointing by investigating the ways in which infants first get into shared activities involving pointing. In the second I investigate the development of pointing within mother-infant interaction. In neither study do I find support for the position that pointing gestures emerge as a result of some conceptual revolution on the part of the infant.

DEDICATION

I dedicate this work to my wife, Corinna Salvail, and to my children, Ellis, born four months before I started my M.A., and Thea, born a month before I started my Ph.D. My children are fortunate that they have no sense of what life might be like had their father not been interested in trying to answer his questions about the mind and the ways in which in children might develop an understanding of such a multiplicity of social situations. Their poor mother has no such good fortune and may also have a much better imagination.

ACKNOWLEDGEMENTS

I thank Dr. Jeremy Carpendale for, among many other things, stimulating my interest in squaring what I knew about Wittgenstein with what he knew about Wittgenstein's relevance for understanding children's social development. I thank Dr. William Turnbull for introducing me to the methods of conversation analysis, and as importantly, for the theoretical rationale for the employment such a methodology. I thank Dr. Michael Maraun for helping to clarify my use of Wittgenstein and for also suggesting other key secondary sources for interpreting Wittgenstein's legacy.

I thank my parents, Phil and Amy Racine, Corinna's parents, Gerry and Jane Salvail, and Corinna's grandmother, Patsy Lewis, for their support and encouragement over the years.

I thank the Social Science and Humanities Research Council of Canada and the Michael Smith Foundation for Health Research for awarding me Doctoral Fellowships to fund the work presented in this dissertation.

I also thank the good folks at Prado Café on Commercial Drive for their tolerance of my many hours of toil on this work in their establishment -- and their awesome coffee.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1 INTRODUCTION

Typically developing infants begin to engage in activities towards especially the end of their first year of life that require simultaneous visual attention on the part of both infant and parent.[1] Developmental psychologists often describe these emergent patterns of activity (i.e., reliable gaze following, social referencing,[2] object-based imitation, pointing, following another's point, early word learning) as episodes of joint attention (see chapters in Moore & Dunham, 1995 and Kita, 2003; for a review, see Moore, 1996b). However, developmentalists seem to use the term "joint attention" not to mean "two people engaged in joint activities that require attention" but rather as shorthand for "the infants' psychological processes and/or states that enable them to participate in joint activities that require attention." Although the latter may be the problem space in which we want our theorists to do their work, this transposition from "outer joint activity" to "inner mental activity" is but one particular way of setting up this problem. And as Wittgenstein (1958, §308) has pointed out:

> The first step is the one that altogether escapes notice. We talk of processes and states and leave their nature undecided. Sometime perhaps we shall know more about them -- we think. But that is just what commits us to a particular way of looking at the matter. For we have a definite concept of what it means to learn to know a process better. (The decisive movement in the conjuring trick has been made, and it was the very one that we thought quite innocent.)

[1] Such activities obviously involve other sensory modalities (e.g., most often audition is implicated as well), but research with typically developing children has focussed almost exclusively on the visual modality. (For research on joint attention that investigates non-visual modalities in atypically developing populations, e.g. blind children, see e.g. Bigelow, 2003.)

[2] In activities involving what is referred to as social referencing, an infant looks towards a parent when confronted with an ambiguous stimulus as if seeking to use the parent's emotional reaction to guide her reactions to the novel stimulus (Walden & Ogan, 1988).

In the research literature on joint attention, the act of pointing has attracted particular interest because: (a) gestural communication has been argued to reveal an "implicit theory of mind" (Bretherton, 1991; Bretherton, McNew, & Beeghly-Smith, 1981; but see Barresi & Moore, 1996; Carpendale, Lewis, Müller, & Racine, in press; Hobson, 1991; Moore & Corkum, 1994; Racine, 2004; Racine & Carpendale, in press; Reddy & Morris, 2004), (b) its relative absence is predictive of childhood psychopathology, in particular, autism (Baron-Cohen, 1989; Carpenter, Pennington, & Rogers, 2002; Goodhart & Baron-Cohen, 1993), (c) the timing of the emergence of its "communicative function" has been linked to language development (Bates, Camaioni, & Volterra, 1975; Bates, O'Connell, & Shore, 1987; Bruner, 1983; Butterworth & Morissette, 1996; Goldin-Meadow & Butcher, 2003; Trevarthen, 1979; Werner & Kaplan, 1963), and (d) pointing has been argued to not naturally occur in great apes and hence its absence has been claimed to demonstrate a bifurcation between the human and non-human primate lines (Butterworth, 1998b, 2003; Povinelli, Bering, & Giambrone, 2003, Povinelli & Davis, 1994; Tomasello, 1999b; but see Krause & Fouts, 1997; Leavens, Hopkins, & Thomas, 2004; Leavens, Russell, & Hopkins, 2005b; Racine & Carpendale, in press; Savage-Rumbaugh, Shanker, & Talbot, 1998; Tomasello, Call, & Hare, 2003).

Pointing can be used in at least three different patterns of shared activity, or to put in the way in which it is most often put in the field, pointing can perform at least three distinct functions: (a) requesting objects or actions from others (the so-called protoimperative function), (b) directing attention to share in an object with another (the so-called protodeclarative function), and (c) informing others about object location (which has yet to be assigned a corresponding precise function, but no harm would be done in

dubbing this the protoimformative function) (Bates et al., 1975; Liszkowski, Carpenter, Henning, Striano, & Tomasello, 2004; Liszkowski, Carpenter, Striano, & Tomasello, in press). Some argue that it is only protodeclarative use that grounds language development (Bates et al., 1987; Camaioni, 2001; Franco & Butterworth, 1996). This claim rests on the belief that protoimperatives signify an infant's understanding of others as only "agents of action," whereas protodeclaratives signify that infants understand others as "agents of contemplation" (Camaioni, 2001). And although the recently identified informative use of pointing (Liszkowski et al., in press) has not been tested in children with autism, of the two more classical functions, it is the absence of protodeclaratives that has been most closely linked with autism (Baron-Cohen, 1989).

1.1 What is Joint Attention?

Despite the fact that quite a bit is known about pointing and the other joint activities in infancy that require attention, there is disagreement in the field over the age at which an "understanding of attention" can be said to exist and what such an understanding consists of (Baldwin & Moses, 1994; Butterworth & Jarrett, 1991; Butterworth & Morissette, 1996; Carpenter, Nagell, & Tomasello, 1998; Corkum & Moore, 1998; D'Entremont, 2000; D'Entremont, Hains, & Muir, 1997; Franco & Butterworth, 1996; Liszkowski et al., in press, 2004; Meltzoff, 1995; Moore & D'Entremont, 2001; Morissette, Ricard, & Decaré, 1995; Müller & Carpendale, 2004; Racine & Carpendale, in press; Scaife & Bruner, 1975; Slaughter & McConnell, 2003; Tomasello, 1995, 1999a). In a parallel vein, there is also disagreement over how to properly characterize joint attention. For example, Butterworth (1998b, p. 171) has

claimed that joint attention is "simply…looking where someone else is looking," which reflects his bottom-up and evolutionary analysis of the origins of gaze following and pointing in particular. In contrast, Tomasello (1995, p. 106) has claimed that joint attention requires that "two individuals know that they are attending to something in common," which reflects his top-down approach to explaining the entire constellation of emergent activities that involve the simultaneous attention. The key point of contrast between these two definitions seems to be that Butterworth is silent on the issue of whether infant and other understand one another in psychological terms in activities that involve coordinated attention, whereas Tomasello's definition requires that an infant *knows* that their attention is coordinated with another's.

Tomasello's upping of the epistemological ante brings into focus a key issue. Namely, how are *we* to know that an infant knows that their attention is coordinated with another's? Well, usually if I want to know whether a person knows something, I ask her. And I might, for example, say, "Betty, did you know that your cat is also paying attention to that toy that you are playing with?" If there were something in Betty's reactions that seemed at odds with what she claimed to know, then I would ask further questions until there were sufficient grounds to accept her avowal. However, this may not be very helpful at present because when infants first get involved in joint activities requiring attention they do not speak. In fact, developmentalists agree that "joint attention" is required for them to learn a language. And as noted earlier, theorists in the main seem to mean that "joint attention is required to learn language" in the sense that a particular psychological competency must be present in the child not that infants must master certain ways of interacting with others. But be that as it may, we are not any closer to

knowing whether it makes sense to say that infants know that they are attending to something in common with another. Very young infants also respond to the attention of others (Reddy, 2001, 2003). Although being the object of another's attention may play an important formative role in social development, what is the relation between these sorts of ubiquitous everyday activities and an infant's understanding of another's attention? And a fortiori, what should we make of claims that "even toddlers *understand* emotions and desires as internal and subjective and understand action and speech as guided by the person's intentions" (Wellman, Phillips, Dunphy-Lelii, & Lalonde, 2004, p. 283, my emphasis)? In this dissertation, I will present one way of addressing these issues that attempts to dissolve the inner/outer dichotomy embedded in social developmental theory and expressed in this quote from Wellman and colleagues (2004).

1.2 Overview of Dissertation

This dissertation contains a conceptual analysis of joint attention and an empirical analysis of the development of one joint attention activity, namely pointing. In the first section, I expand on my present discussion of how joint attention is typically set up as a problem to solve by presenting some arguments on the nature of meaning and mind derived mainly from Wittgenstein (1958, 1969, 1980a, 1980b, 1981, 1992). In the second section, I review and critique the main theories of joint attention from these considerations. In the third section, I consider the relevance of this conceptual work for psychological theory and present a roughly Wittgensteinian account of joint attention. In the next two sections I link this approach to empirical work. In the first empirical study, I consider the issue of how pointing emerges within social practice in the first instance. In

it, I review the research literature regarding various theories of the origins of infant pointing and I present data that tests these theories in the pointing gestures of a sample of 9- to 12-month-old infants assessed with a semi-structured observational measure (the Early Social Communication Scales, Mundy, Delgado, Block, Venezia, Hogan, & Seibert, 2003). In the second empirical study, I draw on ethnomethodological and derivatively conversation analytic approaches to parent-child interaction to show how pointing emerges within particular mother-infant interactions from the same sample. That is, I show what mother and infant are first doing when pointing becomes an interactional resource upon which the infant can draw. In the sixth and final section I summarize my argument and empirical findings and discuss their importance for social developmental theory. In so doing I attempt to address any limitations in the approach that I take to social understanding in infancy.

CHAPTER 2 SETTING UP THE PROBLEM

> Our mistake is to look for an explanation where we ought to look at what happens as a 'proto-phenomenon.' That is, where we ought to have said: *This language-game is played* (Wittgenstein, 1958, § 654, his emphasis).

It is a common belief that a word or phrase has been correctly used if it brings to mind the images and associations in an interlocutor that the speaker had intended to communicate. On this view, understanding a "proposition" is assumed to require a mind or brain state to causally mediate between a thought and some state of affairs. And understanding is taken to be a mental state bridging thought and action. In contrast, Wittgenstein (1958) argued that mental state terms such as understand, know or attention take on their meanings from the role(s) they play in their respective language-games, which are sets of rule-governed conventions delimiting the relations between linguistic terms and the activities with which they are tied up. He suggested that it is in and through the child's immersion in activity with some knowledgeable other that a child develops a mastery of such a given set of rules. And shared activities are patterned around the characteristic regularities in various forms of life. In fact, activity (including speech) is intelligible to others because it is grounded in participation in some shared form of life. Thus, to understand the meaning of a particular mental state term is to know the aspects of circumstances that logically justify the use of the psychological expression in question. Wittgenstein often called these the *criteria* that govern the use of a term. And from a Wittgensteinian perspective on social development, children would need to grasp the criteria for the meaning of mental state concepts (Carpendale & Lewis, 2004; Chapman, 1987; Montgomery, 1997, 2002; Racine, 2004; Racine & Carpendale, in press; Turnbull & Carpendale, 1999, 2001). In the case of what babies understand

about another's attention, then, the question posed in Tomasello's definition is whether or not there are grounds for saying that a baby *knows* that their *attention* is coordinated with another when engaged in joint activity.

Now, some philosophers (e.g., Shotter, 2004) would claim that the phrase "a Wittgensteinian perspective on social development" embodies a deep and egregious paradox that grossly distorts the spirit of Wittgenstein's philosophy. Although it is true that the later Wittgenstein's sole goal was to expose non-sense for what it was, programs of research, which can ultimately be judged as more or less adequate, can be still be grounded in Wittgenstein's way of approaching meaning and mind even if that is not what Wittgenstein intended. And I also don't imagine that Wittgenstein would have been as offended by this as might be the critics of such an enterprise because nowhere does Wittgenstein say that psychological theory is without value. Rather he said it is confused because it often conflates conceptual and empirical issues. But it might be useful to separate a Wittgensteinian philosophy with its goal of conceptual clarification from a "Wittgensteinian psychology" that tries to build a positive account of human development that follows from a clarification of meaning and mind (e.g., Carpendale & Lewis, 2004).

According to Wittgenstein, a feature of psychological predicates is that they possess what he would call a very complicated grammar, meaning that they are used in variety of ways. And although some of these uses are similar, others are not (Bennett & Hacker, 2003; Hacker, 1990, 1996; ter Hark, 1990). However, by subsuming them under a superordinate category we can easily fail to notice this feature. And beyond the different ways in which mental state terms can be manifestly used to perform different

social acts, there are other logical differences of a different order between mental state concepts. For example, attention has genuine duration; intention, desire and belief do not. Attention, intention and desire are grounded on previous natural reactions; belief is not. This cautionary note is here to alert the reader that I am not claiming that mental state concepts really share much in common other than the rather obvious feature of seeming to "refer to the inner." However, Wittgenstein argued that even this aspect of mental predicates turns out to be false.

An important, and logically prior, issue overlooked in Tomasello's definition, and indeed in the prototypical way of approaching social development, becomes apparent when one considers Wittgenstein's (1958) private language arguments, which demonstrate the logical impossibility of being able to reference mental content outside of an ability to lay down a rule (Hacker, 1990; ter Hark, 1990). Wittgenstein understood that it is easy to inappropriately impose a term employed in a language-game that describes objects, and hence demands referential idioms, (e.g., perception) onto another that seems to describe 'inner objects', but cannot accommodate such idioms (e.g., introspection). Wittgenstein uses the example of naming sensations to show that private (seemingly ostensive) definition could not allow for the verification of whether a symbol (i.e., a stand-in for the private mental event in question) is the meaning of the symbol. In such a case there would be no way for an agent to know whether she is following the rules involved in a correct use of a term -- and if one does not possess a rule of application, then one cannot communicate it to others. Some means other than another subjective impression is therefore needed to verify whether the rule has been correctly applied and "in the present case I have no criterion of correctness. One would like to

say: whatever is going to seem right to me is right. And that only means here we can't talk about 'right'" (Wittgenstein, 1958, § 258). This, then, is a problem of normativity, a problem of how a state of affairs can be right or wrong, true or false, in error. And pace Tomasello and virtually all of social developmental theory, an infant cannot introspect upon her mental world to conclude she is experiencing mental event A, because she possesses no standard with which to know she is experiencing A. To do so presupposes, on pain of circularity, that she already is familiar with the meaning of A. But the theorist is well versed in the technique for the application of any mental state term so this paradox is easily overlooked.

When we say that people are attending to some state of affairs, we mean that they are deliberately focussing on it. This is something that animals naturally do. And when we tell a child to "pay attention" we are telling them that they should focus on one activity and not another, or stop what they are doing that would interfere with their focussing on that activity, and so on. When a child is focussed on some state of affairs we afford them a language of attention by the way we talk about their activity. To be in command of such a language is to understand the concept attention. And although it might make sense to speak of attention as a mental *state* given that it has a clear beginning and end, what we often mean by mental state is something "higher order" (e.g., thinking, believing, wanting, intending and so on). Indeed, in addition to a clear beginning and end what we mean by a mental state is a state that can be interrupted by attentional shifts (Wittgenstein, 1980a, 1980b, 1981). However, what we mean by "attention" is more straightforward than what we mean by "know." Although we can know that other people are paying attention to some state of affairs, that is, that we have

criteria for saying that attention can be seen in their activities we have no such criteria for our own activity and as such do not *know* if we are paying attention or not. Normativity does not enter into it because there are no grounds for doubt in our own case (Wittgenstein, 1969). However, this does not mean that we have privileged Cartesian access to our own mental life and have to arguably infer the mental lives of others. Instead, meaning is in the way they are used. And our expressions of mental life do not describe our mental states but rather instantiate and define our "mental states."[3] The challenge in Tomasello's formulation is that he does not claim that *we the observer* know that two individuals are attending to something in common, but rather *that the two individuals themselves* know that they are attending to something in common. This entails that they express their own attentional state *and* their knowledge of their interlocutor's attentional state. This turns out to be trickier than one might think.

2.1 Unpacking Butterworth's and Tomasello's Definitions

Early "forms" of gaze following are evident in the activities of infants as young as three months of age (D'Entremont, 2000; D'Entremont et al., 1997; Scaife & Bruner,

[3] To apply criteria is not an inferential process, which is why I prefer to write of seeing attention in activity instead of applying criteria to judge that concepts are present in activity. I am not of course claiming that observers literally see *the concept* "attention" in activity but rather that observers see *the application of* the particular concept in activity. However, although a concept is not part of natural reality, it might be useful to think of it as part of conventional reality (Kusch, 1997). In any case, although seeing the meaning of activity *is* to apply criteria for a particular concept to activity, I am concerned that phrasing this in terms of "application of criteria/concepts" could be misread as a commitment to a mediated and homuncular picture of psychological meaning that I wish to avoid.

1975).[4] This would seem to satisfy Butterworth's definition given that the baby looks (in at least a rough sense) where someone else is looking, but presumably not Tomasello's definition given that many researchers are reluctant to claim that such a young infant would "know" that their attention is coordinated with another when following their gaze -- that is, they apparently do not have grounds for claiming that the infant knows their attention is coordinated with another's.[5] Tomasello's *explanation* of joint attention, for example, does not allow for such knowledge in the 3-month-old infant. His theory (1995, 1999a) is that all of the "manifestations of joint attention" occur because 9- to 12-month old infants "understand other persons as intentional agents" (Carpenter et al., 1998, p. 5). It is said to be this grasp of, or I would call it insight into, their own intentions and those of others that drives the emergence of these activities. Given that gaze following is typically thought of as an intentional act and 3-month-old infants can follow gaze to some extent, it begs the question why one would claim that babies do not understand their own intentions at three months of age but can presumably be said to six months later -- and accordingly can also be said to know they are attending to something in common with others. What are the grounds upon which we might answer such a question? From a Wittgensteinian perspective, we would need to clarify what "intention," "know" and "understand" mean (a conceptual issue) and then look for evidence of babies

[4] "Forms" is in scare quotes in its first usage to signify my unease with a particular possible way of thinking tied up with its use. The issue is that although researchers have identified earlier and later activities that involve, in this case, gaze following, it is not the case that these are "forms" in some sense of pre-existing taxa. Researchers have created such forms and while these forms express what *we* call different levels of understanding on the part of the infant, it is not clear: (a) what the infant understands of her own activity, and (b) whether we are justified in assimilating all forms into the same hierarchical category. What we have here is activity, and even if, for example, it is structurally isomorphic it need not be logically or functionally so.

[5] Infants seem to do this by first using another's head turn and bodily posture as cues before ultimately relying on eye direction. This is still referred to as gaze following in the field though and I will follow this convention for at least my presently illustrative purposes.

satisfying the criteria entailed in these definitions in their coordinated engagements with others (an empirical issue).

Tomasello claims instead that he has marshalled sufficiently convincing empirical evidence that he is in fact justified in concluding that infants "understand other persons as intentional agents" and that the infants therefore "know they are attending to something in common with another" (Carpenter et al., 1998).[6] In particular, he and his collaborators report that they have discovered that the emergence of the activities involving joint attention proceeds in a roughly identical sequence for most babies and that the age of emergence for the various activities is moderately intercorrelated. These findings suggest to Tomasello that around 12 months infants know that they are attending to something in common with others or put another way that infants grasp that others have intentions like they do. However, the two commentators on Carpenter and colleagues' monograph (Butterworth, 1998a; Moore, 1998) were not as convinced as Tomasello that the data support his position.[7] The real problem here though is that Tomasello's data cannot distinguish what the baby knows from what we the adult observers know because he has assumed a particular (and unexamined) meaning of "know" (and "understand" and "intention"). And at this point all we seemed justified in saying is that although Butterworth's and Tomasello's definitions of joint attention might both be useful, they cannot be used interchangeably. This fact might sensitize us to the

[6] Tomasello (1995) equates these two expressions and thinks that the shared activities show that the infant knows their attention is coordinated with others'.

[7] Carpenter et al. also ignored within-subject sources of variance in their study. Thus, if infants pointed at say 10 months, Carpenter and colleagues assumed infants would also point at 11 and 12. Clearly this tilted the scales towards supporting a theory predicated on the notion that all behaviours emerge from a common source given that the opportunity to see if inconsistent (gradual, "non-insight based") development was disallowed. (I return to this issue in Study One.)

possibility that there are indeed both empirical and conceptual issues involved when investigating the understanding of attention in infancy.

Now, Butterworth's definition at least gives a self-evident ground as to why researchers and people more generally would say that two agents are attending to the same object or state of affairs, namely that they are both looking at that self same object. But suppose we knew one of these individuals to be blind. If this person were to turn her head and "gaze" towards the same object as the other we would not and should not conclude that joint visual attention was occurring.[8] There is also an ambiguity in Butterworth's definition that should be addressed before proceeding. I presume that Butterworth does not intend that two agents could be said to be "engaging in joint attention" by happenstance. That is, it would seem that his definition would be satisfied by two people looking out different windows at the same event without an awareness that the other was simultaneously looking as also constitutive of joint attention. But I will assume that this is not what Butterworth means. However, whereas these active versus passive readings of Butterworth's definition are problematic, Tomasello's definition is not a corrective for them because a claim that an agent is aware that she is engaged in a joint activity requiring attention (which is what I sympathetically take Butterworth to mean) is clearly different than a claim that an agent *knows* that she is engaged in such an activity.

Barring such defeating grounds when we see that two individuals are both looking at the same object we human observers understand that they are jointly

[8] Similarly, if a human and a robotic agent were both "looking" at the same object, pace Nagai, Hosoda, Morita, and Asada (2004), it does not necessarily follow that joint attention is occurring or that a robotic agent can "acquire the ability of joint attention" (p. 212). This is not to say that such claims could be not established, but simply that they have not been by Nagai and colleagues (Racine, 2002).

attending to the same object. This shows us that the relationship between logical criteria and empirical evidence is a defeasible one -- criteria do not guarantee the application of a particular concept to some state of affairs; we can always be wrong (Wittgenstein, 1958).[9] The relation between criterion and use of the term in question is tautological -- it is a definitional or internal relation, not a causal or external one (ter Hark, 1990; Baker & Hacker, 1985; Hacker, 1990). Joint visual attention, by definition, occurs when two agents are (actively) looking at something in common. If observers understand the concept joint attention, then they see joint attention in such shared activities. That is, to see joint attention in such activities is to understand that the concept joint attention is ascribed under these conditions. And to understand such a psychological concept is to have knowledge of the logical grounds, or criteria, for its application -- i.e., what joint attention means. If the grounds for this claim are defeated, then whatever observers do see they do not see joint attention.

Having a psychological concept is having knowledge of some of the multiplicity of particular situations where that mental predicate might apply, and there are obviously a multitude of distinct social activities of which it might make sense to say involve "joint attention" -- some of which have not even yet emerged. Wittgenstein (1981, § 114, his emphasis) noted in this context that:

> One learns the word 'think', i.e., its use, under certain circumstances, which, however, one does not learn to describe. But I can *teach* the use of the word! For a description of those circumstances is not needed for that. I just teach him the word *under particular circumstances*.

[9] Rather than this lack of entailment undermining a model that relies on criterial definition, this actually speaks in favour of such a model because error must be detectable in any kind of a "representational" system, normativity must be possible, for it to work (e.g., Bickhard, 1993; Racine, 2002).

Given this, psychological concepts are not "in the head," but are rather tied up with and instantiated in particular patterns of shared activity. And such concepts are abstracted from the use of words (Bennett & Hacker, 2003).[10] Although there are also neurological-psychological events taking place, one might say presupposed, in both an observer of two people engaged in a particular shared activity involving joint attention and in those so observed that are causally related to these activities, joint attention is defined by the sequential activity of two people jointly attending given *this* particular practice and not by those mind or brain states. Although these latter phenomena are causally related to the activities that partially constitute joint attention they are not definitionally related to such shared activities because we do not cite such events in explanations of the use of these concepts. But nothing can stand in for the internal relation between criteria and concept defined (Baker & Hacker, 1985; ter Hark, 1990). In fact, to be able to empirically map out such external causal relations demonstrates that a logically prior internal definitional relation must exist (Bennett & Hacker, 2003).

Developmentalists who have applied Wittgenstein's criteriological approach to semantics (e.g., Chapman, 1987; Montgomery, 1997) have assumed that the meaning of a mental state concept and criteria are in one-to-one correspondence and a fortiori

[10] Now, clearly there are neurocognitive factors that are causally related to people's acquisition of concepts, but this does not make the concept a mental entity -- it is defined, it "breathes life," in activity. Also, although I endorse the spirit of Bennett and Hacker's description of abstraction of concepts from word use, I want to be clear that my model of the psychological processes at work is one that is predicated on a unitary view of memory-mind (Brooks, 1978, 1987; Whittlesea, 1997, 2003) and in so doing emphasizes particular experiences. However, of course as an agent has more and more similar experiences the "neurocognitive system" on a strictly mathematical basis alone would become less sensitive to atypical experiences and more sensitive to typical ones by definition and in this sense it might function like a system composed of abstracted prototypes. Whittlesea (1997, p. 340), however, contends that "there is no positive evidence for a dedicated mental apparatus that abstracts general, summary properties across particular experiences, or that conceptual knowledge is represented in a separate memory system." At minimum this does away with the paradoxical notion of a "semantic memory system" (see point (b) of my characterization of the commonsense picture of mind below).

that criteria are presumably therefore constitutive of meaning. However, although concepts are defined by criteria, they are partially not solely constitutive of meaning (Canfield, 1981; Bennett & Hacker, 2003). A linguistic (or gestural) token in a language-game, that is, a concept fixed by particular rules, means what it means in part by virtue of the characteristic form of life in which it is rooted. If these aspects of existence change, rules of the corresponding language-games may have no application. Form of life is distinct from the issue of applicability or defeasibility of criteria for particular mental state concepts. And there are a variety of social situations that may serve as criteria for the application of a particular mental state concept. For example, sorrow goes along with certain characteristic social situations (e.g., grieving, separation, and so on). If a person were not in such a situation, one would be much less likely to judge her to be experiencing sorrow. However, if a person were to exhibit sorrowful behaviour, this might well lead one to conclude she is in such a situation. In this sense, criteria are animated in particular situations. That is, not only need a criterion for a judgment of sorrow be met, but it must also be employed in a sorrowful circumstance *and* with respect to the basic function that sorrow plays in our lives. And social situations and background conditions are not contained in the criterion. Rather, the criterion is employed in these circumstances and in the human form of life. Thus, criteria are partially, but not solely, constitutive of mental state concepts.[11]

This brings into focus the issue of following a rule versus seeming to act in such a way that appears to be following a rule. It is seems obvious that rule following involves

[11] This decoupling of concept from criteria would still not presumably satisfy Bickhard, who argues (e.g., 1987, 2004) that Wittgenstein's account of the emergence of normativity fails because Wittgenstein forces a "brute normativity." But the origins of normativity are irrelevant to meaning.

more than just acting in accord with a rule. However, this is to misunderstand Wittgenstein's strictures on language-games as rule-governed activities. Clearly, no psychological intermediary is required to follow a rule. A criterion for understanding a rule is to correctly follow it. And to follow a rule is nothing more than to act in accord with it. However, the satisfaction of criteria for understanding a rule is again conditional upon form of life and defeasibility:

> The illusion that there is something more [to following a rule than merely acting in accord with it] arises from the platitude that it is possible to act in accord with a rule without following it. We are tempted to conclude that the differences must be intrinsic to the act, whereas in fact it turns on the *circumstances* that surround the act....It would make no sense to describe two persons as playing chess (or following the rules of chess) if they belonged to a culture unacquainted with games....And it would undermine the claim that someone in our culture was following the rules of chess if he waited on whispered instructions from another or he had been drilled to perform a single invariant sequence of moves. It is only against a complex background that acting in accord with a rule counts as following a rule. (Baker & Hacker, 1985, pp. 158-159, their emphasis)

On Butterworth's definition, infants and their parents can quite rightly be said to engage in joint visual attention because particular patterns of sequential activity satisfy the simple criterion of both simultaneously looking at the same object or state of affairs. That is, we are following a rule for use given particular circumstances. To the extent that other members of species act like humans, it would make equal sense to see joint attention in the coordinated activities of two individuals of such life forms (Wittgenstein, 1958, § 281). *This is because it is we human adult observers who are seeing their activity in "psychological" terms.* And in so doing we are not necessarily making claims about what these agents know about their activity -- rather we are claiming that their activities license claims about what *we* know. However, unlike Butterworth, Tomasello is not content to use human concepts to describe activity; Tomasello wants to account for

the agents' understanding of these concepts. This difference between these two definitions has implications not only for infancy work, however, but also for work on social cognition in the great apes.

2.2 From Infants to Chimpanzees

Tomasello's definition stems, I suspect, in part from his goal of trying to account for the social cognitive differences between human and non-human primates. Now, until very recently, Tomasello's (1999b) position has been that great apes (and chimpanzees in particular) do not see others' or their own behaviour in psychological terms. As noted, his theory of joint attention in infancy (Tomasello, 1995; 1999a) has been that whereas younger infants and non-human primates may engage in activities involving coordinated attention with others, one-year-old human infants engage in joint attention with others in a manner suggesting that an infant of this age has come to experience others as separate agents who act upon their world as a consequence of their mental states (as noted earlier, specifically their intentions). This provided a bifurcation across the primate line, with the human species as the resultant social cognitive pinnacle (Tomasello, 1999b). Similarly, Povinelli and colleagues (2003) have claimed that although human and non-human primates may gesture in similar ways, there are radically different psychologies at play in the human and non-human lines when such behaviours are performed. However, Povinelli et al. also mistakenly imply (2003, p. 42) that critics of their position are behaviourists (cf. Leavens, 2004). Povinelli and colleagues do not believe that chimpanzees point but they instead hold that chimpanzees only "use their index fingers." The reason for this is said to be that chimpanzees do not "have second-

order mental states." However, there is good reason to think that chimpanzees do point (Krause & Fouts, 1997; Leavens, 2004; Leavens & Hopkins, 1999; Leavens et al, 2004; Savage-Rumbaugh et al., 1998) and although Povinelli et al. seem to mean something very different than I do by a second-order mental state they are quite right that apes do not know that they have attention, intention and the lot. However, it is not clear that humans have second-order mental states either in the way that Povinelli et al. believe.

Although social cognition in the great apes is highly relevant for theories of infant social cognitive development and although I agree with Tomasello and Povinelli that it is humans, and only humans, who understand psychological concepts and variously "see" them in the activities of human and non-human species alike, I argue that their mentalistic line in the sand is predicated on a commonsense theory of meaning and mind that leads to confusions in theorizing about social understanding in both human infants and non-human primates. As Leavens (2005b, p. 292) and his colleagues have similarly noted:

> the authors [have] failed to describe the avenue or channel through which they can identify psychological differences between human children and apes other than overt, publicly observable behaviour.[12]

Although we certainly have grounds for claiming that human infants and also non-human primates alike attend to *x*, want *y* and intend to do *z* for the simple reason that these agents act in particular ways that express their attention, desires or intent and we know the rules of these language-games, we should be very careful when making claims about what non-language using creatures understand of these very same activities. Attention, desire and intent are ways of talking about particular patterns of

[12] I would add that the same issue creeps in when making human infant-infant comparisons.

activity. What patterns? Patterns that express and partially constitute attention, desire or intent. The terms are tokens in language-games. And "the first language-games are located -- 'logically' -- between the complex language-games of adult life and the proto language-games we share with apes" (Canfield, 1995, p. 197). Thus, the ascription of mental predicates (paradigmatically) presupposes particular capacities on the part of the agent so described. Animals attend to things, want things and intend to do the things they do. However, this is not a property of "animal minds" (and a fortiori, not a property of their "inner mental states"), but is rather a property of an agent's activity in a given form of life. That is, where human and non-human primates might differ is in their capacities, not in their minds.

A challenge for Tomasello's and Povinelli's reductive read of the activities of the great apes, and indeed derivatively for Tomasello's theory of infant social cognition, is that he and his colleagues (Call, Hare, Carpenter, & Tomasello, 2004; Tomasello et al., 2003) have more recently questioned their initial scepticism about chimpanzee social cognition and have concluded that in fact "chimpanzees can understand some psychological states in others -- the question is only which ones and to what extent" (Tomasello et al., 2003, p. 156). The reason for this caution is that in a series of recent experiments chimpanzees have been observed to engage in particular types of activity that show knowledge of what a conspecific has seen and knowledge of a human experimenter's intentions. However, we should pause and give thought to what these agents really understand of activity -- that is, to clarify what mental state terms mean -- a prerequisite to a claim that an agent "understands mental states in others."

It is also interesting that although these types of activity investigated by Tomasello and colleagues involve joint attention and presumably require an agent's "understanding of attention," they do not satisfy the more epistemically demanding definition of Tomasello -- i.e., that two individuals *know* they are attending to something in common. We are now back to Butterworth's observer-positioned definition. Tomasello and colleagues may be managing this seeming contradiction by claiming that "beliefs are indisputably mental" (Tomasello, Carpenter, Call, Behne, & Moll, in press, p. 2) and thereby perhaps leaving the "ontology of lower-order mental states" like attention, desire and intention somewhat open to debate. What I take Tomasello et al. (in press) to really want to claim though is not that beliefs are indisputably mental but are rather indisputably human. But this needs clarification. Whereas the *understanding* of belief is indisputably human, it may well make sense to see belief in the behaviour of non-human primates. That is, there may be circumstances that might quite rightly justify the use of the term to describe a chimpanzee's activity (Racine & Carpendale, in press). Sowing the seeds for further confusion though is the fact that Tomasello et al. (2003) claim that future empirical work will clarify what the great apes do and do not understand about mental states.[13] But comparative and developmental psychologists alike need to clarify the concepts they are using before (more) empirical work is done.

[13] Tomasello et al. (2003, p. 156) note in particular that they "believe that the way forward in research on chimpanzee social cognition is to 'turn up the microscope' so as to see which of the many different kinds of primate psychological states chimpanzees are able and not able to comprehend, and in what precise ways." The visual metaphor of turning up the microscope to see with precision which mental states chimpanzees understand is an ironic turn of phrase given the conceptual nature of this issue.

2.3 Problems in Butterworth's and Tomasello's Definitions

Tomasello is not the first theorist to suggest that joint attention entails that two individuals understand each other in psychological terms. In fact, this commonsense interpretation of infants' understanding of attention is the most widely received view (Carpendale et al., in press; Müller, & Carpendale; 2004; Racine, 2004; Racine & Carpendale, in press). A very clear and intriguing statement of this commonsense position can be found, for example, in the work of Bretherton and colleagues (1981, p. 336, my emphasis):

> the emergence of these behaviours [involving coordinated visual attention] is difficult to explain unless one assumes that the child has come to *recognize the psychological similarity* as well as the *separateness of self and other*...and is operating within a *shared meaning system*...

However, although Bretherton et al. seem to assume that shared meaning and psychological "recognition" and distance must be (causally) in place in order to empirically witness joint attention in the activities of a 12-month-old infant, they do not adduce what might logically justify such a claim. It surely cannot be that a baby looks in the same place as another or that she points in the presence of another because these activities would only satisfy Butterworth's definition of joint attention which makes no claims about psychological similarity and the like.[14] And if we are to look more carefully at Bretherton et al.'s claim, does the "recognition" of psychological similarity between self and other not presuppose the recognition of the separateness of self and other? And is

[14] On empirical grounds alone though one should be deeply suspicious about these claims when made about a 12-month-old infant. For example, there is convergent evidence that infants younger than about 18 months do not fully conceive of themselves as a separate object in space (e.g., Butterworth & Jarrett, 1991; Morissette et al., 1995).

the ability to "recognize" psychological similarity to another not parasitic on a shared meaning system? And what exactly is the epistemic status of the psychological predicate "recognition?" It would seem that Bretherton et al.'s conclusion raises as many questions as it answers.

Tomasello also seems to trade on the ambiguity of the term "recognition" when claiming that 8- to 9-month-old infants "recognize that they have goals that are clearly separated from behavioural means" (Tomasello, 1999a, p. 69). Given the private language arguments, I presume he does not mean that infants literally introspect on their mental contents and *recognize* that the mental states they see in their "mind's eye" are a particular kind of mental state called an intention. However, if he doesn't mean that, then what exactly is he claiming? And, as I have noted elsewhere (Racine, 2004), isn't his theory actually dependent on such a view of mental life for it work? Perhaps the unstated presumption is that "recognition" is cognitively weaker than "understanding," which in everyday language it surely is -- despite the fact that Carpenter et al. (1998) use "understand" and Tomasello (1995) uses "know." However, Tomasello would seem to need infants to understand or know rather than recognize though because if infants do not understand their separateness from or psychological similarity to others (i.e., know that they are separate from others) but merely *recognize* this, it is not clear how or why they would act on this information.

Paradigmatically, we have grounds for saying that someone knows something because she has told us something that expressed this knowledge. However, young

infants and other non-language using agents by definition do not speak.[15] We also often attribute knowledge to a person through a demonstration of some sort, such as in the case of the skilful chess player who in her mastery of the game demonstrates that she indeed knows its rules. However, the rules required to correctly play chess are embedded in and acquired through linguistic practice. And if concepts are abstracted from language use, we cannot, on pain of circularity, attribute conceptual understandings to non-language using agents. To claim that infants *know* that they are attending to something in common with another is to claim that they have this concept.[16]

I hope it is clear that that I am not advocating some form of radical parsimony or scepticism on the ostensible grounds that we cannot be sure what non-language using agents understand of their activities. No, I am noting that non-language using agents by definition cannot be aware that they are jointly attending without grasping the meaning of joint attention. Now, as long as we remember that it is just a paraphrase of a satisfying the application of a given set of criteria, there is no harm in claiming that certain non-language using agents have a "practical understanding" of the activities that justify the use of psychological predicates -- i.e., a "non-reflective understanding" that is grounded in practice (which is just a paraphrase of the fact that they cannot be said to have the concept) -- the claims being made about non-language using agents are typically of a

[15] I will not discuss at present the interesting case of enculturated apes that can use a communication system, but I direct interested readers to Canfield (1995) and Savage-Rumbaugh et al. (1998).

[16] This is not to say of course that some agents cannot be said to know more than others about a given concept. My concern here is simply to point out the deeply linguistic nature of concept acquistion and accordingly the incoherence of attributing a conceptual grasp of mind to non-language using agents.

reflective/conceptual rather than participatory nature.[17] But this can easily slip into obscurity and researchers can forget that it is we who understand the concepts, not the agent so observed. In any case, applying Tomasello's definition of joint attention to non-language using agents is a non-starter. And Bretherton's rich characterization of what an infant understands about mental states is in need of clarification and qualification. For example, it is clear that young infants interact with other persons differently than they do with non-social objects. Do we really mean something different or additional to this when we claim that infants recognize their psychological similarity to others? I think not.

However, although Butterworth's definition is useful, it is what we might call psychologically minimalist and does not distinguish the shared engagement I might have with a dog and a stick from the conceptual understandings entailed in Tomasello's definition. However, Tomasello himself does not seem to fully appreciate the commitments in his definition and has advanced an explanation of joint attention that is couched in commonsense, or what I will more precisely call causal-psychological-representational, terms (Carpendale & Lewis, 2004; Carpendale et al., in press; Corkum & Moore, 1998; Frye, 1991; Fogel, 1993; Kaye, 1982; Moore & Corkum, 1994; Racine, 2004; Racine & Carpendale, in press). Rather than activities revealing a prior psychological competency, activity can be seen to create psychological competence -- that is, it can license the applicability of particular psychological predicates. And although the Butterworth 'simply looking' definition of joint attention avoids Tomasello's mentalism, it does so by glossing over the critical distinction between the ways of acting

[17] This might be analogous to what theorists mean by "recognition." But the key point here is that these observer claims are parasitic on observer conceptual understandings and not the infant's. Thus, although the infant becomes involved in progressively more complex activities with others that might justify more refined uses of psychological terms, it is still we, the observers, who have the concept.

that define joint attention versus the understanding of such a definition by an agent so described.

An issue that ties Butterworth's and Tomasello's definitions together is that they are both predicated on what Bickhard (1993) has called an observer semantics. That is, whereas observers can discern that the infant is engaged in coordinated meaningful activities with others, whether the infant herself is making these distinctions remains to be established. (From my point of view this results from the fact that the terms being used in these definitions have not been conceptually analyzed.) Piaget (1954) discussed this as the psychologist's fallacy, and in a related vein, pragmatists such as Dewey (1929/1958) alert us to what they call "spectator theories of knowledge." In fact, Bretherton et al. are also ambiguous as to whether this "shared meaning system" is shared from the infants' or from an observer's point of view. What we need to know though is what is psychologically meaningful from the infant's point of view. And, although one could argue that there are forms, or levels, of shared meaning which it might make sense to say exist between parent and infant and even between non-language using animals, the "level of meaning" entailed in Tomasello's (1995) definition or in Bretherton and colleagues (1981) rich description underscores what I take to be the radically social ontology of what we call "the mind." As George Herbert Mead (1964, p. 194) has put it,

> Mentality on our approach simply comes in when the organism is able to point out meanings to others and to himself. This is the point at which mind appears, or if you like, emerges....The ability to pick these meanings out and to indicate them to others and to the organism is an ability which gives particular power to the human individual. The control has been made possible by language. It is that mechanism of control over meaning in this sense which has, I say, constituted what we term 'mind'.

Without language, an agent cannot point out meanings to self and other, cannot appreciate their psychological similarity to another, cannot have a differentiated self and other. And like Wittgenstein, Mead affords an ontogenetic primacy to activity (Carpendale & Racine, 2005). Before language, in both the phylogenetic and ontogenetic sense, humans have natural reactions and the patterns of activity into which they are thrown. Language affords humans a way to coordinate activity, including activities that are no longer present -- that is, to conceptualize them. Activity is sufficient for observers who have mastered the relevant language-game to apply certain concepts to others' behaviour: "What is the natural expression of an intention? -- Look at a cat when it stalks a bird; or a beast when it wants to escape" (Wittgenstein, 1958, § 647). However, it is persons, not cats, who understand their own intentions or the intentions of others. Persons alone can be said to have the concept in question. And to understand one's own or another's attention presupposes an understanding of the concept attention. That is, to know the criteria that justify the use of this term in particular situations -- a salient one of which is first-person sincere verbal expression of what one is up to. Attention is simply a way of talking about particular patterns of activity that is grounded in the natural reactions of many forms of life. And, whatever else might be causally required for the developing agent to understand mental life, the capacity for humans to understand psychological states, is a *massively* linguistic capacity:

> But what is necessary here is *not* probing deeper. There are no depths here any more than in a mirror; or rather, the only depths are depths of confusion. The neurophysiologist *may* uncover neural correlates of *experiences* (but believing, understanding, thinking, etc. are not experiences) or neural preconditions of the possession of certain capacities. But what he discovers are *at most* correlates of, or structural

> preconditions for, wanting, feeling, thinking, believing, understanding, etc. (Baker & Hacker, 1982, p. 233, their emphasis).[18]

2.4 The Commonsense Picture of Mind

As is often the case in work on preschoolers' social understanding (so-called "theory of mind" research), much is of made of "propositional attitudes" in work on infant social understanding (e.g., Barresi & Moore, 1996; Baron-Cohen, 1995; Hobson, 1994; Moore, 1996b; Wellman et al., 2004). Despite the heuristic usefulness of using this term, clarifying what is entailed in such a proposal is illustrative of this commonsense picture of mind that seems at work in much social development theory. To use an example from Racine and Carpendale (in press), it is commonly held that Joan having an attitude of belief towards the proposition "it might rain" is not only a description of what might be the case for a particular person named Joan, but is also a correct description of Joan's mental process of believing that it might rain. Namely, that Joan has a particular mental state (belief) about a particular contingency (that it might rain) that accordingly might cause her to bring an umbrella along to work. This is assumed to wholly constitute and define her belief and Joan is understood to represent her psychological attitude (a mental state) towards this proposition (a mental content) in order to bring the umbrella to work. Although this may be all well and good in ordinary language use, researchers often seem to forget that this is a metaphorical view of mental activity. Agents are not mentally related to propositions, but rather to what the supposed propositions are about (Bennett & Hacker, 2003; Hacker, 1990, 1996).

[18] I obviously do not mean "linguistic capacity" in the Chomskyan sense. I mean it in the Wittgensteinian sense which points out the dependency of meaning on both language and rule-governed activity.

Although observers of agents can describe beliefs in such a way, taking this metaphor literally results in a false picture of mental states and contents (Bickhard, 2001, 2004; Racine & Carpendale, in press; Sharrock & Coulter, 2004).

Social developmental researchers' typical conceptions of mental states and contents are characteristic of a commonsense causal-psychological-representational (CPR) view of meaning and mind according to which: (a) mental states underlie and cause behaviour, (b) mental representations are mental objects with semantic properties, (c) mental states are a series of relations between agents and their mental representations (propositional or otherwise), and (d) what differentiates an intention from, for example, a belief is that there is a characteristic mental state that occurs when one intends to *x* that does not occur when one believes *y*. Multifarious problems have been identified with this picture of meaning and mind from a diversity of perspectives (see e.g., Baker & Hacker, 1985; Bennett & Hacker, 2003; Button, Coulter, Lee, & Sharrock, 1995; Carpendale & Lewis, 2004; Carpendale et al., in press; Coulter, 1991, 1999; Fogel, 1993; Hacker, 1990, 1996; Garfinkel, 1967; Kusch, 1997; Leudar & Costall, 2004; Martin & Sugarman, 1999; Martin, Sugarman, & Thompson, 2003; Racine, 2002; 2004; Racine & Carpendale, in press; Reddy & Morris, 2004; Savage-Rumbaugh et al., 1998; Shanker, 1998, 2004; Sharrock & Coulter, 2004; Sinha, 1999; ter Hark, 1990; Turnbull, 2003; Turnbull & Carpendale, 1999; Wittgenstein, 1958, 1969, 1980a, 1980b, 1992; Wootton, 1997), and my goal in this dissertation will be: (a) to trace this way of thinking across social developmental theorists who work in very different camps, (b) to show that such problems do not come up in a Wittgensteinian psychology, and (c) to interpret data in light of these two competing pictures of mind. Thus, in contrast to the

disembodied, dualistic, CPR view of mental states I will argue instead for an embodied view of mental life. By embodied I do not mean an emphasis on isolated human bodies, but rather with human embodied experience which is irreducibly social. It is this embodied cultural and linguistic social experience that is the foundation for the development of children's understanding of the mind (Carpendale et al., in press; Racine & Carpendale, in press). And "mental states" are defined and instantiated in activity and have no separable life outside of it.[19] And not only do mental states not play some mediating causal role in our understanding of other persons, to claim, save metaphorically, that they do is logically incoherent because:

> A broader and deeper inspection of the grammars of our concepts of the mental (having a thought, understanding, remembering, etc.) reveals that they all have ineliminable connections with what we say and do in various sorts of circumstances in the social world. The reification and interior projection of the putative 'referents' of these concepts, a practice which is in turn nourished by the kind of theorizing which is erected upon a deep misunderstanding of their grammars, involves *severing* these constitutive connections with public circumstances and criteria (Coulter, 1991, p. 189, his emphasis).

[19] Physical and mental are different ways of talking about the activity of animate beings -- not different ways of talking about parts of animate beings.

CHAPTER 3 THEORIES OF JOINT ATTENTION

What is minimally causally involved for a human infant to engage in sequential activities involving joint attention with another person? At root, encoded in the human genome there seems to be a potential to develop an ability to engage in joint activities requiring attention (Baron-Cohen, 1995, 1999; Bruner, 1983; German & Leslie, 2001; Tomasello, 1999b). But the degree to which the potential to develop this capacity is uniquely human is controversial and turns on how these activities are defined. To my mind, pace Tomasello (1999b) apes have the capacity to jointly attend, but they do not understand the concepts that are tied up with more mentalistic interpretations of joint attention. Where human and non-human primates may also differ is in the role joint attentive practices play in their respective forms of life. For example, many agree that the great apes, like children with severe autism, do not employ protodeclarative points (but see Savage-Rumbaugh et al., 1998). Perhaps then great apes, like children with autism, have no use for this ostensibly ostensive gesture. However, this again turns on how we make sense of this gesture (Leavens et al., 2005). Second, there is an affective-motivational component on the part of the infant that comes into play that seems also to be in some sense biologically based (Fogel, 1993; Hobson, 1993, 2002; Kaye, 1982; Reddy, 2001, 2003; Trevarthen, 1979; Trevarthen & Aitken, 2001; Trevarthen & Hubley, 1978). Another way of saying this perhaps is that the infant must be interested in other people. Third, the infant can be said to begin to distinguish self and other through the coordination of self-object and other-object spatial relations (Barresi & Moore, 1996; Carpendale & Lewis, 2004; Frye, 1991; Müller & Carpendale, 2004). That is, the infant needs to experience their own perspective and that of others

(cf. Mead, 1964). This seems to develop apace with decentration and the appreciation of another as an independent centre of causation (Piaget, 1952, 1954). But this is no more a property of infant's internal cognitive development anymore than are the settings in which the infant displays such "psychological structures." Fourth, the infant has engaged in progressively more complicated forms of sequential activity with a conspecific. And in so doing has come to have sequential expectations in relation to the other party around some basic human practice (Canfield, 1993, 1995; Carpendale et al., in press; Kaye, 1982; Racine, 2004; Racine & Carpendale, in press; Turnbull, 2003; Turnbull & Carpendale, 1999; Wootton, 1994, 1997). Although the roughly biological, cognitive and social theories of joint attention that I presently review emphasize different aspects of this characterization, these aspects minimally enter into an infant's ability to engage in such interactions.[20] However, the glue that seems to hold most of these theories together is their vulnerability to Wittgenstein's private language arguments and their reliance to a larger or a lesser degree on a CPR picture of mind.

3.1 Biologically Driven Theories

3.1.1 Starting State Nativism

Meltzoff and his colleagues (Gopnik & Meltzoff, 1997; Meltzoff & Brooks, 2001; Meltzoff, Gopnik, & Repacholi, 1999; Meltzoff & Moore, 1995) suggest that the grasp of

[20] I will not systematically review the dynamic systems approaches to joint attention (e.g., Fogel, 1993; Fogel & Hannan, 1985; Fogel & Thelen, 1987; Kaye, 1982). These emphasize developmental continuities in the first year of life and as such root the activities that are said to involve joint attention in much simpler dyadic exchanges. Because of this, this approach has little to say about joint attention that would make for any kind of a one-to-one counterpoint with the theories reviewed below.

mental life rests on an innate mechanism for social learning, an ability to represent experience in the first months of life, and an amodal representational body scheme that allows infants to represent and compare perceptual inputs to proprioceptive outputs. The links between self and other are provided by imitation of others' actions and expressive behaviour (see also Gallese & Goldman, 1998). The imitation of others' behaviour is said to naturally lead the infant to experience the concomitant emotion states that are tied up with the behavioural expression of emotion. This form of reasoning is known as the argument from analogy (e.g., Russell, 1921). The success of this account rests in large part on the interpretability of Meltzoff and Moore's (e.g., 1977) well known, but controversial, neonatal imitation research and the validity of the analogical argument. And the twin lynchpins of this theory are both potentially problematic.

In the first instance, infants may only reliably match tongue protrusion at birth, but not mouth opening or lip pursing (Anisfeld, 1991, 1996). Given that one rather than a range of behaviours may be what the infant is "matching," simpler explanations become more tenable. To this end, shortly after Meltzoff and Moore published their landmark work, Jacobson (1979) showed that infants would respond as much to a black pen being moved back and forth from the baby as they would to an adult model. Jones (1996) has suggested that tongue protrusion may not be an intentional matching of an adult-modelled act but rather simple exploration of the environment by the infant in order to find, for example, a nipple from which to nurse. Given that this changes the issue to stimulus rather than cognitive properties, the adult and infant's seemingly coordinated activities may only seem imitative but are instead incidental and non-imitative. However,

even if neonatal imitation actually occurs, Kaye (1982) observed some time ago that a low-level solution is much more plausible than the high-level interpretation that Meltzoff and Moore give. And Myowa-Yamakoshi, Tomonaga, Tanaka, and Matsuzawa (2004) have also recently reported evidence of neonatal imitation in chimpanzees.

Second, the argument from analogy is a way of "solving" the problem of other minds, that is, of bridging the gap between one's own mind and that of another. Essentially the argument entails that because an infant perceives others as similar to the self (possibly established in this case by Meltzoff and Moore's empirical work on neonatal imitation), the infant will credit others with the same psychological experiences the infant experiences when others are doing similar activities. However, when I am happy I am not reading off or describing some mental state called happiness that I can reason about in relation some event it has caused such as, for example, a smile. Rather my smile is an expression of my happiness -- constitutive of it. And your or my understanding that my activity communicates happiness presupposes a community of language users, a particular linguistic practice and that expressions of joy play some role in our form of life. As Müller and his colleagues have also observed (Carpendale et al., in press; Müller & Carpendale, 2004; Müller & Runions, 2003), even if Meltzoff et al.'s way of thinking about mental states were coherent, at best the analogical argument would lead to the conclusion "there goes one of my mental states again" for the reason that the distinction between self and other has been smuggled into the argument even though the argument was advanced in order to explain self/other differentiation.

However, nativist accounts of social development seemed to have been given a boost by the recent discovery of so-called mirror neurons in macaque monkeys (Gallese

& Goldman, 1998). These are a class of motor neurons that are preferentially activated when macaques perform an action that is similar to that performed by a conspecific. The presence of these neurons has also been confirmed in humans and it has therefore been argued that this suggests a neural basis for intersubjectivity (Gallese, 2005). Be that as it may, there is no evidence that the mirror system is innately specified. And even if such a system were to exist in older children this would not in and of itself make Meltzoff and colleagues' theory more (or less) plausible. The problem is that the role of these types of correlative externally related phenomena is easily misunderstood in general. What would it mean, for example, if an fMRI revealed that a particular part of the cortex is differentially active when a child is engaged in intentional activity? It would be the simple identification of a *possible* structural precondition for intentional behaviour. The apparent homology between "the mental state of intention" and an underlying brain state is I have argued also a mistaken one. And if through some unfortunate accident a crucial area of the prefrontal cortex were lesioned in a person and that person lost the capacity for goal-directed activity what would this tell about how children learn about or understand the mind? And what if as a consequence of such a neurological event the afflicted person no longer hoped? (I do not mean in the sense that they were depressed.) Would that mean that the area responsible for hope was lesioned in the process?

3.1.2 Modular Maturation

Simon Baron-Cohen (1995, 1999) builds upon Alan Leslie's (1987, 1994; German & Leslie, 2001) work by accepting Leslie's suggestion that a Theory of Mind Module (ToMM) comes on line towards the end of the second year of life that is

responsible for coding the mental relations that are said to occur between agents and propositions. Baron-Cohen argues that an Intentionality Detector (ID), Eye-Direction Detector (EDD), and Shared Attention Mechanism (SAM) would however precede it. Baron-Cohen's claim is that early in life EDD and ID are jointly responsible for constructing dyadic intentional relations between self/other and some object -- ID processes the directedness of the agent's behaviour, EDD processes eye movement in particular. Later in the first year of life SAM comes on line and codes the fact that self and other are attending to the same event by using lower level input from EDD or ID to generate triadic representations. The developing agent is said to understand propositional attitudes because ToMM adds the requisite psychological attitude to the lower level representations generated by SAM. Despite the fact that Baron-Cohen (1995) is careful to distance himself from claims of extreme modularity by instead wanting to call his modules "mechanisms," his neurocognitive circuits satisfy all of Fodor's (1983) classical criteria (Moore 1996a).

Moore (1996a), however, wonders what specific adaptive problems these modules could have evolved to solve when "what is most remarkable about human psychological processes is the virtually unlimited breadth of their applicability" (Moore, 1996a, pp. 614-615). He also points out that the fact that in gaze following infants first rely on head turns and postural cues before they rely on eye direction problematizes the function of the EDD module. Tomasello (1999b) has argued there is not enough evolutionary time to have evolved a series of modules selected for their capacity to solve particular adaptive problems. He argues instead that it is more reasonable to look for one adaptation rather than several. And that adaptation is reasonably said by Tomasello

to be a capacity for joint attention. Baron-Cohen (1995, p. 24) also claims "there really are mental states inside our own and other organisms' heads." I have instead characterized mental states as embodied in a multiplicity of social situations -- some of which do not yet exist. If I am right, this creates the rather monolithic problem of how and why evolution would favour such a storehouse of potentially irrelevant facts. However, if Baron-Cohen is right, he needs to explain how it is that humans could have evolved what I can only take to be "non-material entities," a question that is independent of the modules that might subserve such states.

Now, we should pause and underscore the fact that clearly biological factors play a critical causal role in infant's dawning appreciation of others' mental lives. For example, infants seem to be biologically primed to display early preferences for attending to social stimuli. The logic of Baron-Cohen's approach, however, goes beyond biological preparedness and ultimately entails that particular mental states could be computed by evolved algorithms. This again places one of the key assumptions in the CPR picture of mental states directly on the table. Namely, that it makes sense to conceive of beliefs or intentions as discrete attributes of mind that can be directly computed because their meanings are fixed. This is a CPR view of mental states because it is the internal machinery that is said to discriminate "mental life" rather than this occurring through our bodily engagements in a world of social activity. The problem here is that the social situation in which a gesture or a mental state utterance takes place determines its meaning -- and this is in an ongoing sequence of interaction. And although meaning is indexical (i.e., dependent on a particular social situation of use), algorithms at least at present cannot be (Racine, 2002; Shanker, 1998). Fixed a priori

meanings, however, are not possible from an embodied view of mental states such as the one presently advocated. Further, if one wants an example of communication systems in which meaning is fixed by evolutionary pressures, that is, of non-indexical meaning, looking at infants is problematic. Better examples would be found in the calls of vervet monkeys or the dances of honeybees because these seem to have fixed meanings indicating particular predators such as snakes, eagles or leopards or directions to new sources of nectar (Racine & Carpendale, in press). However, despite a fixed morphology the gestures that infants use such as pointing can convey different meaning to an observer (Bates et al., 1975; Liszkowski et al, in press, 2004) as a function of the social situation in which they are employed.

3.1.3 An Evolutionary-Ecological Approach

Butterworth and his colleagues (1998b, 2003; Butterworth & Jarrett, 1991; Butterworth & Morissette, 1996; Franco & Butterworth, 1996) have focused on accounting for the origins of two activities involving joint attention (gaze following, pointing) and in so doing have offered an account of their continuity with language acquisition. This is an evolutionary-ecological approach because evolutionary mechanisms are said to play a role in pointing production -- and indeed that pointing is claimed to have a specialized communicative function shaped by evolutionary pressures (Butterworth, 1998b, 2003) -- while at the same time much of the account relies on simple ecological principles that are perhaps not as contentious. Butterworth et al.'s description of gaze following, for example, nicely captures how an infant might find herself in such episodes of shared attention (Moore & Corkum, 1994; Corkum & Moore,

1995, 1998) and how these might form the basis for other joint activities requiring attention.

To explain the emergence of gaze following, Butterworth and Jarrett (1991) postulated the existence of three separate mechanisms (ecological, geometric, representational). They argue that an ecological mechanism develops first (at about 6 months) that allows the agent to select a correct referent of two possible objects on the same side of a room if and only if the object possesses particular stimulus attributes (e.g., motion, brightness contrast) that would draw the infant's attention in the absence of another's gaze. The geometric mechanism (in place at about 12 months) is so named because the infant is said to infer and project a tangent from the line of sight of the other person to the referent. Butterworth claims that this capacity allows infants to both understand and produce pointing gestures because both are said to require extrapolation of a vector into space.[21] At about 18 months, a representational mechanism allows the infant to operate in not just a visual, but also an auditory, space located behind the agent. This is said to explain why infants are able to follow gaze to referents in space behind them despite being unable at 12 months.

Butterworth's description of forms of gaze following is compelling, and it sits well with demonstrations of forms of pointing (e.g., Carpenter et al., 1998; Franco & Butterworth, 1996; Moore & D'Entremont, 2001; Morissette et al., 1995). However, rather than offering up his mechanisms of gaze following as descriptions (or even definitions) of activity, Butterworth and colleagues seem to conceive of these as causal

mechanisms proper and focus upon the presumed cognitive developments that underlie the infant's abilities. Although I question the neat cause-effect relationship presumed by Butterworth et al., such developments do of course occur. However, far from internal mechanisms causing an overt behaviour, from my point of view "overt behaviours" are part of a matrix from which an infant's (often presumed) "internal competencies" cannot be separated. Further, Butterworth's account of pointing is complicated by the fact that he believes pointing emerges from gaze following, yet is also a species typical manual gesture selected for by evolution to serve a communicative function and which also primes the infant's understanding of mind. For example, Butterworth and colleagues state that pointing:

> is a species typical gesture for humans whose evolution is tied to both the instrumental functions of the hand and mind...and the socialization of internal states. (Franco & Butterworth, 1996, p. 309)

Butterworth's (1998b, 2003) argument that the shape of the manual gesture in human ontogeny may have evolved (or co-evolved) out of concerns related to fine motor control is plausible (see also Povinelli et al., 2003). However, if the fact that humans point in a species typical manner reflects an a priori genetic endowment, this would not give them some leg hold on other minds.[22] Butterworth knows this given that he claims (2003) that while the pointing gesture is evolved, it still requires a communicative context in which to be deployed. But this seems to trivialize activity and in a manner analogous to Chomsky's (1995) most recent theory of language acquisition reduces the role of "the

[21] Butterworth (2003) notes that in cluttered (i.e., typical) environments, successful "understanding" of pointing is said to require stimulus cues for even adults to follow pointing. Ecological and geometric mechanisms are therefore said to interact. One might wonder then just how neat a line should be drawn between these two 'mechanisms' if the line can only be drawn in a laboratory setting under very restricted conditions.

context" to a mere trigger. A species typical hand shape does not replace the need for a shared practice within which to communicate, but it rather presupposes such a practice (Goodwin, 2003; Schlegloff, 1984; Wilkins, 2003). And pointing, as an activity, is clearly more than the unfurling of some fixed action pattern. As Goodwin (2003, p. 218) puts it,

> Pointing is not a simple act, a way of picking out things in the world that avoids the complexities of formulating a scene through language or other semiotic systems, but is instead an action that can only be successfully performed by tying the act of pointing to the construals of entities and events provided by other meaning making resources as participants work to carry out courses of collaborative action with one another."

3.2 Cognitively Driven Theories

3.2.1 Secondary Intersubjectivity

Trevarthen and his colleagues (e.g., Trevarthen, 1979; Trevarthen & Hubley, 1978; Trevarthen & Aitken, 2001) suggest that infants are born with a conscious awareness of the psychological states of others -- a so-called initial psychosocial state -- and a powerful motive to attend to this awareness. He argues that at even two months of age, infants can understand another's communicative intentions and respond to these. For this reason, Trevarthen terms this primordial state "primary intersubjectivity," which is said to provide the infant with an a priori means and motive with which to participate in the extant dyadic relational structures in which infants find themselves.[23] This is said to be because infants have adult-like social cognition but their ability to

[22] It is also relevant here that language trained chimpanzees typically point with their index finger rather than the open hand point that is more common for non-language trained apes (Leavens & Hopkins, 1999).

[23] Intersubjectivity is a term that roughly means "shared meaning," but this term is used in a variety of ways. Reimers and Fogel (1992, p. 82) conceive of intersubjectivity as "a shared understanding of what an interaction is about," and this is a usage upon which I will capitalize in Study Two.

express and act upon this understanding is limited by maturational factors, which as we have seen is a standard nativist solution to the "problem of other minds." Trevarthen suggests that towards the end of the infant's first year of life, curiosity about her caretaker's attention and intentions leads the developing agent into a state of "secondary intersubjectivity" (i.e., an ability to engage in joint activities requiring attention) which is said to explain how an infant finds herself in the midst of triadic relational structures. Despite its nativism, then, this is a strongly psychologically mediated account of joint attention, which is why I have categorized it as a cognitively driven account.

Trevarthen is to be commended for emphasizing the role of affective arousal and motivation in joint attention, as it is often overlooked. Infants are actively emotionally engaged in their world, and are motivated to explore and interact with it (Hobson, 1993, 2002; Reddy, 2001, 2003). Their intentions are well known to others quite early in their life. However, despite the fact that they react to others with well-differentiated emotional responses, they do not understand others' intentions at such an early age (Moore, 1996b; Racine, 2004; Racine & Carpendale, in press). This seems to commit Trevarthen to a theory of neonatal reflection despite the fact that by most accounts the infant is experiencing a form of primary (non-reflective) consciousness. As a consequence, Trevarthen inserts underlying competencies into infant behaviours that are seen to drive the emergence of the behaviour in question. For example, protoconversations, which to my mind evidence the child's increasing ability to coordinate their interactions with others (and which in so doing might license Trevarthen to talk about their intentions) are offered up as support for the notion that infants understand others' intentions. However, although I do not agree with Trevarthen's

solution, he avoids the problematic of the argument by analogy by at least rooting intersubjectivity in the genome. Many accounts do not avoid this problematic, including the by now familiar one to which we shall turn next.

3.2.2 Intentional Insight

As noted earlier, Tomasello (1995, 1999b) has advanced a theory that is meant to account for the infant's participation in *all* joint activities that require attention. His argument is that when infants come to understand that their actions cause things to happen in the world they do not at first apprehend how they do this. However, through continuing to act upon their world they come to distinguish means and ends and in so doing they are said to experience their intentional (goal-directed) mental states. That is, infants of 8-9 months of age (i.e., Piaget's sensorimotor stage IV) "recognize that they have goals that are clearly separated from behavioural means" (Tomasello, 1999a, p. 69). Because they perceive others as 'like me', they automatically apply this recognition of their own intentionality on analogy to the behaviour of others (Meltzoff & Moore, 1977). In so doing they deduce that others also act upon their intentional states. This (automatic) insight into others' intentions therefore enables the multifarious "manifestations of joint attention."

Now, more recently, Tomasello et al. (in press) have rejected simple identification with others based on neonatal imitation and they speculate instead that intersubjectivity must depend on the forms of emotional interpersonal engagement proposed, for example, by Hobson (2002). Irrespective, the success of this account minimally rests on the demonstration of a tight synchrony of onset of the various joint attention behaviours (which Slaughter & McConnell, 2003 do not find), the plausibility of infants being able to

recognize their goal-directed states, as opposed to just employing different behavioural strategies in response to fluctuating environmental contingencies (Moore, 1996b; Racine, 2004), and third again on the validity of the analogical argument (and a fortiori on the invalidity of the private language argument).[24]

In contrast to Tomasello, although Piaget (1954) has demonstrated the goal-directedness of 8-month-old infants as an emerging skill at coordinating means-ends relations, he would have likely had difficultly with Tomasello's analogical solution because infants "[do] not yet recognize positions and displacements as being relative to one another, but only as relative to [themselves]" (Piaget, 1954, p. 183). However, it would seem that what distinguishes sensorimotor stage III and IV from Tomasello's point of view is the arrival of a capacity to experience the mental events that cause infants to act intentionally, that is, the experience of a mental state called an intention. Further, Tomasello's account is predicated on a literal internal causal mental entity called an intention (Racine, 2004; Racine & Carpendale, in press). However, as the private language arguments show, an infant or young child cannot introspect upon her mental world to conclude she is experiencing an intention because to do so presupposes that she already is familiar with the meaning of an intention. Thus, although human infants and members of many other species act intentionally -- that is, it makes sense to apply the concept 'intention' to their activities -- we have to be careful about assumptions about what this means for the child's understanding. This only shows what could be at best called a sensorimotor or practical competence in interacting with others and

[24] Although not offering a full-fledged account of joint attention, Bruner (1983) also claims an infant has an a priori "intent to refer" (cf. Trevarthen, 1979) and an ability to recognize "that intent in others" via analogy (cf. Tomasello, 1995, 1999b). This is Bruner's solution to the fact that "some basis for referential intersubjectivity must exist before language proper appears" (p. 122).

responding to their directedness toward aspects of the world and people's efforts to do things. But understanding in the sense of thinking about intentions depends on the ability to talk about activity (Carpendale & Lewis, 2004; Racine & Carpendale, in press).

It is worth pausing to consider the difference between Tomasello et al.'s (in press) proposal and the present one. The difference can be clarified by considering Tomasello's statement "what could it mean to say that language is responsible for understanding and sharing intentions, when in fact the idea of linguistic communication without these underlying skills is incoherent" (p. 27). The difference again turns on what is meant by "understanding." It may be correct to state that "language is not basic, it is derived. It rests on the same underlying cognitive and social skills that lead infants to point to things and show things to other people declaratively and informatively..." (Tomasello et al., in press, p. 27). However, the question is how should these skills be characterized? If the word "understanding" is used, and this is what seems to be aimed at in Tomasello's claim that infants recognize their intentions, then it is a practical understanding (Racine & Carpendale, in press). As Hobson (2002) puts it, it is understanding in action or feeling, not reflective understanding. And it is certainly not what we mean in everyday language use by "understanding." However, it is competence in these forms of shared activity that is the ground upon which language is built. Language is required to move from a practical or lived competence in interacting with others based on expectations about others within shared practices to an ability to conceptualize the intentions that are defined in patterns of shared activity (Carpendale et al, in press; Racine & Carpendale, in press). And whereas the foregoing may be the case, we must also not lose sight of the fact that the above considerations are parasitic

on the lack of applicability of the concept understanding to non-language using agents. "Internal practical understandings," soon to be ostensibly replaced by internal conceptual understandings, are not what render activity intelligible. Practical understandings are nothing more (or less) than capabilities presupposed by the judicious application of mental state concepts.

Differing views of intentions can be further clarified by unpacking the ambiguity in the following claim:

> Understanding intentions is foundational because it provides the interpretive matrix for deciding precisely what it is that someone is doing in the first place. Thus, the very same physical movement may be seen as giving an object, sharing it, loaning it, moving it, getting rid of it, returning it, trading it, selling it, and on and on -- *depending on the goals and intentions of the actor* (Tomasello et al., in press, p. 2, my emphasis).

Although this might be true in an everyday sense in that what we call intentions are defined and partially constituted in shared activities such as the ones listed by Tomasello and colleagues, another way to interpret their position is that to grasp the meaning of an action involves understanding the mental state that gave rise to it (Russell, 1921). This is a restatement of the commonsense CPR view of (human) activity. From a Wittgensteinian perspective the meanings of the multifarious practices listed in the above example are not defined by the mental state of an actor, but by the mastery of these self-same practices. "Intention" in fact is one such practice. To say an agent has acted intentionally is to understand the concept intention as situated in particular activities. And persons see an agent's activities as intentional in and through the practices that Tomasello and colleagues list. Thus, the way of thinking that seems to manifest itself in Tomasello et al.'s work is one that "takes behaviour to be just bodily

movement and so strips it of its intentionality, relocating all that is alive and intelligent in the hidden mind" (Leudar & Costall, 2004, p. 603).

3.3 Socially Driven Theories

3.3.1 Mental State Matching

In contradistinction to the above classes of approach, Moore and his colleagues (e.g., Barresi & Moore, 1996; Corkum & Moore, 1995, 1998; Moore & Corkum, 1994; Moore & D'Entremont, 2001) adduce evidence that infants do not, for example, understand pointing until approximately 18 months of age. Their account proceeds roughly as follows. Moore and Corkum (1994) claim that various mechanisms are implicated in leading the infant to engage in shared activity with others and to therefore match attentional states with them. On this theory, such activities do not require social understanding; rather, they facilitate it. An instance of what Moore and colleagues conceive of as matching of attentional states would be set up, for example, by an infant learning that following a parent's head turn results in seeing interesting things. In this way, when the infant follows parental head turns the infant will find rewarding objects or events in their line of sight (Corkum & Moore, 1995, 1998). Thus, without recourse to strong bottom-up nativism or top-down cognitivism Moore and colleagues can suggest that a simple conditioning explanation explains how an activity involving joint attention can be set up. Moore et al. can also account for other aspects of joint attention (e.g., social referencing) in the same manner. What Moore et al. claim is important about setting up such coordinated activities is that they are said to couple the infant's first

person psychological experience with a third person experience of others' psychological states, which ultimately affords the infant an understanding of self-other equivalence. The final component in their model is their claim that infants do not understand mental life outside of particular interactive contexts until they are able to construct multiple models of self, other and some state of affairs. It is the development of this meta-representational capacity that is said to confer upon infants an understanding of joint attention (Perner, 1991).

I agree that once an infant finds herself in this and other similar shared activities an opportunity has been afforded for the developing infant to learn about the mental life of others. And I clearly endorse Moore's (2005) recent criticism of Tomasello's theory on the grounds that "mental states are embodied in infancy." Where I depart with Moore is that I believe that mental states are embodied. Full stop. To understand why this is so, it is instructive to pause and consider an ambiguity in Moore and colleagues' account. Although it makes sense to claim an infant can coordinate attention -- a legitimate mental state with a clear duration and other statelike attributes -- with others, what might work for joint attention could not work for intention, desire, belief and so on. As Wittgenstein remarked, "An intention is embedded in its situation. If the technique of chess did not exist, I could not intend to play a game of chess" (Wittgenstein, 1958, §337). What does it mean to claim that an intention is "embedded in its situation?" Consider the following example taken from Racine (2004). My moving forward to stand at a bus stop as a bus pulls closer is an expression of an intention to get on the bus for which the criterion lies in mastery of the practice of waiting for a bus. Whether some neurons fired or some psychological events occurred during the execution of this

intention is beside the point: If this particular practice did not exist and if this skill had not been mastered, these activities would not have partially constituted this intention. That is, these activities would not mean that a person had acted intentionally (Shanker, 1991). Similarly, we could see attention in that person's activity as well, for example when they may have scanned the vista to see if a bus was actually coming.

Wittgenstein's remark shows that children need not "peer into other people's minds" to understand their attention, intention, desire or belief but rather need to see these concepts in activity. To put it differently, *this is* peering into another's mind. Another way to bring out problems with viewing mental state language as referring to inner causal entities is to consider the number of beliefs that a person can be said to have. Hacker (1996, p. 419) has noted that "there are indefinitely many things that I believe at a given time, but I am not in indefinitely many different mental states at a given time." As Bickhard (2001, 2004) points out, most of us believe, for example, that stepping in front of an on-coming truck is dangerous, but we also believe that blue trucks, pink trucks and even stripped trucks are dangerous. Thus, it is possible to describe a seemingly unlimited number of beliefs that most of us hold. This shows that just as beliefs cannot be particular mental states they also cannot be particular mental contents (Racine & Carpendale, in press). Children do not have the time to learn all of these beliefs nor the capacity to hold this unbounded number of beliefs. Instead, talking about beliefs, intentions or attention is really just to talk about activity. Thus, even if Moore et al. can ground the development of an understanding of attention, no amount of "matching of psychological states with another" could facilitate the developing agent's understanding of other aspects of mental life.

Even if Moore and colleagues were to grant that mental states like intention, desire and belief are abstracted from language use, their explanation of joint attention is still faulty. The infant is learning about a concept called "attention." And in so doing the infant can come to understand that their attention is engaged in some joint activity while using criteria to understand that another's attention is also engaged in this activity. There is no harm in saying that the infant can now "reverse first person and third person perspectives on an object" when this understanding is in place. However, this is a linguistic understanding and not per impossible some matching of psychological states. Moore et al.'s assumption, like Tomasello's, is that there is a one-to-one correspondence between psychological activity and the meaning of that activity and that attention is defined by the psychological events that give rise to it. Not so. The ways in which infants come to perceive other persons as social agents necessarily involve the practices that are built on natural reactions and which are rooted in the form of life of humans and many other higher animals. Like any other mental predicate, if the private language arguments hold we can only *understand* that our attention is coordinated with another's through linguistic practice.

Although it may be important to underscore the role that information-processing abilities might play in transitioning the infant who participates in joint activities requiring attention into one who understands that another's attention is also focused in the exchange, the strong implication in Moore and colleagues' account seems to be that an infant's increasing sophistication in coordinating their interactions is simply a matter of cognitive development given some shared activity. In their account, then, it would seem that any shared activity would do. However, particular social activities partially constitute

and define intention, whereas others partially constitute and define desire. It is human language that creates the capacity to understand that agents attend, intend, desire or believe. It is not shared activity in general that matters, then, but rather particular activities upon which particular words and concepts are grounded. If social understanding develops within such activities, Moore's et al. solution of an emerging meta-representational capacity is only acceptable to the degree to which it reframes shared practices and the language by which they are defined. Whereas language and the embodied practices upon which it is based are a massively important part of the infant's interactional matrix, despite the fact that they advocate an activity-based explanation of joint attention, practice is presupposed and is not elucidated in Moore and colleagues' account. And what distinguishes the infant who participates in joint activity requiring attention from the infant who understands attention is not their abililty to process information. It is linguistic practice.

3.3.2 Piagetian Approaches

As seems to be the case with Moore and colleagues, the developmental starting point is conceived of by Piagetians as one of relative undifferentiation between subject and object (social or non-social). The infant is said to gradually construct generalized "representations" of self, other, and object in and through their practical activities with social and non-social objects. With regard to cognitive skills supporting joint activities that require attention, then, infants would construct relations between other people and objects that become more complex as a function of differentiating and coordinating action schemes (Piaget, 1952, 1954). As a result of decentration, social and non-social objects eventually are experienced as non-self, and persons become experienced as

independent agents. The emphasis is different for Moore et al. because although he also stresses activity in the world, he conceives of social development as a process of matching others' psychological states. In the Piagetian tradition, the development of schemes (expectations) about social and non-social objects entails constructing knowledge of the *relation* between the experiencing subject and the object experienced, which is said to force a constructivist theory of knowledge and an interactive form of representation (Bickhard, 2001).

This is also one way of talking about the infant's experience of perspectives. And as noted at the beginning of this section the psychological details of such a process are causally related to the infant's understanding of attention. However, we must remember that the understanding of attention is not reducible to the infant's disembedding of their psychological engagement from that of another's. Rather, the understanding of any psychological concept presupposes such an ability. Chimpanzees, for example, evidence gaze following skill that shows they are able to take another's perspective, but they do not understand the psychological concept "attention." Mead (1964) has argued that perspectives are embodied in the social acts that agents enact with one another (Carpendale & Racine, 2005). This is not the "in the head" view of perspectives that is often attributed to Piaget. In any case, language also affords further perspectival development by enabling children's participation in more complex contexts (Carpendale et al., in press).

Unlike the above accounts, to my knowledge there has been no systematic analysis of joint attention from a Piagetian perspective, nor has a programme of research been conducted to support such an account to this point. However, Frye

(1991) offers a Piagetian model that, like Tomasello's, emphasizes the subjective experience of means-ends relations. Frye instead focuses on the possibility that these experiences might provide a context within which to discriminate social and non-social objects. Where Frye's account seems to differ from Tomasello's is that there is no reliance on an argument from analogy. Müller and Carpendale (2004) offer an explicitly Piagetian account of a particular activity involving joint attention, namely gaze following. In contrast to Tomasello, they in particular seem to highlight the intentionality *in* activity and are clear that such developments are not mediated by children's understanding of particular mental states like intention. However, although their approach may be appropriate for detailing the development of gaze following, it stops short of being an acceptable model of all aspects of joint attention. The cognitive processes they detail are causal cognitive "mechanisms" that may be required for an understanding of attention, but they are not sufficient for such an understanding because they are not linked up to linguistic practice. Thus, the criticism of Moore et al.'s socially based account would apply here as well.

3.3.3 Emotional Engagement

Hobson (1993, 1994, 2002) offers an account of social development that also emphases infants' early activity, but unlike Moore et al.'s and Piagetian accounts it places much greater emphasis on the emotional context of parent-infant interpersonal relations. This may be because in many respects Hobson's is a reverse-engineered account of typically developing children's social understanding that is derived in

contradistinction to the developmental course for children with autism.[25] And what becomes most striking about children with autism is indeed their lack of emotional engagement with others. The approach I take to the infant's understanding of mental states has the most in common with the socially based approaches, and what is nice about Hobson's theory in particular is that he explicitly draws on Wittgenstein as do to a smaller extent Müller and Carpendale (2004).[26] This will allow me to foreshadow the contrast between Hobson's approach and the one I explicate in more detail below.

Hobson's (1993, pp. 4-5, his emphasis) general thesis is that:

> knowledge and understanding of persons, or to put this differently, a conceptual grasp of the nature of 'minds', is acquired through an individual's experience of affectively patterned, intersubjectively co-ordinated relations *with* other people. A young child comes to know about people's psychological states through having subjective experiences that are shared with, opposed to, or otherwise articulated with the experiences (and not merely the 'behaviour') of others.

The first part of this dissertation is a restatement of the socially based view, with, as noted, stress added on the affectivity of relation, of which I am in agreement. And also I wholeheartedly agree that there is no such thing as "mere behaviour" that is dualistically stripped of its relation to others. However, I have painted a very different picture of how agents come to know about the other's psychological states. As I understand it, we learn about other's psychological states not by matching or contrasting our inner subjective worlds, but rather by learning the circumstances that justify the use

[25] In *Autism and the Development of Mind*, Hobson (1993, p. 5) suggests that "The phenomenon of autism may help think about these issues [e.g., the relation between self and other], if not to solve them." Given that I take the problem of other minds to be a conceptual rather than empirical matter I do not agree that facts about autism can help to solve this problematic. However, Hobson does make a good case that considering how children with autism develop can help us to think about such issues.

[26] However, the activity-based account of Moore and colleagues (e.g., Barresi & Moore, 1996; Moore, 1996a) explicitly sets itself up in contradistinction to a Wittgensteinian approach -- which is identified with Hobson's.

of mental state terms. When we apply such criteria to people we speak about particular aspects of activity. Although Hobson (1994) is careful to point out that the infant does not have a conception of what such psychological attitudes are, my criticism of the other socially based approaches would seem to apply Hobson's account as well. Of course, such emotional engagement is necessary for social development, but it is again not sufficient.

Hobson's theory is more aligned with the one I will discuss below than is Moore's though. For example, Hobson (1994, p. 72) notes that "the infant can be observed to respond to meanings perceived in the bodies and behaviour of other people, meanings that we as adults know to have significance as expressions of a person's psychological state." And, unlike Baron-Cohen (1995) who posits the existence of things called mental states inside people, Hobson (1994) points out that instead there really is such a thing as emotional contact between people. And there surely is and we have criteria for adjudicating when such contact is and is not present. Hobson (1994, p. 76, my emphasis) also claims that:

> The infants' capacity to imitate conventional actions and gestures *reveals* how they can identify with and assume the attitudes and actions of another person. Infants *appreciate* not only that only that they and others have the capacity to share, but that *they are also psychologically distinct*.

If the parallels to Bretherton and her colleagues were not already obvious, Hobson cites Bretherton et al. (1981) after this statement. However, in contradistinction to Bretherton et al., Hobson (2002, e.g., p. 79) is careful to point out that whereas such observations prove to him that infants of this age experience others in a new way this is without yet understanding others as selves.

What seems to me to be absent in Hobson's account is how such an understanding of mental life can actually occur. In my case, I draw on linguistic practice; in his case, he draws on psychological intermediaries. In his (1994, p. 81) paper, he re-states one of the themes of his 1993 book, which is that "an adequate treatment of the matter [of other minds] would begin with an analysis of what it means to acquire a concept of anything at all, and then to apply the results to the special case of 'concepts of mind'." And although Hobson then goes about detailing a series of plausible developmental insights that must occur for the developing child to acquire a concept, as I have pointed out elsewhere we seem to have something quite different in mind by "concept" (Racine, 2004). This difference stems, I think, from Hobson's focus on what must be "psychologically in place" for the infant to understand the intersubjective experiences they share with others versus my focus on how psychological concepts are embedded in embodied activity. However, our languages in many ways predispose us to think in terms of seemingly oppositional categories like physical and mental, body and mind, in-dividual and in-relation, and so on, which are all a series of false dichotomies. And perhaps the disjunction between my more logical and Hobson's more psychological analysis is another one of such dichotomies.

Nonetheless, I believe that Hobson's main thesis, indeed the main thrust of extant socially based approaches in general, of rooting mental state understanding in the infant's psychological engagement with the world may be fundamentally wrongheaded. Mental state concepts require -- and can tolerate -- no such subjective matching. The meanings of mental state terms are not taught to children by getting them to associate the words with some inner correlative phenomenon but rather by simply teaching rules of

use. Although teaching rules of use presupposes the infant's mastery of a given practice, it is the practice that grounds the concept. However, as noted, what we call emotional or more broadly psychological engagement with others is required for (or I would have it, presupposed in) an "understanding of mind," but it cannot be sufficient because such an understanding is tantamount to an understanding of a multitude of particular patterns of activity that are tied up with their respective linguistic practices.

CHAPTER 4 ACTIVITY, ATTENTION, MEANING AND PRACTICE

Meltzoff and his colleagues (1999, p. 17) take it as self evident that while "a few radical philosophers and psychologists may deny the existence of mental states, most regular 'folk' feel sure that they themselves and others have them." Although I would like to see the data showing that people really are as convinced as Meltzoff et al. take them to be that there are actually literal mental entities with validly state-like properties controlling their behaviour, I hope it is clear that I am respecifying, not denying, the existence of mental states. Even the most hardened cognitivist would have to agree that activity is primary -- this is, after all, what all of the talk about people having to "infer mental states from behaviour" is about. Folk psychological idioms are indispensably intertwined with and grounded in virtually all activities in the human form of life (Bennett & Hacker, 2003; Canfield, 1999; Carpendale & Lewis, 2004; Carpendale et al., in press; Montgomery, 2002; Racine, 2004). These concepts and the activities with which they are bound up define what we think of as the mind. The trouble with Meltzoff et al.'s statement is that it is clearly predicated on a disembodied Cartesian picture of mind with its inevitable parsing of (human) activity into a mental and a physical component. As Hacker (1997, p. 14) has noted, the belief

> that a human being is a composite creature consisting of body and soul (or mind, or spirit) is an ancient one. It is bound up with our fear of death, with the craving for an afterlife in a happier world, with our grief at the death of our loved ones and our longing to be reunited with them.... And it is deeply rooted in the grammar of our languages.... This conception, in different forms, was articulated in the religious and philosophical thought of antiquity and the Middle Ages. It was given its most powerful philosophical expression in our era by Descartes. According to Descartes, a human being is composed of two distinct substances, the mind and the body.

For my part, I have tried to demonstrate in the foregoing that virtually all social developmental theory (including that of Meltzoff and colleagues) is reliant to a greater or lesser extent on what I have called the CPR picture of mind, a roughly Cartesian one. My overriding goal in the present section will be to elucidate a developmental account of psychological understanding in infancy and beyond that is predicated on an embodied non-dualistic understanding of mind that grounds mental state understanding in shared activity and natural reactions.

4.1 On the Meaning of (Joint) Attention

Wittgenstein's approach is notoriously at odds with a correspondence (word-object, and as I argue below a gesture-object) notion of meaning and reference, but what exactly is meaning from this point of view? Consider Wittgenstein's (1958, § 2-21) approach to evaluating the classical correspondence theory of meaning. He asks his readers to consider a primitive language such as that used by a builder in telling an assistant what kind of building material is required, such as blocks, slabs and pillars. This primitive language would require labels for these objects as well as numbers to indicate quantity.

But even this simple language immediately grounds out because what, for example, does the utterance, 'five slabs' mean? It could be a statement, a report, an order or some other social activity (Wittgenstein, 1958, § 21). That is, the same utterance can be used in different social situations to convey different meaning. Therefore, its meaning cannot simply be due to a connection to a referent because the same utterance can have different meanings. This is an argument against a mechanistic

view of meaning and reference and argument for an indexical view -- that is, that meaning depends on the location in a sequence of interaction, in a practice or a form of life (Goldberg, 1991). The meaning of 'five slabs' is fixed by linguistic rules. If one likes, it is the social act contained therein as mutually defined by a particular social practice and sequence in ongoing interaction, not the reference to a particular object or set of objects.[27] Wittgenstein's notion of a language-game therefore completely respecifies the referential relationship -- reference does not inhere in the relation between sign and object, but rather "reference consists in a sign's having a role in a language-game" (Proudfoot & Copeland, 2002, p. 338).

I contrasted Butterworth's and Tomasello's definitions of joint attention to draw attention to the distinction between describing how observers can see the sequential activity of two agents as constitutive of joint attention versus the agents' understanding that they are engaged in such an activity so conceived. In so doing I claimed that: (a) although the criteria for Butterworth's notion are obvious, this third person reading was not what most psychologists seem interested in -- and his definition also did not discriminate an active from a passive use of the term, and (b) Tomasello's notion requires that the criteria for the application of the term "know" be satisfied and like all mental state predicates such an understanding cannot be applied to non-language using agents -- and indeed, as we will see, only to agents who understand the multifarious uses of a term in question. When people speak about a baby attending to such-and-

[27] As noted earlier, what might be called levels of meaning become apparent to a human observer of activity with respect to what seems attributed to activity on the part of the agent, but I am interested in a notion of meaning that would justify the use of the expression "shared meaning system" in the strong sense that would, for example, include a recognition of psychological similarity when applied to parent-infant interactions. Given the claims made in the field, this notion of meaning requires a conceptual, or put another way, reflective/non-practical, understanding of activity on the part of the agent.

such at the same time as another person they are not referring to some covert inner mental operation of the child's but rather a particular pattern of coordinated sequential activity in which baby and other are engaged and in and through which such a concept can be applied. That is, if we were to ask researchers and other observers to justify why they thought infant and other were jointly attending to some state of affairs the observer might say something to the effect that both parties in the exchange were looking at something in common. Or, in a more complex practice, that an infant might have looked at her mother before pointing to "make sure" that the mother was paying attention before pointing. Now, some might protest that the workings of the mind are more mysterious than this, but are they? Take any mental state concept and you will find that the criteria that justify the application of the term in question are no less tautological or obvious to an observer who knows the rules of the language-game in question.

If, for example, you were to ask a mother why she said that her baby intends to climb the stairs, she will likely tell you (perhaps with a somewhat quizzical look on her face) that her baby was crawling towards the stairs -- perhaps she might add that there was a look on her baby's face that expressed defiance. Now, if we had grounds to doubt the sincerity of the mother's statement we might do so, but these grounds would not be found in that person's head but rather in their activity. It should also be clear that the mother in this example is not involved in some computational exercise of hypothesis-testing in trying to figure out what mental state the child "is in," she sees this activity as intentional -- because it is. Unlike her baby, however, her answer shows that she has the concept of intention, that she knows how to express her intentions and understands the many patterns of activity that warrant ascription of this mental predicate. And I hope

it is obvious that I distance myself here from Dennett's (1987) intentional stance argument. Ironically, Dennett is a more suitable target for Meltzoff et al.'s (1999) disabusement of radical philosophers and psychologists than are Wittgensteinians because Dennett advocates a functionalistic view of mental states in which their ontology is moot. I am certain that mental states exist when they are properly conceived.

Attending to aspects of one's environment is a natural reaction common across most species. Animals must "pay attention" or they shall surely perish. Infants too must pay attention to their environment in order to ultimately come to make sense of it. The infants' "understanding of attention" as it is often called though is not some separate cognitive wheel that is somehow distinct from, or, even more problematic, temporally prior to the various social practices that require paying attention to aspects of their environment. That is, attention is not some inner state that causes our bodily activity yet is somehow dualistically distinct from it. No more than is a frog's attention a function of their inner states. The infant's *capacity to attend* is presupposed in their execution of joint activities that require attention. I will also note that it is potentially misleading to talk about forms of joint attention because this can obscure the fact that it is forms of social practice within which we attribute particular skills to the infant. We can see that infants get involved in more sophisticated forms of activity involving, for example, more sophisticated use of pointing gestures and following gaze, and that is all. To repeat, there is also a causal story to be told here that might involve the particular brain circuits that "subserve" the infant's engagements, the infant's developing sensorimotor skills and more general cognitive developments. However, as the above should make plain, I am

quite suspicious about the "direction of causality" in these causal mechanisms. What I mean by causal is simply externally the ascription of joint attention. To be sure the agent could not participate in sequences of coordinated activity without causal input from these various mechanisms. However, far from the infant's activity issuing forth from neurocognitive factors, I see such factors as part of a matrix that is conditioned by the infant's engagements in the world.

4.2 Proto Language-games as Bedrock Activities

> Consider the interaction pattern connected with request words. At the earliest stage the child simply cries when hungry, or cold, or wet and so on. Then the mother responds, say, by bringing it to her breast, whereupon the child does its part by suckling. Similarly, there is the interaction pattern of the child trying to get something by reaching for it, and the mother's response of handing it to the child. These are an example of basic, ground-level interaction patterns between child and caretaker that support the development of language. Without such patterns of interaction that arise naturally, without any drill or explicit instruction, between child and caretaker, language would not arise -- logically it could not arise, since it only exists inside contexts of interaction (Canfield, 1995, p. 198).

Canfield (1993, 1995) labels this ground-level pattern of sequential interaction a proto language-game, which is a primitive instance of what I am calling a shared practice. And it clearly presupposed a capacity for agents to attend. Although Canfield points out the continuity between such activities and language, indeed the *dependency* of understanding a language upon such activity, the infant's understanding of attention itself also emerges from what he calls proto language-games. Although word learning requires coordinated attention, joint activities that require attention emerge -- naturally, spontaneously -- in the context of baby and caretaker getting things done. They are proto-phenomena for which no ultimate explanation is required or can coherently be

given. Of course, proximate and causal explanations can be given, but such causal factors do not, in my view, implicate the emergence of say joint attention over the emergence of the child's "stage iv understanding" of the object. These are what are often called domain general factors.

Another way of discussing practice is to point out the importance of routines in grounding a baby's understanding of their world. Theorists who take a social pragmatic approach to language development (e.g., Bruner, 1983; Tomasello, 2003; Turnbull, 2003) have noted that a simple way to counter a strong nativist approach to language development is to capitalize on the disambiguating structure of interaction rather than appealing to innate internal constraints.[28] Imagine a baby who is engaged in some activity. What Canfield (1995) conceives as a proto language-game consists of a caretaker watching what her charge is up to in particular routines, using the routine to sequentially anticipate what her child will do next and then responding appropriately. The baby's role in this primitive game is to sequentially respond to the caretaker's response. However, within such a shared interaction, gesture, for example, can naturally emerge:

> What I shall call the natural gesture arises in the context of, and is inseparable from, the proto language-game. The natural gesture is a stylization of some action that occurs in the proto language-game. That is, such an action is in some way modified, emphasized, or added to, in a way that brings it to the other's attention, and thus it becomes a natural gesture (Canfield, 1995, p. 199).

Canfield draws on examples from the chimpanzee and human forms of life to show how gestures can non-problematically emerge from local interactional contingencies without any a priori intent. Tomasello (1999b) has called these instances

[28] Yes, these are the same Tomasello and Bruner who were argued earlier to hold cognitivistic accounts of joint attention (see Study One for more on their respective views of pointing).

of dyadic signs because they are not reversible from the point of view of the recipient. An example of such a dyadic/natural sign is the "arms up" gesture that becomes transformed from a shared activity where the infant is the recipient of an action into a sign that can signal a request to be picked up. And again the issue of naturally constructing perspectives in and through interaction involving shared practices arises (Carpendale et al., in press). Tomasello (1999b) calls this process of social shaping *ontogenetic ritualization* and acknowledges that some shared activities requiring attention can get off the ground in such an unfettered manner (Tomasello, 1999b, 2003). However, although he allows for this in his theory, Tomasello still maintains that recognition/understanding of the mental states that give rise to the goal-directed activity drives the emergence of most of these activities. However, if social shaping is in principle possible for the emergence of some shared activities requiring attention, then other than unambiguously exposing its philosophical commitments it is unclear what additional purchase Tomasello's theory can deliver. And irrespective of the emergence of this particular shared practice, a dyad cannot develop and use a ritualized gesture to signal a desire to be picked up if their form of life is not one where group members anticipate and respond to requests. This is the only grounding that is provided from practice. In the end, this is what we do. However, form of life is a source of what we do. Instinct is the ground upon which joint activity flourishes, not reason.

In proto language-games the child is of course immersed in sequential interaction involving the type of interpersonal engagement to which many developmentalists both classically and more recently have drawn our attention. But I agree with Canfield (1995, p. 198) who notes that "such actions and reactions arise from the cellular depths of

human nature. Clams don't behave that way." We need to be clear therefore about the role that psychological engagement might play in understanding mental state concepts. In the human form of life, this is how babies and caretakers naturally respond to one another. And once these patterns of sequential interaction begin, simple language-games emerge from -- are grafted onto -- these proto language-games. Therefore, although infants have particular capacities, some of which are often conceived of as "mental," the understanding of mental life is grounded in practice -- not in the psychological events that are theorized to: (a) predispose infants to participate in practice (Baron-Cohen, 1995, 1999; Butterworth, 1998b, 2003; Butterworth & Jarrett, 1991; Gopnik & Meltzoff, 1997; Meltzoff & Brooks, 2001; Meltzoff et al., 1999; Meltzoff & Moore, 1995; Trevarthen, 1979; Trevathen & Aitken, 2001; Trevarthen & Hubley, 1978; Tomasello, 1995, 1999a, 1999b, Tomasello et al., in press), or (b) take place while infants are engaged in practices with others (Barresi & Moore, 1996; Corkum & Moore, 1995, 1998; Hobson, 1992, 2002; Moore, 1996b, 1998; Moore & Corkum, 1994; Moore & D'Entremont, 2001; Müller & Carpendale, 2004).

4.3 Highlighting the Role of Shared Practice in Activities Requiring Attention

As noted earlier, observers can see an infant is following another's gaze as early as the infant's third month of life. The other shared activities that involve joint attention all depend upon, presuppose, skill at this early practice. Moore and his colleagues may be right to account for the emergence of gaze following with learning principles. Ironically, Moore (1999) claims that he has been taken to task for relying on behaviourist

principles in accounting for the emergence of joint attention. But, as Reddy and Morris (2004, p. 650) note, criticism on such grounds is odd because it mistakenly implies that "anything learned from reinforcement cannot constitute genuine knowledge." One can also see how social referencing could get off the ground in the same way with an infant happening to look at a parent after an ambiguous situation and coming to associate over several experiences say a smile with a positive outcome and a frown with a negative one because of attachment reasons (Baldwin & Moses, 1996). This presupposes infants' attunement to their parents' attentional and emotional states before activities involving joint attention even begin to emerge (Hobson, 1993, 2002; Reddy, 2001, 2003).

There is also a social shaping hypothesis of the emergence of imperative pointing (Vygotsky, 1978), which holds that when a parent sees a child's intent in the child's reach for an object, and responds by retrieving the object for the child parents gradually shape the act of reaching into a pointing gesture. Such a notion also makes sense for the emergence of imperatives in non-human species (Canfield, 1995; Leavens, 2004; Leavens et al., 2004). It is easy to see then how learning principles could account for the recruitment of the child's sensorimotor activity into a request practice. Shinn (1900/1975) has suggested that pointing emerges from an alternative sensorimotor route, namely that in the baby's interest in their physical world they begin to touch things at hand with their index fingertip. This "tipping" behaviour is transferred from proximal to distal objects and then we have the child engaging in activity in which reference can also seem to occur. However, the child is not referring according to Shinn but is rather trying to touch things that are far away. Also this is not yet a gesture given the lack of shared practice. But it may be a route to declarative activity that springs from

a separate behaviour antecedent. Although the emergence of many or even all joint activities requiring attention may be able to be "explained" (but not therefore "explained away") by learning principles, these principles can only have the application that they do because of capacities that predispose typically developing infants towards joint activities that are rooted in the human form of life -- and possibly other life forms to a large extent as well. This should not be confused with the position that the *understanding* of mind -- at any level -- is a biological or cognitive a priori primitive. I agree with Piaget that knowledge cannot exist a priori. And although I agree with Piagetians that the children's "cognitive constructive activity" may not be reducible to mere associations, such cognitive activity is externally related to the children's understanding of their world. Thus, although Piaget's view of language has been criticized (e.g., Chapman, 1988), squaring Piaget with Wittgenstein is not just a matter of interjecting a meaning-as-use component into Piagetian theory. Meaning-as-use only makes sense within the overall system of Wittgenstein's thought.

It is worth pausing for a further moment and thinking about declaratives and imperatives. Most theorists see the imperative pointing act as somewhat trivial because apes and children with autism can use this gesture. The idea here is that it can be socially shaped and therefore can come to be through either a mentalistic or a behavioural route. That humans point seems to support the mentalistic gloss; that apes point tilts the scale towards the behavioural one. With this gesture they do not necessarily need to think of imperative pointers as those who experience a prior goal state and then go about pursuing their goal. With the declarative act, there seems to be no other solution other than the mentalistic default. And the default view is that *because*

of their motivation state to share attention with others they declaratively point. But there is a confusion here about the reason for an act and the cause of an act (Wittgenstein, 1958). Just as the activity of imperative pointing expresses a desire for an object, the activity of declarative pointing expresses a desire to share attention in the same manner. There is no causal mental state in one case and not the other. Liszkowski and colleagues (in press, 2004) claim that babies point *to* share interest. But their points express their interest given this practice and give observers (including Liszkowski and colleagues) the possibility to see this interest. And to repeat, imperative or declarative pointing acts are defined by sequential interaction around some practice not by the pointer's state of mind. We say an infant has this state of mind by virtue of the practice. But where exactly is this state of mind? It is embodied in the entire sequence of interactivity.

Activity does get more related to what people think of as more characteristic of agents with minds when we move past the earliest emerging "joint attention behaviour." In contrast to early gaze following, social referencing has been argued to reveal the infant's engagement with a caretaker's emotional reaction to an ambiguous stimulus. This triadic relation may be meaningful to the human or nonhuman infant in the sense that it signals something about a stimulus by virtue of the caretaker's reaction. However, it is a form of meaning that is likely grounded in attachment relations between caretaker and infant (Baldwin & Moses, 1996), and is comparable to the meaning that an infant might attach to a bottle or nipple. Joint activity involving pointing seems to afford a level of meaning that is not limited to the binary approach-avoid significance of the caretaker's "psychological attitude" towards a stimulus as is the case with social referencing, but is

flexibly created in and through the activity of interacting with a conspecific (e.g., making a request through a pointing gesture). This is different than an infant using another's reaction as a danger signal, and it requires that the child develop more flexible sequential expectations about interaction (Kaye, 1982; Wittgenstein, 1958; Wootton, 1997). For an infant to engage in such interactional contingencies minimally presupposes that the infant is skilled at: (a) engaging in the forms of sequential turn-taking behaviour that are well in place when the infant is about 6 months of age (e.g., Trevarthen, 1979) and (b) reconstructing a small number of local sequences (Fogel, 1993; Piaget, e.g. 1954; Turnbull; 2003; Wootton, 1997).

Kaye (1982) has carefully analyzed the parent-infant system in terms of how turn taking is managed. He has come to the conclusion that for the most part psychologists have radically underestimated the asymmetry in such interactions. Kaye argues that whereas there is clear evidence of turn taking at 6 months and even earlier, the parent is providing an elaborate apprenticeship through enacting universal parenting practices which he calls frames that scaffold the infant's development. Pace Bruner (1983), Tomasello (1995, 1999b) and Trevarthen (1979), becoming part of such a system introduces the developing agent to others' intentions by first being deeply embedded within such intentions. However, this is not to say that infants do not respond to the attention of others quite early in the first year of life (Reddy, 2001, 2003).

Kaye describes 4 phases of parent-infant interaction in infancy. In brief, there is first a period of *shared rhythms and regulations* from birth to about 3 months. In effect, the parent is constructing 'turns' around the infant's own natural rhythms and activities, and in this manner the infant begins to learn about turn taking practices. At about 2

months (each period overlaps the previous) the period of *shared intentions* begins which is typified by the parent attributing intentions to the infant's activities. By speaking for the infant as parents often do (e.g., "He says, 'It's time for my nap, Dad'"), for example, the parent is also incorporating the infant into an already existing social system involving characteristically human practices. Kaye argues that at roughly 8 months infants begin to share in another's intentions in the sense that they begin to anticipate the direction of parent's behaviour, not in the sense pace Tomasello (1995, 1999b) that they apprehend something of the mental states that ostensibly led to the behaviour: "The infant's schemas, differentiated through experience in certain behavioural frames imposed by adult behaviour, allow [the infant] to anticipate the most likely direction of that behaviour" (Kaye, 1982, p. 67). Once this begins to occur, and the infant begins to understand signs, a phase of *shared memory* is ushered in wherein parents' actions presuppose that the infant remembers certain experiences that they have shared. Parents use words and gestures to refer to those prior occasions, and in so doing they signify them for themselves, and eventually for the infant. In repeatedly evoking such signifiers, parents are teasing apart signifier-signified relations such that they reduce to the infant being able to predict what a parent will next do. The final phase is one of *shared language* when the infant begins an apprenticeship into the world of human culture at a deeper level (Vygotsky, 1978).

I have argued that children learn about their own minds and the minds of others in and through learning the circumstances that justify the use of words referring to the psychological world -- i.e., the criteria for the use of such words. Such psychological talk is built onto natural reactions depicted in Canfield's (1993, 1995) proto language-games

and the prior practices or patterns of human interaction in which the child participates. Caretakers do not interact with their babies by recourse to some conscious strategy. Non-human primate parents naturally treat certain of their infants' activities as what we human observers see as expressing desire, but this is not reversible from the point of view of both chimp parent and offspring -- because chimpanzees lack the conceptualizing structure that language affords.

4.4 The Role of Criteria and Avowals in Mental State Understanding

To illustrate the general point of the approach I take to child development in more concrete terms, consider the following examples with an older language-using child (T, age 2; 2) and her mother C who are interacting while sharing a meal (cf. Canfield, 1993). The examples were transcribed according to the conventions of conversation analysis (e.g. Atkinson & Heritage, 1984). Square brackets represent overlap in speech. Numbers in round brackets represent pauses in seconds. Dots in round brackets indicate noticeable pauses of under 0.2 seconds in length. Words in brackets represent (in this case) a description of an action. A colon represents a drawn out syllable and the upward arrow represents rising intonation in a pitch contour. Although these examples show that T can distinguish a specific request context from a specific intention context, Example 1 also retrospectively demonstrates all how child, caretaker, object and practice all commingle in ongoing interaction to "display" a particular mental state, in this case an intention.

Table 1: Discriminating an Intention from a Request

Example 1		
1	C:	the [skin is really soft on those kind]
2	T:	[too sour mommy] i han havva drink
3		(.5)
4	C:	[you're going to have a drink↑]
[(T drinks from cup on table)]		
Example 2		
1	T:	MOMMEE i wan more WAT:er in uh BODdle
2	C:	in a bottle or in a cup?
3	T:	in uh:bODDle
4	C:	okay (.) where's (.) oh i see your bottle (.) okay

Consider how a CPR picture of mind would explain the above. Searle (1983), for example, claims intentions are not attributions made by actors in ongoing interaction, but are rather causal mental states that drive behaviour. He holds that grasping the meaning of an action involves understanding the intention that gave rise to it (Racine, 2004). However, grasping the meaning of an action typically involves understanding the practice that is embedded in the local circumstances in which it occurred in ongoing interaction: i.e., *[too sour mommy] i han havva drink* followed by *[(T drinks from cup on table)]*. To grasp the local meanings of actions is to see intention in this particular social situation and in so doing to attribute intentions to the actor who performed the act in question: i.e., *[you're going to have a drink?]*

As noted earlier, human (and most non-human) agents act intentionally early in life. Kaye's (1982) notion of shared intentions beginning at roughly 2 months is consistent with Piaget (1954). And Kaye's notion of the emergence of shared memory is

consistent with Piaget's unambiguous observations of infants' intentions at about 8 months. But intentions are not particular mental states infants develop a capacity to experience, but rather are ways of describing activity (human or otherwise) which presuppose particular capacities on the part of the agent so described. Intentions are embedded in activity. By contrast, Searle assumes that intention utterances are mapped onto actors' underlying goal representations.

Intention utterances (e.g., *i han havva drink*) do not describe a mental event, but rather license expectations about a proposed course of action. A parent might say 'I'm going to pass you the ball' and then pass a ball to a child (a simple language-game built on the proto language-game` discussed earlier). The child learns to do say 'I'm going to x' and goes about doing it. `The child learns that an intention is internally related to the act that satisfies it (Racine, 2004; Racine & Carpendale, in press; ter Hark, 1990; Wittgenstein, 1958, 1980a, 1980b). To repeat. This is a definitional not a causal relation; intentions and actions are mutually defining and one does not exist without the other. To claim that intentions cause actions is like claiming left causes right.

From the point of view of the present account, doing justice, then, to the meaning and nature of phenomena like attention, intention (as shown in Example 1), desire, or belief forces one to collapse the inside and outside dichotomy that we reflexively apply to the activity (Fogel, 1993; Fogel & Thelen, 1989; Kaye, 1982; Racine, 2002; Racine & Carpendale, in press). Two consequences of this view are that the meaning of such concepts: (a) cannot be located in some ostensible constituent of this space of relations (i.e., subject, object, other, practice) and (b) are instantiated anew in ongoing activity. Thus, joint attention can be seen as a paradigmatic case for a way of understanding

development that necessarily forces shared meaning, concept and mind outside of the strict bounds of the human agent's psychological and/or neurological activities. It is a view of mind where episodes of joint attention are meaningful because both agents understand a request is being made, not because both agents might appreciate the correspondence between an action and a referent. From this point of view, reference presupposes, and is parasitic upon, practice.

Human observers rightly describe certain patterns of non-human primate interaction as expressing desire -- that is, of constituting a request. These proto routines are the ground upon which developing agents can base language. Phylogenetically, basic desire practices like doing a request would have emerged from the coordination of our hominid ancestors' activities in order to get basic needs met. However, there would have been no priori practice for our ancestors to draw upon as a cultural resource. Thus, a distinction can be made between emergent local practices, cultural practices that are free to vary to some extent and even cultural practices that are so tied up with the human form of life that they are not free to vary.[29] In the case of emergent local practices that are created within, for example, a human family unit, whatever meaning that might be interactionally manifest when drawing on a shared local practice is still parasitic upon a priori cultural practices that a caretaker is bringing to bear upon the emergent meaning. Only when speaking on an evolutionary scale could one justifiably speak of an emergent "original" local practice.

[29] Folk psychological practices seem to be of the latter variety; one might rightly call these, and the concepts with which they are tied up, universal (Canfield, 1993; Carpendale et al., in press; Montgomery, 2002; Racine, 2004).

The implication of the present approach for understanding social cognitive development is that a caretaker's use of language and the practices with which it is bound up within particular patterns of shared activities with their infants are determinants of the forms of "infant behaviour" known as joint attention. If this account if correct, joint attention behaviours need not and perhaps should not be classified together given that they may reflect the child's immersion in and mastery of different basic human practices. This is not to claim that adults model behaviours like pointing (Butterworth, 1998b). Rather, caretakers do not need to teach their children how to request because they *naturally* treat activity that may not initially have a clear goal as requestive nonetheless (Canfield, 1993, 1995; Kaye, 1982; Vygotsky, 1978). By rejecting a strict modelling account, I am not claiming that learning theory has no place in accounting for social cognitive development. I believe, for example, that Moore and his colleagues may be quite correct that initial episodes of some joint activities requiring attention may well be best explained by a variety of learning principles. However, information processing limitations of the infant cannot alone distinguish infants who engage in joint activities with others that require attention. And with respect to activity-based approaches in general, no amount of matching others' psychological states can stand in for the developing agent's mastery of a natural practice like a request. Such ground level practices necessarily involve coordinating attention with caretakers *and* they are the ground upon which an understanding of the mental state concept desire might ultimately emerge. But although concepts are often grounded on natural reactions, when a child shows mastery of a psychological concept, they are not verbalizing something they already knew.

As I understand them, the main problem with socially driven views as they currently stand is that it is not clear that there is any fundamental disagreement about the ontology of mind between the social, cognitive or biological views. They seem parasitic on a roughly Cartesian picture of mind to the extent that the difference between a cognitive/biological and social approach is that the latter claims that although the mind is not CPR at 12 months it certainly is at 18 months.[30] Although there may be grounds present at 18 months for attributing certain capacities to infants that are not present at 12 months, from my point of view this need not and should not be taken to herald some fundamental revolution in the child. Although the infant may be able to participate in more complicated aspects of the human form of life and although we may have grounds for granting them knowledge of certain concepts, this does not create "a representational mind." What it would suggest though is that such infants know more than their younger counterparts. This is not to say that infants cannot be said to represent, or even more strongly that we have clearer grounds for such a claim for 18-month-old infants. But attention, desire, belief and so on are not inner representations.

4.5 Clearing the Ground for Study One and Study Two

In the empirical portion of this dissertation, I will look at one particular joint activity requiring attention, namely non-spontaneous pointing. I have focussed on this particular activity because it is seen as a canonical "joint attention behaviour," and it is also linked in the literature to the development of both children's understanding of language and mind -- which I clearly think of as inseparable. Now, the theorists reviewed earlier tend

[30] Moore and colleagues hold this position. Whether the other socially based views do or not is less clear.

to treat pointing as a special case of an overall competence to engage in episodes of joint attention (Baron-Cohen, 1995; Gopnik & Meltzoff, 1997; Hobson, 1993, 2002; Meltzoff et al., 1999, Meltzoff & Brooks, 2001; Tomasello, 1995, 1999a; Trevarthen, 1979; Trevarthen & Aitken, 2001;Trevarthen & Hubley, 1978), or as a discrete behaviour requiring a discrete explanation (Butterworth, 1998b, 2003; Corkum & Moore, 1995, 1998; Frye, 1991; Moore; 1996b, 1999; Moore & Corkum, 1994; Müller & Carpendale, 2004). Although I agree that children acquire an understanding of attention in and through any joint activity requiring attention, their *understanding* of attention is in these practices and is nothing more than having knowledge of these particular practices. To put this another way, one might say that this understanding is piece-meal and requires immersion in all activities and thus I would want to account for the emergence of any "joint attention behaviour."

One needs to be careful in making claims about the meaning of the pointing gesture. For one thing, this may ignore or downplay the previous function that pointing might have played in development (Fogel & Thelen, 1989). Also, whereas pointing is canonically a "social gesture" for adults that is meaningful in ongoing interaction, it is not a social gesture when infants first point because it is something that the infant is doing without it is being a gesture for others -- and indeed possibly without an appreciation of others as independent agents.[31] It is also clear that 'pointing' as a discrete manual action occurs early in ontogeny. Hannan and Fogel (1986) observed a first "point" at 18 days of age. Fogel and Hannan (1985) observed spontaneous index finger extension in 2-month-old infants that took place during face-to-face interactions before or after the

[31] Most theorists address this latter problematic when accounting for the emergence of (social) pointing.

infant opened her mouth or vocalized and which were linked to arousal states. Other manual actions (e.g., curling the fingers) did not follow this trajectory. What researchers are interested in then is the development of intentional index finger extension, which can in so doing become a social activity. I report two empirical studies on the development of intentional pointing. In the first I will consider how pointing emerges; in the second I will describe what mother and child are doing when pointing occurs.

CHAPTER 5 ORIGINS OF THE POINTING GESTURE

Fogel and his colleagues have noted that spontaneous fingertip extension occurs disproportionately often within social interaction long before parents begin to treat it as a social pointing gesture (Fogel, 1981; Fogel & Hannan, 1985; Hannan & Fogel, 1987). Similarly, Masataka (2003) has reported that fingertip extension in 3-month-old infants is more reliably associated with their syllabic speech-like output than vocalic non-speech like output or silence. Many theorists interpret these findings, and similar ones by Trevarthen (1979), to demonstrate a primitive coupling between the pointing gesture and reference -- and ultimately language (Bruner, 1983; Butterworth, 2003; Fogel, 1993; Trevarthen, 1979; Trevarthen & Hubley, 1978). Werner and Kaplan (1963) have also argued that infants initially point for the self and they have also linked this gesture to reference. It is not obvious that such early activities should be thought as a means to orient the infant's own attention in the way that Werner and Kaplan theorize, but the link to reference is at least assumed. However, the fact that infants need to realize that people are talking *about* things does not entail that gestures or words are referential in and of themselves (Wittgenstein, 1958). Although an infant can single out an object of desire or interest as a referent, gestures, like words, refer by virtue of their role in practice not by their correspondence to their referent. And in either case we would still need to account for how the "primitive function" we ostensibly witness in 2- to 3-month-olds can become a social gesture in the 10^{th} or 11^{th} month (Butterworth & Morissette, 1996; Carpenter et al., 1998). As it turns out there are different ideas about the ways in which these earlier more sensorimotor activities link up to practice. And as is the case

for explanations of joint attention in general, theories of the origins of pointing tend to emphasize roughly biological, cognitive or social factors.

5.1 Introduction to Study One

5.1.1 Gesture or Gestures?

Pointing occurs within different communicative contexts; it plays different roles in different practices. Does pointing appear and then get recruited into different practices as one might expect from the observations of Fogel and colleagues, or do different practices actually select the same roughly canonical gesture and its functional significance is assigned therein? Given the morphological isomorphism of protodeclarative and protoimperative pointing in at least Western cultures (Butterworth, 1998b, 2003; Povinelli & Eddy 1994; Povinelli et al., 2003) answering this question on empirical grounds is not as easy as it seems (Wilkens, 2003). However, it is most typically assumed that function precedes practice. And the function at work in the line of argument of Fogel, Masataka and Trevarthen is a protodeclarative one stemming from infants' interest and motivation in sharing attention with their caretakers around objects of mutual interest. It is this particular function that has been argued to ground word learning (Bates et al., 1987; Camaioni, 2001; Franco & Butterworth, 1996).

Now, infants are clearly interested in others and motivated to engage in joint activities with them that ultimately come to involve attention in the sense that their interest and motivation is defined in these self same activities, but what specific adaptive problem is solved by referential pointing? From the above it would seem that the

adaptive problem is word learning. And the fact that children with severe autism and great apes have been both argued to lack protodeclaratives and are also both very impoverished in word learning is offered up as evidence for selection of this function in the human line (Baron-Cohen, 1989, 1995; Butterworth, 1998a, 2003; Povinelli et al., 2003; Povinelli & Davis, 1994). However, Savage-Rumbaugh and colleagues (1998) have reported anecdotal protodeclarative pointing in the enculturated and language-trained bonobo, Kanzi. They suggest that declarative pointing and at least rudimentary language emerge in and through the particular entire pattern of life within which Kanzi has been raised. These functions on this view are properties of particular life histories. And it is also not clear that that imperative and declarative gestures issue forth from a separate ontogenetic motivation and phylogenetic function (Hobson, 1993, 2002; Moore & Corkum, 1994; Leavens, 2004; Leavens et al., 2004, Leavens et al., 2005b; Moore & D'Entremont, 2001).

The parsing of pointing into protodeclaratives that are said to show genuine prelinguistic reference in 12-month-old infants versus protoimperatives that serve only an instrumental function only makes sense if one presumes that that meaning of these activities is to be found in the agent's psychological motivation (or in the case of protoimperatives, as some would argue, the lack thereof) (Liszkowski et al., in press, 2004; Moore & Corkum, 1994; Moore & D'Entremont, 2001). And in contradistinction to cognitively driven accounts of pointing like Michael Tomasello's, socially based accounts (e.g., Moore & Corkum, 1994) have stressed the role of activity and engagement.[32]

[32] I am reminded while writing this of a 13-month-old infant who upon recently meeting my 4-year-old daughter pointed at her after a short while (without gazing at her parent or any other adult) and then proceeded some minutes later to hug my daughter for an extended period. Why do researchers in the main seem to think that the point reveals something more psychologically telling than the hug?

Dave Leavens (personal communication) believes that whereas protoimperative pointing emerges in an instrumental context common across primates, it is the affective context of human mother-infant interaction that creates the social situation in which protodeclarative pointing can emerge (Hobson 1993, 2002; Moore & Corkum, 1994; Reddy, 2001, 2003; Savage-Rumbaugh et al., 1998).[33] Early pointing then, whether imperative or declarative, may share an identical function -- namely, to cause some change in an interlocutors' activity. In the case of imperatives, it may be to get an object or event; in the case of declaratives, it is to "get attention." In this sense, they are both imperatives.

5.1.2 Biologically Driven Hypotheses

Although Butterworth (1998b, 2003) argued that pointing occurs in human ontogeny because it is a species-typical hand gesture, he (2003, p. 16) noted:

> The recent upsurge of research on pointing in chimpanzees suggests that it is not possible to maintain an absolute divide between humans and other higher primate species with respect to the gesture. Some aspects of indexical pointing may be shared with other primates...

I should first clarify that by "indexical pointing" Butterworth means "pointing with the index finger," not the more radical proposition that the meaning conveyed in an activity involving pointing is itself indexical -- that is, that it is critically dependent on, breathes life in, *this* social situation. In any case, Butterworth suspects that differences in the need for referential precision are responsible for the fact that chimpanzees often

[33] However, the fact that Kanzi might learn to declaratively point in such a context need not mean that human infants do. Leavens and colleagues (Leavens, Bard, & Hopkins, 2005) have recently presented some data showing that human mothers smile when they themselves are declarative pointing. But Leavens also needs to show that mothers treat infant pointing in the same manner and more importantly that this response is preferential to declarative over imperative social situations.

indicate with a "whole hand point," whereas human infants employ the index finger for this indicative function. Povinelli and his colleagues (Povinelli & Davis, 1994; Povinelli et al., 2003) have argued that the presence of a particular tendon in the human hand that is absent in the hand of apes predisposes humans to index finger pointing.

Butterworth and Morissette (1996) have also found that the use of a pincher grip in human ontogeny precedes the onset of pointing. And although such a grip is present in apes, it is utilized earlier and more frequently by humans. Butterworth (2003) suggested that the pincher grip and pointing gesture are coevolved for precise action in the former case and precise communication in the latter. In fact, Butterworth argues that the grasping and pointing functions are antithetically related in evolutionary terms (Darwin, 1904/1965). That is, when gripping, one's attention is on the hand; when pointing, one's attention is on the target. These lines of argument and the findings that open-hand indicative gestures (i.e., reaching) and pointing are unrelated (Franco & Butterworth, 1996; Lock, Young, Service, & Chandler, 1991) suggested to Butterworth that the human pointing gesture likely occurs spontaneously in its fully functional form given a communicative situation. Thus, he argued that the gesture itself is not socially shaped.

I have argued that the fact that pointing might naturally emerge in human ontogeny need not mean that it was selected for its communicative-referential function. For example, on evolutionary grounds it is plausible that Darwin's (1859/1964) principle of covariation rather than his (1965/1904) principle of antithesis could apply here. That is, rather than these functions being coselected (for ostensibly antithetical functions), the human tendency to point with the index finger (and the evolution of the tendon noted by

Povinelli and Davis, 1994) could be epiphenomenal of the selection of the pincer grip, possibly to do with more specialized demands of tool use. That is, the ostensible priming given to the canonical index finger point by the tendon reported by Povinelli and Eddy (1994) could have been happenstance. However, once selected for specialized tool use, Homo sapiens may have then become more likely to ultimately gesture with the index finger rather than open hand. Arguing that humans have evolved a communicative-referential function also seems to ignore that fact that many cognate species communicate and refer within particular shared activities. And even if humans are biologically predisposed to gesture by canonical pointing (but see Goodwyn, 2003; Wilkens, 2003), this does not mean that the index finger has been imbued with a capacity for reference that has been naturally selected. The shape of the gesture also may be less important than the function (Krause & Fouts, 1997). And being biologically prepared to point may turn out to at best a structural precondition for pointing within particular shared activities.

Leavens and his colleagues (2004) have shown that although chimpanzees typically point with the whole hand rather than the index finger, they can perform "canonical points" when the need arises. They argue that the reason apes may not gesture very often in the wild is not because they cannot but because they do not typically encounter a "problem space" like being in a cage with an object out of reach and a human experimenter within view. And in some human cultures, pointing is done with the whole hand or with the fingertips spread out, or even the lip. In Western cultures we also use our thumb when pointing behind or to the side to indicate location (Wilkens, 2003). These are all "referential gestures" in their particular social situations.

However, I doubt that Leavens would ask a chimpanzee for directions. Such a practice is parasitic on the human form of life. And what is natural and unnatural for chimpanzees defines their form of life. Natural, of course, need not mean strictly genetically determined. From my perspective, the attributes of human interaction that define the mind reflect more cultural than biological lines of influence. However, the categories of biological and cultural are yet another false dichotomy. And the Wittgensteinian notion of form of life is meant to cut a swath between both constructs. However, Wittgenstein's approach can be thought of as a form of social naturalism.

Leavens (2004) seems to want to settle such matters on empirical grounds and criticizes Povinelli et al. (2003) because their arguments reduce in Leavens' view to the position that:

> humans are humans and apes are apes -- [and] one would be hard-pressed to find any clearer example of an essentialist argument, an argument that can never be addressed through empirical inquiry. (Leavens, 2004, p. 158)

I hope it is clear that this does not by implication make my argument an essentialist one; the essentialist claim of bifurcation between the human and non-human primate lines is parasitic on a Cartesian ontology (Savage-Rumbaugh et al., 1998). However, although I agree that Povinelli et al.'s mentalistic divide between the human and non-human primate line is problematic, I obviously do not agree with Leavens that such matters can be settled on (strictly) empirical grounds.

Butterworth's and Povinelli's positions are tenable to the extent that they claim that human infants are biologically prepared to point. In particular, they give good reasons why babies might be primed to display canonical pointing gestures in a way that other primates are not. Further, in my view pointing is a natural gesture formed through

natural reactions between child and caretaker and perhaps it is more natural for humans to use the index finger to gesture than apes given tendons, pincher grips, and so on. A consequence of Butterworth's and Povinelli's approach though is that it assumes that protoimperative and protodeclarative pointing are ontogenetically and phylogenetically distinct. And it is the protodeclarative that is selected by evolution, whereas the protoimperative we share with other members of the primate line. But from my perspective, to the extent that there is a difference in pointing across the human and non-human primate lines it is declarative activity that may not exist in great apes rather than a particular type of pointing function being absent. From my perspective, function presupposes and emerges from practice -- in both phylogenesis and ontogenesis.

In terms of empirical hypotheses, from Butterworth's point of view one would expect the frequency of reaching and imperative pointing to positively correlate in the first year because they are both imperative acts. However, from Butterworth's point of view one would not expect a correlation between reaching and protodeclarative pointing because these are supposed to have evolved from different functional contexts. Indeed, Franco and Butterworth (1996) argued that the fact that they found no correlation between reaching and declarative pointing supported their position. And there is no reason to think that babies who frequently reach would also frequently protodeclaratively point because the former is argued to stem from an instrumental motive whereas the latter is not.

5.1.3 Cognitively Driven Hypotheses

Tomasello (1995, 1999a, 1999b) has two competing ideas about the origins of pointing. One is a social shaping account (1999b); the other is a cognitive account

(1995, 1999a). Tomasello (2003, pp. 34-34) does not specify which route is more likely in human ontogeny, only that both routes are possible:

> we do not know whether infants learn to point via ritualization or imitative learning or whether some infants learn in one way (especially prior to their first birthdays) and some learn in the other.

And unlike the biological accounts, he does not explicitly tie function to either route. Although in this quote he seems to have no particular stake in a cognitive theory of the origins of pointing, as noted in earlier sections his theory of the origins of the "manifestations of joint attention" is that the reason that all of these activities emerge at 9- to 12-months is that infants "recognize" that they and others have mental states called intentions when they engage in goal-directed activities. However, although he allows that a single joint activity (in this case pointing) can get off the ground through a non-cognitive route, the implication of his general theory is that it is highly unlikely that all such activities could get off the ground in that manner. But logically this does not follow. And as I have speculated earlier, this position may be motivated by his attempt to account for differences between apes (wherein he once believed that pointing and joint attention in general had to emerge by non-cognitive means) and human infants (wherein he still believes that the reason that joint attention occurs in humans and not other primates is cognitive in origin). Although Tomasello and his colleagues now realize that it is not this simple (Call et al., 2004; Tomasello et al., in press; Tomasello et al., 2003), his theorizing is still thoroughly cognitive in character and does not recognize the conceptual underbelly of the issues being investigated.

Tomasello's cognitive account of pointing is that the reason that infants *begin* to point, then, is that they have experienced the mental states that give rise to their goal-

directed activities and because they have a previous biological (Meltzoff & Moore, 1978; Tomasello, 1995, 1999b), or social (Hobson, 2002; Tomasello et al., in press) identification with others they therefore automatically apply this understanding (which he calls a recognition) of their own mental states to others. I am not measuring other joint attention behaviours in the way done by Carpenter et al. (1998) or Slaughter and McConnell (2003) and I will therefore not investigate the synchrony of emergence between several joint activities requiring attention. But in any case the interpretations of these data can be richer or leaner (Butterworth, 1998a; Carpenter et al., 1998; Moore, 1998). One seeming way to test a conceptual hypothesis of pointing is to simply note whether infants continue to point once they begin to point. If the emergence of pointing is based on some understanding (or even recognition) of another persons' attention or intention, then once babies point they should presumably continue to point. As I noted in an earlier section, this sort of commonsense check on Tomasello's theory was disallowed in Carpenter et al. (1998) because they ignored within-subject sources of variance in the onset of joint attention behaviours. And within-subjects variance certainly wouldn't be expected from Tomasello's cognitive theory as it might be applied to the emergence of any one joint activity requiring attention.

5.1.4 Activity Driven Hypotheses

Tomasello's social shaping account of pointing can also be tested. As noted, he (1999b) calls this a process of ontogenetic ritualization. What he envisions by the term is the shaping of a form of intentional communication that initially only works in one direction. That is, although an infant can use the "hands up" gesture once it has been socially shaped as "a request to be picked up," the infant would not necessarily

understand such a gesture outside of the particular interactive context in which it was ritualized (Tomasello, 1999b; Canfield, 1995). So, an infant might not automatically understand another infants' use of such a gesture and a parent could not necessarily automatically use such a gesture towards the child to query whether the child might want to be picked up. Similarly, a baby chimp and mother might ritualize a nursing routine where an entire sequence of behaviour is abbreviated to a tap on the shoulder (Vygotsky, 1978), but the baby chimpanzee would not understand a tap on the shoulder from its mother as an opportunity to nurse (Canfield, 1995; Tomasello, 1999b). Another way of putting this is that the meaning of the gesture is not reversible from the point of view of both interlocutors (Baressi & Moore, 1996; Moore, 1996b).

Two specific activity driven accounts of the origins of pointing have been briefly discussed in earlier sections. Vygotsky's (1978) explanation of the origins of pointing is an instance of "ontogenetic ritualization," whereas Shinn's (1900/1975) is not obviously so. Shinn has suggested that pointing emerges from the baby's exploration of their physical world. Specifically, the hypothesis is that the infant first begins to touch nearby objects with their index fingertip ("tipping") and then this kind of exploration on objects that are out of reach is attempted. We therefore have a transfer of tipping from proximal to distal objects and then we have a baby who is performing canonical pointing "gestures." There is an irony in calling this a gesture though because according to Shinn it is an asocial act. Like the aforementioned biologically driven accounts, this activity is not shared activity. And Shinn's explanation as it stands is not linked up to an interactive context. Thus, while we could test this theory by seeing if infants do engage in proximal and then distal tipping, it is still not an acceptable account of the origins of the pointing

gesture as a social act.[34]

Vygotsky (1978) argues that intentional pointing is socially shaped through unsuccessful grasping.[35] In the proto language-game of the child trying to grasp an object that is out of reach a caretaker naturally makes a request out of a non-requestive act. In so doing, they complete the act for the child, or as Vygotsky would put it, they mediate the child's intent. Through this shaping process the entire activity becomes abbreviated into a pointing gesture. Thus, like Shinn, Vygotsky argues that the origins of pointing are to be found in the child's sensorimotor activity. And what ultimately becomes intentional behaviour is a byproduct of the child's engagement with their physical world. Were it not for a social world and primate form of life, this act could not emerge within shared practice. It is contingencies of natural reactions that create this activity and not some conscious intent or strategy on the part of a caretaker to teach a child how to make a request, anymore than a great ape needs to be taught to treat their infant's attempts at nursing as signifying intent or desire in order to react to them. Vygotsky's, unlike Shinn's, then, is one account of how infants are "apprenticed" into a shared practice of, in this case, requestive pointing. And given my theoretical orientation, this particular hypothesis and the notion of social shaping in general would be expected.

[34] I will have more to say about this in Study Two.

[35] Franco and Butterworth (1996) argue that the fact that 3-month-old infants have been reported to extend their fingertips ("point") during social interaction (cf. Fogel & Hannan, 1985) *before* being able to reach and grasp casts doubt upon Vygotksy's hypothesis from the get go. But this assumes that the function of pointing is already set in stone. Whatever fingertip extension might mean at 3 months, imperative pointing could still in principle emerge from failed grasping.

5.1.5 Previous Research on the Origins of Pointing

Although various ideas have put forth about the emergence of the pointing gesture -- the most central of which have just been reviewed -- there has been very little data generated. Indeed, this may be one of the few areas of developmental psychology that is theory rich but data poor. In two early case and multi-case studies, Bates, Benigni, Bretherton, Camaioni, & Volterra (1979) reported observations that are consistent with Shinn in that they found that proximal "poking" or "scratching" by infants was ontogenetically prior to pointing and Bruner (1983) concluded that pointing does not, pace Vygotsky, emerge from reaching. Like Shinn, Bates et al. (1979) interpreted early pointing activity in sensorimotor terms as object exploration. Bruner (1983, p. 75) on the other hand claims that early pointing is "part of a primitive marking system for singling out the noteworthy." And although this claim can be understood in different ways, Bruner (1983, p. 122) clarifies that:

> the 'intent to refer' is unlearned and....so too is the recognition of that intent in others.... Some basis for referential intersubjectivity must exist before language proper appears.... It is a primitive that 'other minds' are treated as is they were like our own minds.

What Bruner is noting is that language develops within joint activities that require attention, which is not a contentious claim. However, this does not require that this "basis for referential intersubjectivity" be a biological primitive. To the extent that joint attention is biologically specified, it seems more likely to implicate priming towards interest in conspecifics that is not present to the same degree in other primate lines (e.g., Hobson, 2002; Tomasello et al, in press). Being interested in conspecifics is not the same things as having a nascent recognition of others' intent. And, pace Bruner *and* Tomasello it need not be tied up with an implicit or explicit form of mentalism. Joint

activities requiring attention develop in and through human interaction.

To my knowledge, there are four other published studies with larger samples that explicitly look at the origins of pointing. There are also a few other studies that speculate about the emergence of pointing as a function of the topic of interest in their study, but as I understand it only the four I presently review were conducted to investigate this issue and test different hypotheses. Leung and Rheingold (1981) cross-sectionally studied four groups of 12 infants ranging from 10- to 16-months of age. They concluded that pointing replaces reaching as a referential gesture, and propose that the pointing gesture may be acquired by modelling.[36] Lock et al. (1990) cross-sectionally studied fourteen groups of 10 mother-child dyads with infants ranging from five- to twenty-four-months of age. They concluded that pointing does not have its origins in failed reaching, but they claimed, in a somewhat Vygotskyan manner, that an adult's response to a failed reach may still provide the baby with information about how to influence another. It is also notable that they concluded their paper on the cautionary note that "the origins of pointing are not straightforward." Indeed.

Like Lock et al. (1990), Franco and Butterworth (1996) took the fact that reaches and (imperative) points were uncorrelated in two cross-sectional samples (22 infants in one study, 25 in the second) of 12- to 18-month old infants and the fact that while pointing increased exponentially reaching remained at a steady and low rate as inconsistent with Vygotsky's hypothesis. Franco and Butterworth also brought attention to a problem in earlier studies, namely that Vygotsky's hypothesis can be used to account for the emergence of only imperative, but not declarative, pointing. And as such

[36] Kaye (1982) has similar ideas, but he argues that modeling occurs via frames rather than hand shapes.

heretofore investigations into the origins of pointing needed to be reinterpreted as investigations of a "mixed referential and request-inducing context" (Franco & Butterworth, 1996, p. 311). Finally, Masataka (2003) studied eight 3- to16-month-old infants longitudinally. He found that reaching behaviour was also at a flat line low rate and that an increase of index finger extension followed by a decline in extension occurred as pointing emerged. (Note that it was index finger extension not exploration which was measured (Shinn 1900/1975). Also, Masataka did not discriminate the function of pointing in relation to reaching.)

I want to first clarify that what has become known (and dismissed) as "Vygotsky's hypothesis" in the field is really little more than an example that Vygotsky used to demonstrate his general theory. I do not think that Vygotsky had anything invested in this particular example and I also doubt that he had an overly simplistic view of pointing that was tantamount to it being a strictly imperative act. And if this particular hypothesis were not supported, it would not undermine his general theory. In any case, it would seem that this hypothesis has been tested and summarily dismissed. But was an adequate test conducted? As Franco and Butterworth note, Vygotsky's hypothesis can only account for the emergence of an imperative act. And only their study looked at imperative pointing in relation to reaching. Irrespective of the function of pointing, as Franco and Butterworth also point out, if we are to hypothesize that pointing emerges from failed reaching and that what is mediating this effect is the role of a caretaker in getting the desired object, then we should also be attentive to the infant's eye contact with the other while they are reaching. However, although Franco and Butterworth (1996) took pointing function and the role of gaze into account and still found no support

for the reaching-pointing hypothesis, they focussed on the second year of life. Although they did include a small group (n = 7) of 10-month-old infants, we should be interested in reaching as pointing first emerges not after it is well established in the second year. Further, even their second year finding of a statistical flat line for reaching frequency should be investigated in the first year of life. And as I have suggested above, if reaching and imperative pointing are the acts that Butterworth takes them to be, then one would expect them to be positively associated when they first emerge. Also, a longitudinal rather than cross-sectional study would be best. To address these issues and to look more broadly at the relations between pointing, imperatives and declaratives in the first year of life, a group of infants was followed longitudinally with monthly visits at the 9th, 10th, 11th and 12th months. As well as reaching with or without eye contact, imperative and declarative gesturing (again with or without eye contact) were distinguished in the administration of a measure of social communication skill (Mundy et al., 2003).

5.1.6 Expected Results

A negative association between the frequency of reaching and imperative pointing acts across time in the end of the first year of life would be consistent with Vygotsky's hypothesis. (And if support for this hypothesis is found, then support for Tomasello's ontogenetic ritualization/social shaping notion is found by definition.) However, if reaching and protoimperative pointing were positively associated across time in the first year this would be consistent with Butterworth's position because Butterworth conceives of both as imperative acts. On this logic, children who are more imperative, so to speak, should show higher levels of reaching and imperative pointing.

Also, because Butterworth argues that reaching and imperative pointing are functionally distinct from protodeclarative pointing one would not expect an association across time between declarative pointing and reaching. Although I will not directly evaluate Tomasello's cognitive hypothesis, a finding that pointing (whether declarative or imperative) emerges at a given month and continues to be in the infant's gestural repertoire at subsequent visits would be consistent with Tomasello's conceptual hypothesis because once an infant has had this insight it should persist. (Note that if there were support for both a socially based and a cognitively based view, it would be consistent with the position that imperative pointing emerges out of failed reaching. However, it could also appear to an observer that this resulted from some cognitive insight because of the abrupt behavioural shift.)

Given the low frequency of the declarative and imperative pointing gestures in the first year of life, the 9 and 10 month ("early visits") and the 11 and 12 month ("later visits") visits will be grouped together when evaluating the above associations. And one would expect, for example, a negative correlation between early (9 and 10 month) reaching and later (11 and 12 month) pointing according to Vygotsky. Mean summary data will also be presented and changes in frequency of all child activities will be noted.

5.2 Method

5.2.1 Participants

Twenty-nine mother-infant (15 girls, 14 boys) dyads were observed when the infants were approximately 9, 10, 11, and 12 months of age.[37] As is standard in the field and reflecting the difficulties of research with infants, visits were scheduled within two weeks of the given month of age (e.g., Carpenter et al., 1998). This was a broad urban community sample selected from a western Canadian city by newspaper advertisements and word of mouth. As a recruitment incentive, all dyads were compensated for participating in the study. All mothers were partnered. Given the educational and occupational status of the families in this study, this was a primarily middle class sample. Twenty-four of the infants were Caucasian and 5 were Asian. Eighteen of the infants were first born. All were full-term. Six infants did not point by 12 months and were excluded.

5.2.2 Design and Procedure

Infants were videotaped from two angles as the Early Social Communication Scales (ESCS, Mundy et al., 2003) were administered. In order to improve sound quality, 2 PZM microphones were inputted into the cameras instead of relying on the camera's internal microphones. The ESCS is a semi-structured observational measure that takes 20 minutes to administer on average. Through this instrument, non-verbal competence at declarative and imperative activities can be noted as well as competence

[37] One mother-infant dyad withdrew from the study after the 10-month visit. Their data was not included.

in social interaction. Declarative acts refer to the child's skill in using nonverbal behaviours (e.g., eye contact, showing, pointing) to share the experience of objects or events with other people, whereas imperative acts refer to the child's skill in using nonverbal behaviors to elicit aid in obtaining objects or events from people. Social interaction acts refer to the child's skill in engaging in "playful, affectively positive turn-taking interactions with others" (Mundy et al., 2003, p. 1).

All three classes of act can be classified as to whether they are used to initiate a declaration, request or social interaction or whether they are a response to another's declarative, imperative or social interaction bid. Initiating a declarative refers to the frequency with which a child uses eye contact, pointing and showing to initiate shared attention to objects or events. Responding to a declarative act refers to the child's skill in following the tester's line of regard and pointing gestures. Initiating an imperative act refers to the child's skill in using eye contact, reaching, giving or pointing to elicit aid in obtaining an object, or object related event. Responding to an imperative act refers to the child's skill in responding to E1's gestural or simple verbal commands to obtain an object or action from the child. Finally, initiating social interaction refers to the child's skill at initiating turn-taking sequences and the tendency to tease the tester. Responding to social interaction refers to the frequency of eye contact, gestures, and turn-taking exhibited by a child in response to turn-taking interactions initiated by the tester (Mundy et al., 2003). From this entire measure, frequency counts of infant reaches, reaches with eye contact, declarative points, declarative points with eye contact, imperative points and imperative points with eye contact were conducted.

As suggested by Mundy and colleagues, the order of presentation of the tasks was as follows: (a) Ball (Turn-Taking Task), (b) Wind-Up Mechanical Toy (Object Spectacle Task #1) (c) Glasses (Response to Invitation Task), (d) Hand-Held Mechanical Toy (Object Spectacle Task #2), (e) Social Interaction Task #1, (f) Gaze Following Task #1, (g) Balloon (Object Spectacle Task #3), (h) Book Presentation Task, (i) Wind-Up Mechanical Toy (Object Spectacle Task #4), (j) Car (Turn-Taking Task), (k) Hat (Response to Invitation Task), (l) Wind-Up Mechanical Toy (Object Spectacle Task #5), (m) Hand-Held Mechanical Toy (Object Spectacle Task #6), (n) Comb (Response to Invitation Task, (o) Plastic Jar Task (Initiate Imperative Act), (p) Social Interaction Task #2 and (q) Gaze Following Task #2.

During testing, infants were seated on their mothers' laps behind a 3' by 5' fold-up table while a research assistant (E1) paced the infant through the ESCS procedures and a second research assistant (E2) recorded the observations. E1 and E2 were trained in the administration and scoring of the ESCS materials by utilizing a training protocol and videotape provided by Peter Mundy, the first author of the ESCS. If E2 was unclear about how the infant responded to a particular procedure, she checked the videotapes after the fact in a research laboratory. The first camera angle showed the infant only in full profile so that a complete record of their reaching and gestural activity would be recorded, whereas the second showed the infant's and E1's faces from a side profile so that the infants' gaze could be retrospectively verified if needed. All recording sessions were scheduled at the mother's convenience and took place in the dyads' homes. Mothers were told that the study would investigate how healthy infants develop

ways of communicating with their parents. If a mother had to discontinue a visit because their infants were too fussy or tired, such visits were re-scheduled as soon as possible.

5.2.3 Materials

The toys and other materials used in the ESCS were selected because of their potential to elicit social interaction and declarative and imperative acts. All toys used in the ESCS were positioned within view but out of reach of the child and the toys are presented one at a time during the administration of this measure. Three sets of similar but non-identical sets of toys were used such that no set was used in consecutive visits. Each set included: (a) 5 small wind-up mechanical toys (3 of which were presented individually, two of which were put inside a plastic jar), (b) 3 hand-operated toys (a jack-in-the-box, a pop-up puppet and a balloon), (c) a small car and a ball that could be easily rolled across the table, (d) a book with large distinct pictures on its pages, (e) a toy comb, hat, and glasses and a 6" tall clear plastic jar with screw-on lid.

5.2.4 Scoring

The tasks of interest for testing the hypotheses regarding the origins of pointing were the object spectacle and book tasks. Points to objects within a given sequence of infant-E1 interaction were coded as declarative unless additional information (e.g., reaching or verbalization) indicated an object request. Alternatively, if the infant used an extended index finger to express desire for an object or event within a given sequence of interaction then it was coded, by definition, as imperative. As per ESCS conventions, a point during any of the object spectacle tasks was coded as a protodeclarative if it occurred while a toy was *active* and protoimperative if it occurred *before* or *after* the

activation of the spectacle. Thus, it is assumed that given this social situation the infant would not point while the toy is active to request that it be activated again but rather that the infant is attempting to share their interest in the toy. However, if the toy is not active it is assumed that the child would not be attempting to express interest in a non-active toy but is rather prompting E1 to activate or re-activate it. All points at items depicted in the picture book task are considered declarative because it is again assumed that the child is not expressing a desire, for example that the pictures be ripped out of the book and handed to the child but is rather expressing interest in the items depicted. There are six object spectacle tasks (3 mechanical, 3 hand-operated), each of which was activated 3 times in a given task. This gave babies 18 opportunities to point in ways that could be coded declarative or imperative and to reach. There were also six pictures displayed in the picture book to which the infant could declaratively point.

The number of observations of a given behaviour was tabulated for all infants across the administration of the procedures. In some instances a child reached and then turned the reach into a point or vice-versa. In such cases, the higher-level behaviour (i.e., pointing) was coded. Thus, all categories of reaching and forms of pointing were mutually exclusive. However, for statistical analyses, the overall categories of reaching, imperative pointing or declarative pointing did not distinguish eye contact. The category of reaching, for example, included the subcategory of reaching with eye contact, and so on. The reason for this was that reaching with eye contact is a form of reaching and thus to appropriately test the hypotheses one should test the relations between the mutually exclusive categories and at the level of the more refined category that took eye contact into account.

5.3 Results

Alpha was fixed at .05. All tests were two-tailed. Non-parametric statistical methods were employed in this study because these data were decidedly non-normal.

5.3.1 Differences across Visits

The Wilcoxin Signed Ranks procedure, the non-parametric analogue to the paired t-test, was used to test mean differences.

There was a significant increase in reaching while making eye contact with E1 at 12 months when compared to 9 months, $z(22) = 2.122$, $p < .05$. There was a significant increase in declarative pointing at 10 months when compared to 9 months, $z(22) = 2.716$, $p < .05$. There was not a further significant increase in declarative pointing at 12 months. The increase in gazing at E1 while declarative pointing at 12 months was not significant, $z(22) = 1.633$. There was a significant increase in imperative pointing at 11 months when compared to 9 months, not significant, $z(22) = 1.342$. The decrease in reaching frequency at 10 months was not significant, $z(22) = -1.391$.$z(22) = 2.320$, $p < .05$. There was not a further significant increase in imperative pointing at 12 months. The increase in gazing at E1 while imperative pointing at 12 months was not significant, $z(22) = 1.342$. The decrease in reaching frequency at 10 months was not significant, $z(22) = -1.391$.

Table 2: Descriptive Statistics for all Variables by Month of Visit (n = 23)

	9 month		10 month		11 month		12 month	
Variables	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Reach	9.91	5.30	8.09	6.27	9.96	7.18	10.61	5.80
Reach with Gaze	1.17^{A}	1.34	1.83	2.15	1.74	1.71	3.39^{A}	3.53
Declarative	0.00^{A}	0.00	0.83^{A}	1.85	0.70	1.11	1.22	2.65
Dec Pt with Gaze	0.00	0.00	0.04	0.21	0.09	0.42	0.39	1.47
Imperative Point	0.13^{A}	0.46	1.61	4.45	1.57^{A}	3.01	2.39	4.40
Imp Pt with Gaze	0.00	0.00	0.26	0.86	0.35	1.03	0.30	1.26

Values sharing the same superscript differ at $p < .05$

5.3.2 Correlations between Early and Later Visits

To test correlations across time, Spearman's rho was computed across the variables (see Table 3). Given the low frequency of declarative and imperative pointing with eye contact, these variables were not analyzed.

With regards to Vygotsky's (1978) proposal, early reaching was not negatively correlated with later imperative pointing even when looking at only the end of the first year of life and with a larger sample. With regards to Butterworth's proposal, both early and later reaching were in fact associated with later declaratives (cf. Franco & Butterworth, 1996) (see Table 3). With regards to Tomasello's proposal, 10 of the 23 infants (43%) pointed (either declaratively or imperatively) at a 9, 10, and/or 11 month visit but did not point at 12 months.

Table 3: Correlations between Variables at Early (9 and 10 m) and Later (11 and 12 m) Visits

	ER	ED	EI	LR	LD	LI
Early Reach (ER)	---	-.10	-.34	-.01	.49*	.16
Early Declarative (ED)	---	---	.74****	-.34	-.27	-.17
Early Imperative (EI)	---	---	---	-.36	-.42*	-.28
Later Reach (LR)	---	---	---	---	.59***	.47
Later Declarative (LD)	---	---	---	---	---	.55**
Later Imperative (LI)	---	---	---	---	---	---

**** $p < .001$; *** $p < .005$; ** $p < .01$; * $p < .05$

5.4 Discussion

There were no significant differences in reaching frequency across the observations in this study. Thus, the statistical flat line of infant reaching that Butterworth and Franco (1996) documented in the second year seems to extend back into the first year. This seems to make "Vygotsky's hypothesis" untenable and in either case reaching and pointing frequency were not negatively correlated across the early and late visits as one would expect if this notion were correct. Despite their imperative utility, reaching and imperative pointing were in fact not significantly correlated at either time point or across time points in this study. Reaching and declarative pointing were significantly positively correlated such that early and later reaching were associated with later declarative pointing. Early reaching was however not significantly associated with

early declarative pointing. Franco and Butterworth (1996) reported that these variables were unrelated in the second year of life and Butterworth (1998a, 2003; Franco & Butterworth, 1996) has used these data to argue in part that declarative pointing is independent of an imperative motive. The present study suggests that they are not independent towards the end of the first year of life -- the time at which protodeclaratives first emerge.

Considerable within subject variability in pointing was observed in the present study, with 43% of the sample pointing at one or more than one of the 9- to 11-month visits and then not pointing in the 12-month visit. Despite the non-empirical nature of such a hypothesis, this is not what would be expected from a theory predicated on the view that an insight into the nature of intentions drives the emergence of pointing. As mentioned earlier, by ignoring within-subject differences in the age of emergence of the various joint activities that require attention (including pointing), Carpenter and colleagues (1998) tilted the statistical scales towards supporting Tomasello's (1995, 1999a) theory. And Carpenter et al. would not have had the statistical support for these conclusions if one assumes that a comparable amount of within-subject variability would have occurred in their studies.

Some might claim that some support for Tomasello's is found in the fact that the two forms of pointing observed in this study were quite highly intercorrelated across both the early and later visits. However, Tomasello's theory is based on the idea that there is a lot of overlap in the onset of different joint attention behaviours -- not that particular forms of a given behaviour would be highly correlated. What the present data suggest is that infants that point a lot declaratively, also pointing a lot imperatively. Or to put it

another way, some infants point a lot and some do not. As I understand it, Tomasello's theory would have no stake in this one way or the other.

A limitation of the present study is that the sample is still quite small. And one should be conservative about drawing any conclusions based on such a study. However, the sample size was comparable to Butterworth and Franco's with the added power of a within-subject design. It was also the same size as Carpenter et al. (1998). In their study of 7 younger infants, Franco and Butterworth (1996) reported that infants rarely reach in a strictly declarative context. Thus, another limitation of the present study might be that the "imperative and declarative contexts" were confounded. Further work looking at this issue in the first year might help to resolve the discrepant findings in the present study and in Franco and Butterworth (1996). For now, however, the present results should give us pause to consider whether protodeclaratives evidence a pre-existing primitive nascent prelinguistic language-specific referential system.

This raises the related issue of whether imperative and declarative acts should be seen to demonstrate the existence of independent motives on the part of the infant. As noted earlier, a standard claim is that infants declaratively point *because* they have a motive to share attention with others. Liszkowski and colleagues (2004, p. 306), for example, conclude a recent study on declarative pointing by arguing:

> this study suggests that infants at 12 months of age do understand something about attention.... With this understanding, infants actively engage in mutual informational exchanges about events in the outside world with the goal of sharing attention and interest.

But this empirical conclusion conceals logical confusion. Infants are able to socially interact in particular ways at 12 months of age that they are unable to earlier in life. Because of this, researchers make claims about what they understand about other

people. And perhaps the claim that they "understand something about attention" is a harmless, even quite accurate, one.

The next step in this line of reasoning, however, is the causal and Cartesian one, namely that infants do such and such in the world *as a consequence* of this ability that we attribute to them as a function of their activity given their shared practice with another. Because of this, what distinguishes an imperative from a declarative act is no longer some shared pattern of activity but rather the mental state of the child. And the differences in ability that we then see in developing agents become attributable to the infant's mind instead of that "mind" being a non-local distributed property of an agent's activity in a world of people and objects. Liszkowski et al.'s (2004) conclusions, like many other similar ones, are correct descriptions of activity. But they are parasitic on a particular view of mind, a view that they cannot help but perpetuate.

The results of this study do not clarify the exact manner in which infants get into shared activities that involve pointing. Although it adds counter-evidence against the grasping-pointing particular view of social shaping, the amount of within-subject variance in pointing seems most consistent with a gradual, non-insight based, account of social development with which the present account runs parallel. Also, Tomasello (1999b, 2003) himself argues that pointing can occur through social shaping. But perhaps in deference to his own theory and its philosophical presuppositions he mainly seems to use ontogenetic ritualization to only account for across species differences. In any case, there is no support for some abrupt stage-like shift in infant pointing gestures that one might expect if it were mediated by some "cognitive revolution" on the part of the infant.

CHAPTER 6 DEVELOPMENT OF POINTING IN INTERACTION

In the second study in this dissertation I employ a micro- rather than macroanalytic strategy. This is because other than Bates and Bruner's pioneering work, the issue of how pointing comes to be recruited and used as an interactional (and what is often, and derivatively, understood to be a referential) device is largely absent in empirical psychological investigations of pointing. This is addressed in Study Two by relying chiefly on participant meanings as an indicator of communicative function. Thus, the sequential turn-by-turn structure of mother-infant interaction is analyzed in order to determine the ways in which mother and child manifestly treat one another's communicative bid. A premium is placed on providing examples to justify any conclusions that might be drawn, and by being careful to distinguish the analyst's role from the participants'. In this sense, the present study is therefore different than the extant diary studies mentioned earlier. The goal is to chart the development of the interactional resources that one 9- to 12-month-old infant has at her disposal when communicating with her mother. To accomplish this, the interactional roles that infant pointing, pointing-like and reaching play in interaction are displayed. And the mother's skilled use of pointing gestures becomes an interesting contrast for the infant's. In fact, a most striking thing about these interactions is the degree to which the mother's pointing gestures are seamlessly intertwined with the sequences of joint activity.

6.1 Introduction to Study Two

I take an inductive approach that relies on the methods of social pragmatics (Turnbull, 2003; Turnbull & Carpendale, 1999), which is in turn derived from ethnomethodological (EM) and conversation analytic (CA) approaches to social interaction (e.g., Antaki, 2004; Atkinson & Heritage, 1984; Coulter, 1991, 1999, 2004; Garfinkel, 1967, 1988; Jayyusi, 1991; Sharrock & Coulter, 2004; Wootton, 1990, 1994, 1997). Conversation analysts study the orderly structure of conversation in order to show how interaction is accomplished. Unlike EM/CA, a social pragmatic approach places explicit emphasis upon the structuring of environments so as to observe interactional patterns of research interest. I will make claims about participant meanings because CA produces evidence that interlocutors manifestly treat/do not treat each other's contribution (i.e., prior turn) with an identical understanding. Although participant meaning is defeasible in the sense that a participant might opt to not correct a misunderstanding, it does presuppose the ability to experience a shared state of affairs from the point of view of another. This, at the very least, avoids the indeterminacy of an purely observer frame of reference.

Wootton (1990) departs from a standard conversation analytic approach when investigating parent-infant interaction because he seems wary of the standard 'second turn proof procedure' wherein the meaning of a previous infant turn would ordinarily be seen as conditioned by the subsequent response turn of the parent. This is understandable given that parents might treat behaviour with no clear goal as e.g. a request (e.g., Kaye, 1982). Wootton's solution is to focus on third position repair sequences (i.e., where an infant has acted, the parent has responded to that action in

the second turn/position, and the infant responds to that response in the third turn/position) wherein an infant is an unambiguous position to display her congruence or lack thereof with the parent's previous turn, a contingency that Wootton reports 12-month-olds can manage. However, my interest is in part in those very sequences wherein a parent is treating what might to an observer appear a random behaviour or an unclear bid as purposeful nonetheless. As such, I will pay attention to how the parent treats the infant's prior turn while nevertheless being careful about what this means about the infant's contribution. A prior concern for me, however, will be to first reconcile the logic of ethnomethodological investigation with a Wittgensteinian orientation.

6.1.1 Ethnomethodological Approaches to Meaning

As detailed in earlier sections, to Wittgenstein, psychological meaning is a process of direct perception given a mastered practice wherein logical criteria justify the application of particular psychological predicates to particular patterns of activity. EM/CA approaches to psychological meaning assume that the psychological meaning that participants attribute to their interactions is manifest in the details of that sequential and negotiated activity. What I will call the weaker, and to my mind dualistic, EM/CA thesis is that its approach renders the ontology of mind moot because their empirically driven method cannot address such issues given the impossibility of specifying what, for example, a person's intentions are without recourse to the details of their interactions with others (e.g., Bilmes, 1992). However, the stronger and non-dualistic EM/CA thesis is that intention is a way of speaking about activity and there is no separable mental event of which to speak (e.g., Coulter, 1991).

Despite the fact that some EM/CA approaches explicitly draw on Wittgenstein (e.g., Antaki, 2004; Coulter, 1999, 2004; Jayyusi, 1991; Sharrock & Coulter, 2004), the direct and logical Wittgensteinian approach and the more interpretative and empirical EM/CA tact might be seen to contradict one another.[38] Some practitioners of EM/CA emphasize the discovery of meaning, which makes meaning seem empirically bound and derived. However, logical criteria are involved in both approaches due to: (a) the dependency of particular interactional sequences, and (b) the relation between the meanings of linguistic terms and the particular circumstances that justify their use. With regard to (a), the notion, for example, of turn adjacency pair structure in conversation is such that a question conditions the next turn such that an answer is made relevant. With regard to (b), the meaning of interaction (linguistic or otherwise) is inseparable from the enacted embodied practices with which it is tied up. These two related notions are instantiated in what conversation analysts have termed 'participant meaning' -- the meanings that interlocutors attribute to their joint activities. And participants are able to have the meaningful exchanges that they do by virtue of (a) and (b).

As Garfinkel (1988, p. 108) put it, in an admittedly different context, "these phenomena are locally and endogenously produced, naturally organized, reflexively accountable in and as of detail, and therein they provide for everything that detail could be." And the practices in and through which we understand psychological concepts are active ongoing achievements that are accomplished "for another first time through" each time they are performed (Garfinkel, 1967). I take Garfinkel to use this idiosyncratic turn

[38] Sidnell (2003) compares Wittgenstein and Garfinkel with respect to their positions on rule following and notes that while these approaches are non-identical they are nonetheless complementary. See also Coulter (1999, 2004).

of phrase (and many others) because to do otherwise would raise the spectre of idealism. Similarly, from a Wittgensteinian perspective: "The rules here are immanent, not transcendent" (Baker & Hacker, 1985, p. 154). In a clarification of his use of Wittgenstein and Garfinkel (in Coulter, 1999), Jeff Coulter (2004, p. 337, his emphasis) notes that:

> Garfinkel's sociological programme explicitly eschews appeals to what is 'beneath the skull' of agents...but such a position needed (I argued) to be explicated as either (1) a methodological postulate or (2) an intellectual commitment with ontological status. *I argued for the latter*, and claimed that Wittgenstein's later logical work grounded the case.

I would argue that to get his ideas of practice off the ground Wittgenstein presupposes the "interactional structures" that ethnomethodologists and conversation analysts have "discovered." Conversely, conversation analysts are reliant on such practices to license expectations about interaction. And although humans directly see the meaning of psychological concepts in and through their mastery of particular practices in which the concepts are grounded, there are frequently occasions where the meaning of an action is underspecified by practice and interlocutors then find themselves in "a process of interpretation" until they either co-construct an emergent practice or else apply an extant one to the activities in question (Wittgenstein, 1958, part II).

As noted in earlier sections, developmental theorists often claim that clear signs of "intersubjectivity' are when infants engage in joint activities requiring attention with others (e.g., Bretherton et al., 1981; Bretherton, 1991; Trevathen, 1979; Trevathen & Hubley, 1978). And although Tomasello (1995, 1999a) does not predicate his theory on the use of this term, he theorizes that infants recognize their psychological similarity to others and attribute to them a capacity to recognize that others assign the same

meaning to their actions as they themselves do. I noted though that a problem with non-language using agents is the impossibility of first person avowal as criteria for understanding. And I argued that far from this being a technical problem it was in fact fatal for theories that claim infants understand their own mental states or those of others. Thus, the issue in front of us is not whether infants understand mental states (or their psychological similarity to others and so on) but rather whether there are grounds for claiming that shared meaning exists in their interactions. Or what we might alternatively call a practical understanding of shared practice or an understanding in feeling or perhaps some other term. The point is that we are claiming that infants can interact with others in ways that might justify the use of mental predicates not that the infants themselves understand such psychologica. We could settle this criteriologically, but given the slippery use of understanding that abounds in the field this runs the risk of keeping us in an observer frame of reference. This is where a Wittgensteinian CA can be useful.

As Reimers and Fogel (1992, p. 82) have put it, "intersubjectivity is a shared understanding of what an interaction is about." Participant meaning gives manifest evidence that interlocutors do/do not attribute roughly the same meaning to their interactions. Participants co-construct particular practices with one another in their interactions and make sense of each other in so doing. Another way of thinking about conversation analysis then is as a means of using an interactional sequence to instantiate criteria in ongoing interaction. For example, in the sequences of interaction involving my daughter T discussed in a previous section, I showed how criteria for the application of intention and desire were warranted in the contingencies of T's

interactions with her mother *and* I also showed that T's mother attributed intention and desire to her daughter's activity. Mother and child were both manifestly following the rules of these games. Thus, in this dissertation I need not generate explicit sets of criteria to justify the conclusions I draw because CA provides a convenient shortcut. We simply need to look at the degree to which shared meaning is apparent in sequential understanding. In so doing, we and the participants, to the degree that they possess reasonably complex language, can see the criteria that are contained therein.

Shared meaning ("intersubjectivity") however, cannot be some amorphous construct that we speak of in infancy. There will not be shared meaning but shared meanings. And to speak of states of primary (or secondary) intersubjectivity from my point of view is a non-sequitor. How can an infant have a shared understanding of what a particular interaction is about before having participated in? No, we need look at the infant's understanding of particular activities, and in this case particular joint activities that require attention. The canonical activity of interest is pointing. We would have grounds for saying that an infant has a shared understanding of pointing when she follows another's pointing gesture and when she uses pointing gestures to direct another's attention. This will necessarily involve an investigation of the relation between gaze and gesture. Using CA, we can look for evidence of this. And shared meaning will be to respond appropriately to the previous turn and to treat the other's turn as non-problematic; that is, to not initiate a repair of that sequence. Where this differs from some other accounts is on the central role of sequential understanding as opposed to discrete behaviour.

As noted, conversation analysis takes an inductive approach to investigating interaction. Thus, advocates of this methodology do not believe it is advisable or indeed possible to specify a priori in any concrete sense the outcome to which an investigation might lead. However, the manner in which the interactions are minimally structured should elicit spontaneous communicative bids by the infant within the constraints of the infant's abilities to so do. And I will focus on the particulars of mother-child interaction when reaching, pointing and cognate activities occur.

6.1.2 Expected Results

As noted, conversation analysis takes an inductive approach to investigating interaction. Thus, advocates of this methodology do not believe it is advisable or indeed possible to specify a priori in any concrete sense the outcome to which an investigation might lead. However, the manner in which the interactions are minimally structured should elicit spontaneous communicative bids by the infant within the constraints of the infant's abilities to so do. And I will focus on the particulars of mother-child interaction when reaching, pointing and cognate activities occur.

6.2 Method

6.2.1 Participants

One mother-daughter dyad chosen at random from Study One was observed when the infant was again 9, 10, 11, and 12 months of age. The dyad in this case was middle-class and Caucasian and the infant was second-born.

6.2.2 Design and Procedure

A research assistant videotaped a free play interaction and a meal interaction for 15 minutes each. Interaction situations were chosen so as to maximize the opportunity to observe instances of infant gestural communication and its antecedents (Turnbull, 2003; Turnbull & Carpendale, 1999). The mother was asked to encourage her infant to stay within the field of the cameras during the free play sessions. The sequence of the two interaction tasks was not counterbalanced because the experimenter could not control when the ideal time to feed the infant might be. The structured observation described in Study One was administered after completion of the interaction tasks. Thus, separate camera angles were again used so that I could select the best angle on a given interaction when transcribing the sequence.

6.2.3 Free Play and Meal Interaction

A research assistant provided the dyad with 3 similar but non-identical sets of 6 toys (colourful stacking cups, a doll, a puzzle, a picture book, a toy telephone, a rattle with movable parts) at each session based on those used by Carpenter et al. (1998) and Bakeman and Adamson (1984). The sets were rotated such that on no two consecutive visits did the dyad play with identical toys. The mother was asked to play with their infant as she normally does. After 15 minutes had elapsed, the other interaction task was set up and taped. The mother was asked to feed her infant a meal at some point in the visit while the infant was seated in a highchair.

The video records of interest were approximately 240 minutes in total observed from the two camera angles when the infant was 9, 10, 11, and 12 months of age. These videotapes were extensively reviewed and the best angle of the two possible was

used to transcribe the interaction. The entire sessions were not transcribed verbatim from beginning to end but rather particular sequences of interest were selected for transcription after many passes.

6.2.4 Transcription Conventions

The conventions of conversation analysis were designed for use with language users, and therefore turn-by-turn transcription also included gaze (see e.g., Schegloff, 1984) and manual gesture channels (see e.g., Wootton, 1990, 1994). M represents mother; C represents child. Vocal (V) elements of the interaction, gaze (G) and positions of the right hand (RH) and left hand (LH) are identified on separate lines. The temporal dimension of the interaction is represented by movement from left to right along these lines and the horizontal relations between the various channels were kept as tight as possible. Manual actions are surrounded by ( ). Overlap in actions (including speech) is surrounded by [ ]. Some less relevant aspects of the interaction are simply summarized as gloss transcriptions. Mother's hand gestures and positions are noted only when they make relevant the infant's behaviour. In cases of an empty line (typically when either M or C did not vocalize), the line is not listed. Examples from same visit will be given the same numeral with differing subscripts to distinguish the examples within a visit. Examples from visit one will be referred to as Example 1a, 1b and so on, whereas examples from visit two will be referred to as Example 2a, 2b and so on, etc.

6.2.4.1 Vocal

All talk is transcribed orthographically and there are not phonetic transcriptions in the case of infants' utterances. The reason for this is that precise phonetic

representation was not necessary to capture the aspects of the interaction that are relevant to the arguments presented. Timings (in round brackets) have meaning only within the vocal line. Pauses under half a second are marked in the following manner, (.). Further conventions include: [to represent point of overlap, : to represent stretching of a sound in an utterance and ↑ to represent rising intonation in pitch contour. Cases where a vocalization was whispered are sounded by °°. Audible inbreathing or outbreathing accompanying utterances is represented by > and < respectively. Cases where a vocalization was inaudible the utterance are surrounded by ( ).

6.2.4.2 Gaze

Gaze was inferred from orientation of the head. Cases where the target of gaze was less clear are surrounded by ( ). When targets are not stated in the gaze line, M represents gaze at parent and C represents gaze at infant. Dashes (- - -) represent sustained gaze at a target, whereas commas (, , ,) represent shift of gaze.

6.2.4.3 Hand Positions

Dashes and commas are used in a manner analogous to gaze with (----) representing static positioning of the hand (RH or LH) in initial location designated in the line and commas representing movement of the hand from an initial position to another in the sequence. The following additional conventions are employed: M = parent, C = child, Re = reach, Gr = grasp, Po = point, PoT = tipping point, and PoS = slip out point. Tipping points occur when the infant uses her index fingertip to explore aspects of their world at hand (Shinn, 1900/1975). Slip out points occur when a pointing hand occurs ("slips out") during the infant's activities but which the infant does not incorporate into

their ongoing interactions (e.g., Lock et al., 1990).[39] In cases where a hand position is indicated with no accompanying referent, this is meant to signify the presence of a particular hand shape.

6.2.4.4 A Note on the Coding of Affect

Before I present the results I want to explain why affectivity is not explicitly taken into account in each interaction. I could have coded for smiling and other "expressions of emotion" in both parent and child and possibly have found some contingent relations. What I will say is that affect was certainly omnipresent in these interactions and I would have thrown out these data as highly suspect if it were not. However, affect is not some separate aspect of interaction to be accounted for but is rather a deep part of interaction. And when it is absent, it is very apparent to an interlocutor (Hobson, 1993, 2002).[40]

6.3 Results

I present the findings first grouped by month of visit. When discussing general themes below I collapse across the various sessions. My goal is to first allow the initial grouping of these transcriptions to serve as a record of what is typical of mother-child interaction for this particular dyad at a particular visit. I discuss the themes of these interactions in the Discussion.

[39] Lock and colleagues (1990) would refer to both of behaviours as "pointless points" because they serve no apparent function. Thus, it will be interesting to see how they are treated in ongoing interaction.

[40] There was in fact only one sequence (1e) where I noted emotion because it is possible that the reason that the infant in that case began to eat again was that her mother was obviously quite ticked. However, even here the M does not treat it that way and I did not make note of it in my commentary.

6.3.1 Ninth Month Visit

6.3.1.1 Example 1a

M and C have just been given a bag of novel toys by E2 at the first visit. C has just pulled the bag of toys in front of her and is sitting in front of it with her hands on the strap of the bag. During the sequence C places her hands down on the bag (palm first) and pushes off of the bag presumably to elevate herself. What is notable about this interaction in general is not just the way the mother initiates and monitors the activity but indeed the way she shapes the interaction so that her baby discovers the properties of a toy to which the latter will subsequently point. C's concomitant tipping behaviour is not linked to a prior turn nor does it project to a subsequent one.

1 C: (V) <mhha <hhh

(G) bag - , , M's RH - - - , bag - - - -

(LH) Gr strap - , bag , open palm over bag , , on bag - - - - - - - - - - - - - - - - - -

(RH) Gr strap - , bag , open palm over bag , , on bag - - - - - - - - , , at side - - - -

M: (V) °whats at↑° (2.0) boobie↑ (1.5) rachelrachel (1.25) whats this↑

(G) C; bag - - , , C -

(pulls doll from bag)

---

2 C: (V) uhh

(G) doll -

(LH) at side , , towards doll fingers apart - - - - - - , Gr doll - - - - - - - - - - - - - - - -

(RH) at side , , PoS , Po, PoT - - - - - - - - , to side , , Gr doll - - - - , PoT - - - - - -

M: (V) <hhh:hhh> (2.0) its a spon:gebob

(G) C -

(holds doll 12" from C's face) (releases doll)

---

3 C: (G) doll -

(LH) places doll on bag and supports , , faces doll towards self - , to face - - - -

(RH) Po - - - PoT - doll face - - - doll body - - - - - - - - - Gr doll, , to face - - - - - -

M: (G) C -

6.3.1.2 Example 1b

After C has sucked on the doll and explored it with her fingertip on two more occasions (once a simple touch, once a sequence of touches) over a span of about 40 seconds, M directs C's attention to the further contents of the bag of toys. Unlike most of the later visits, although C overshoots the referent C still treats M's point as an attempt to direct C's attention. And again, the tipping seems unrelated to prior or subsequent turns.

1 C: (G) doll - - - - - - - - - , , M - - - - - , , , M's RH - - , , (beyond bag) - - - , , bag - - -

(LH) on leg - Re withdraw -

(RH) on leg -

M: (V) °what else is in the bag↑° (3.0) um (.) hum:uh hum

(G) C; bag - - , , C - , bag - - - - - - - - - - - - - - -

(RH) on ground - - - - - Po - Po in bag - -

---

2 C: (G) bag - - - - - - - - - - - - , phone - - - - - - - - - - - - , , M - , phone - - - - - - - - -

(LH) on leg -

(RH) on leg -

M: (V) °whats tat↑°

(G) bag - - - - - - - , , , , C -

(pulls phone from bag, puts on floor) (dials with index fingertip)

---

3 C: (G) phone -

(LH) on leg -

(RH) on leg - - - - - - - - - - - - - - , , dialling with fingertip - - - - - - - - - - - - - - - -

M: (V) it's a rotary phone (1.0) youve probably never seen a rotary phone

(G) phone , , C -

(RH) (dials again with index fingertip)

(LH) Po - - - (circular dialling motion in air)

6.3.1.3 Example 1c

M and C have been playing a game of "talking" into the toy phone. Although again unrelated to prior and subsequent turn, the tipping behaviour featured in this sequence is treated by M as a request for an object label. Although this suggests that Wootton's caution regarding the third term proof procedure is well advised, it also shows that parents richly interpret their children's activity and in so doing may help to create the very expectations that they, the parents, perceive in the interaction.

C: (V) ahh

(G) M - - - - - - , , phone - - , , (forwards, no clear referent) , phone - - - - - - - -

(LH) on leg - , , , PoT - - - - - - - - - - - - - - - - - - -

(RH) on leg , Gr phone - holds parallel to ear - - - , , leans on M's leg - - - - - - -

M: (V) here (.) its grandpa (.) hello rachel↑ (1.) (inaud) yu:h (.) that's the phone

(G) C -

(hands phone to C)

6.3.1.4 Example 1d

M and C are seated in front of the remainder of the set of visit 1 toys that M has now pulled out of the toy bag. A set of nested stacking cups was pulled out first, followed by a soft small truck, a picture book, a cardboard puzzle and a rattle. After picking up the book and unsuccessfully attempting to open it while M retrieves the remainder of the toys, C orients to the stacking cups. C is seated within M's legs while M builds the tower. M uses a pointing gesture to direct attention, but C is unresponsive. M treats C's behaviour of grasping a cup in the stack as contingent upon hers, but it is not. On a frame-by-frame analysis C is grabbing a particular cup from the stack and in so doing the stack is falling over. M reacts to C's activity as an attempt to push over the stack. (M reacts this way in this sequence and one subsequent that is not transcribed.) As in 1c, we can see how parents' "misattributions" of their children's intentions may help to structure the latter's mastery of practice. However, I would claim that M is not cognizing that the infant is thinking such and such but is rather enacting a particular practice in and through which a child will develop.

1 C: (V) [ooh:a]

(G) C's cup - - - - - - - - - - - - - - - , , (forwards, no clear referent) , M's cup - - -

	(LH)	on leg -
	(RH)	Gr C's cup - - , brings to mouth -
M:	(V)	one [(2.2)] °oh is it my turn↑ (.) here° (3.0)
	(G)	C -
	(RH)	Gr cup , puts on floor , Gr cup, floor
	(LH)	Gr stack -

---

2 C:	(G)	(forwards, no clear referent) , C's cup -
	(LH)	on leg - , , Gr Cup - - , to mouth - - - -
	(RH)	Gr C's cup - , in mouth - - - - - - , , in front of face - - - - - - - - , to mouth - -
M:	(V)	oh these are good ones (1.0) they go ve:ry tall (.) we dont have these kind
	(G)	stack -
		(builds tower of cups)

---

3 C:	(V)	o[oh:a] (.5) ohah (.) owah
	(G)	stack - - - - , C's block - - - , , stack -
	(LH)	holding cup in mouth - , Gr stacked cup - - - - -
	(RH)	holding cup in mouth - - - - - - - , hold/drop C's cup, Gr stacked cup - - - - -
M:	(V)	[ho ho wha ho] (3.5) push↑ (1.5) can you push it over↑ (0.5) good girl
	(G)	stack - - - - - - - - - - , , C - , stack - - - - - -
	(RH)	Po stack - - - - - - - , , PoT stack (demonstrating push) - - - - - , ground - -

---

6.3.1.5 Example 1e

In this sequence, M is feeding C a meal. The sequence and four subsequent ones that I have not transcribed are notable for the presence of extended periods (10 – 15 seconds) of tipping behaviour. In all of these cases, the tipping is again unrelated to prior or subsequent turn.

1 C: (G) M - - - , , tray - - - - - - - - - - , M - - - - , camera - - - - - - - - - - - - - - - - - - -

(LH) Gr edge tray -

(RH) drops melon - - - - - - - - - - , PoT melon -

M: (V) hum (.) thats good stuff (2.0) no (.) nat one hit the ground that's gahhbage

(G) C - , , (towards melon on floor) - - - - - - - - - - - - - - - - -

(M gets ups and picks up melon)

---

2 C: (G) camera -

(LH) resting on tray -

(RH) -

M: (V) now this ones garbage now (.) hum↑

(G) C -

6.3.1.6 Example 1f

C has been playing with food bowl for approximately 90 seconds. C drops it on the table and it lands upright. As M moves it back towards C, the latter throws a piece of melon onto the floor followed by the bowl. There are three LH slip out points during this exchange, all of which occur while C is grasping an object with the RH. As this RH

tipping in combination with LH grasping behaviour was easier to observe in these sequences, I went back to earlier ones and found that in RH tipping the LH was in all cases already grasping. Again, the tipping and slip out pointing might be better thought of as "pointless behaviour" than as shared pointing activity. Although it may be non-random, it does not seem have any interactional significance to either M or C. In this sequence, M also demonstrates a novel use of a reaching gesture in this sequence, namely to get C's attention after C drops the bowl.

1 C: (G) M - , ground , , bowl - - - -, , ground -
(LH) PoS , , open hand - - - - - - - - - , PoS - , open hand - - , PoS , , open hand
(RH) Gr melon - - - , , drops off tray , Re + Gr bowl , , drops , Gr side tray - - - -
M: (V) mm (.) pretty good (2.0) not so much↑ (mouth opens in surprise)
(G) C -

---

2 C: (G) ground -
(LH) PoS , on tray - - - - - - , , over tray - - - - - - , on tray - - - - - - - - - - - - - - - -
(RH) - , over tray - - - - - - - - - - - - -
M: (V) (jaw clenched in irritation)
(G) -
(LH) tray - - - - - - - - , , Re , brings to C's cheek - - taps cheek with fingers - - - - - - - - -

---

3 C: (G) M - - - - - - - - - - - , melon - , tray - , M - - , (to side) - - - - - - - - - - - - - - - -
(LH) over tray - - - - - - - - , Gr melon - - , to mouth - - - - , , Gr melon , to mouth
(RH) -

M: (V) ahaw:ooh

(LH) taps chin with fingers , , back to lap

6.3.1.7 Example 1g

M treats C's outstretched and raised right arm as hands up gesture. It is not clear that C treats it in the same manner. When compared to example 3e, this asymmetry becomes particularly apparent.

C: (V) <<uh (.) oAAHHH aaoooo hhahhh ahhh (.) ma [(.5)] ma

(G) tray - - - - - - - - - - - - - , M -

(LH) pushing bib , , Gr tray - - - - - - - - - - - - - - - - - - , on M's hand - - - - - - - - -

(RH) pushing bib , , Gr tray , , , Re M - - - - - - - - - - - - , on M's hand - - - - - - - - -

M: (V) oh i know dodo (5.5) [i know] ma ma ma ma

(using bib to wipe C's hand) (undoing belt in highchair) (pulls C out of chair)

6.3.2 Tenth Month Visit

6.3.2.1 Example 2a

M again sets up the phone game. C's eventual mastery of this practice becomes a situation within which she can point in a subsequent visit. C is able to coordinate her activity with M's relatively seamlessly within this routine.

1 C: (V) euhh=qa

(G) phone base - - - - - - - - , , phone receiver - - - - , , base - - - - , , phone - - -

(LH) on floor - - - - - - , , on lap -

(RH) Re + Gr phone , bring to R ear -

M: (V) =hello↑ (.) whos on the phone↑

(G) C -

(puts C in lap) (pulls phone from bag, puts receiver on it)

---

2 C: (G) receiver - - - - - - - - - - - - , M -

(LH) on lap -

(RH) holds phone to M's ear -

M: (V) hello (.) its dadi (.) say hi dadi (1.0) ey:ll talk (.5) hi (.) hi dadi how are you

(G) C -

(grabs C's hand holding phone and moves to own ear)

6.3.2.2 Example 2b

M tries to distract C from the phone game. After mixed results, M uses a pointing gesture to direct C's attention to the other toys in the bag. C does not overtly respond to the pointing gesture by shifting her gaze.

1 C: (G) bag -

(LH) at side -

(RH) Gr phone (in mouth) -

M: (V) you wanna see whats in here↑ (.) whatsat↑ (.) °looka looka itsa surprise°

(G) bag - , , C - - , , bag - -

(LH) ground, , Po (bag)- - - - - - - - - - - - - - , Re bag , Gr puzzle piece, on chin

---

2 C: (G) bag - - , , M - - - - - - - - - - - - - - - - - - - , , bag - - - - - - - - - - - - - - - - - - -

(LH) at side -

(RH) holds phone to M's ear - , , to C chin - - - - - - - - - - - - - - - - - - - , to bag -

M: (V) (1.) look (.) oh (.) hello (.) yes (.) hi grandma (.) im just playing (.) do you

(G) bag - - - - , phone - , (forward) - , C - - - - - - - - - - -

(LH) hand under chin - , Po bag - -

---

3 C: (G) bag -

(LH) at side - , ground - - - - - - - - -

(RH) Gr phone - - - - - - - - - - - - - - - - - , , , drop phone , , , ground - - - - - - - - -

M: (V) want to look and see whats in there↑ (.) whats in there? (3.5) what is it↑

(G) C- , bag -

(LH) - , Re + Gr doll - - - - - - - - - - - - - - - -

6.3.2.3 Example 2c

M has taken a large 4-piece puzzle out of toy bag. M is placing pieces in the puzzle board while C has inspects a rattle M has previously brought from the bag. C tosses it aside and approaches the puzzle. M repeatedly points during this activity to direct C's attention to the pieces of the puzzle. And M provides object names as she does so. There is again no evidence however that C is looking at the pictures as a result of M's point because she does not again take the point into account in her gaze.

1 C: (V) [erhs] (2.0) heu

(G) puzzle board - - - , , piece in hand - , board -

(LH) at side - , ground - - - - - - - - - -

(RH) at side - - , , Gr piece - - - - - - - - - - - - - - - - -

M: (V) [mm] (2.5) yeah i know (4.5) (this is) rainy (.) sunny (1.0) cloudy (0.7)

(G) C- , bag -

(RH) building puzzle - - - - - - - - - - , Po puzzle piece -, Po piece - , Po piece - ,

---

2 C: (G) piece in M's hand - - - - - - - - , M - - - - - - - - - - - - - , (down) - , , rattle - - -

(LH) at side - , ground - - - - - - - - - -

(RH) Gr phone - - - - - - - - - - - - - - - - - - , , , drop phone , , , ground - - - - - - - - - -

M: (V) what was this one↑ (.) windy >>phh (.) right back (.) this ones windy

(G) puzzle - - - - - - - - - - - , C- - - - - - - - - - , puzzle -

(picks up piece dropped by C)

---

3 C: (G) puzzle board -

(LH) at side - , (crawling) - - - - - - - - - - - - -

(RH) Gr rattle - - - - - - - - - - , , drops phone - - - - , (crawling) - - - - - - - - - - - - -

M: (V) you like that one↑ (2.5) aw:h youre just on that microphone thing

(G) puzzle -

6.3.2.4 Example 2d

M holds out a piece of a sandwich while seated in front of C at the highchair. While M withdraws it, C points with the index finger prominent but with the hand still

somewhat open, as M begins to reach for the food. M, but not C, treats this as a gesture for food.

1 C: (V) ENAh[hah:ha] ahhhh
(G) M - , , M's hand/sandwich - - - -
(LH) at side -
(RH) rubbing eyes - - , to lap - , PoS + Re - - - -
M: (V) ohh (.) need to have a nap↑ (2.7) [i know] (1.0) dyou want this one↑
(G) C- , bag -
(holds out food) (begins to withdraw)

---

2 C: (V) umm ma mi
(G) - , piece on tray
(LH) at side -
(RH) Gr food - - - , drop bread - , Re food - - , (to mouth) - - - - - - - - - - - - - - -
M: (V) this one↑ (2.0) you just want the cheese out of it (.) is that what you want↑
(G) C -
(holds out food again)

6.3.2.5 Example 2e

C has turned to the left and right in her highchair and made a series of high-pitched sounds that seem to express frustration to which M responds with a query about food. When M's request is made, C is leaning to the left hand side of her highchair with her RH grasping the tray. As opposed to the above, C places her elbow on her tray and

displays her open hand without any attempt to grasp the object and looks at her open palm while M does same. On this occasion, both M and C treat this as a gesture for food.

1 C:	(G)	M/banana - - - - , hand - , M - - - - - - - - - , , , banana - - - - - - - - - - - - - -
	(LH)	at side -
	(RH)	on tray , , open hand (knuckles down) - - - - - - - - - - - - , , turns hand over
M:	(V)	want banana↑ (4.5) yup (.) in your hand
	(G)	C - - - - - - - - - - , , C's hand - - , banana (peeling) - , , C - - - - - - - - - - - - -
		(touches C's open hand with her index finger)

6.3.2.6 Example 2f

C opens her hand and reaches for the banana in turn one of this sequence, but unlike the last sequence does not seem to intend it as a request. Although a slip out point occurs in the second turn that is orderly within the sequence, C and M do not treat is as a request for the banana. Rather, M is already in the process of providing C with the food when the slip out point occurs.

1 C:	(G)	(side, no clear referent) - - - - - - - - - - - , banana - - - - - - - - - - - - - - - - -
	(LH)	in lap - , on armrest - - - - - - - - - - - - -
	(RH)	in lap - , Re (open hand, banana) - - -
M:	(V)	you want more (.) you want banana↑ (2.7) do you want more banana↑
	(G)	C -
		(picks up banana and holds up to C)

---

2 C: (V) [deh] [heh]

(G) (side, no clear referent) -

(LH) -

(RH) - - - - , arm on tray , [PoS] , back on tray

M: (V) [(2.5)] yup [(2.0)] here ya go

(G) C -

(peeling banana)

6.3.3 *Eleventh Month Visit*

6.3.3.1 *Example 3a*

A striking thing about the 11 month visits in general is the changes in ability that C displays, which do seem to afford additional degrees of freedom for her contributions to the interaction. She is now walking quite ably. And other obvious aspects of dawning motor control are apparent in the fact that it is C not M who takes charge of the bag of toys at this visit. In the first example from this visit, M treats C as having uttered a word during play.

C: (V) [tullah]

(G) M - - - , stuffed toy - - , truck - - - - - - - - - - , , M - - - - - - - - - - - - - - - - - -

(LH) at side -

(RH) (pushes away stuffed toy offered by mother) , hand in front - - - - - - - - - - -

M: (V) nope (.) not that one (.) this one↑ (.5) truck [(2.5) (nods head)] truck

(G) C -

(RH) (taps truck with index finger 3 times)

(rolls back and forth) (rolls to C)

6.3.3.2 *Example 3b*

C does what would be probably coded as a declarative act with outstretched arm and index finger prominent but without all other fingers in palm. However, C seems to treat this as "asking for permission" and M clearly treats it to minimally signal intent. Thus, the "underlying motive" of C pointing gesture could be said to be imperative. The function of M's point seems to be a confirmation of some sort given that it is accompanied by the utterance "I know."

C: (G) camera - , , , M - - - - - - - - - - - - - - - - - - -

(LH) at side -

(RH) open hand (shoulder level) - - - - - - - - - , Po (camera) - - , lowers arm - - -

(walks towards camera, ~7' from where M is seated, stops 2' from target)

M: (V) no I don't think we can (chuckle) (1.0) yeh (.) i know (.) thats a camera

(G) C -

(LH) Re (towards C) - - - - - , (withdraws) - - - - , Po (camera) - - - , , , (lap) - - -

---

6.3.3.3 *Example 3c*

M points again during book reading. If C were to have done this act, it would have been coded as declarative. However, although C does not treat it this way, as the

analyst I would claim that it is taking on an informative function with respect to again providing object names.

1 C: (G) M/book - - - - - - - - - - - - - , , book -

(LH) at sides elbows bent - - - - - - - , , at side -

(RH) at sides elbows bent - - - - - - - , M's shoulder - - - - - - - - - - - - - , to book

(walks to M seated on ground with book, steadies self on M's shoulder, moves to book)

M: (V) wanna see the book↑ (.) yep (1.5) kay it says (.) table (2.0) thats the chair

(G) C - , book -

(LH) PoT (book-table) - , PoT (book-chair)

---

2 C: (G) book - , , (to side) - - -

(LH) (attempt grab book with open hand) -

(RH) (attempt grab book with open hand) -

(approaches book) (slips and unable to grasp) (loses interest)

M: (V) you climb on (.) and then that one (1.) theres the toychest charlotte has

(G) book -

(LH) - - , , PoT (book-toychest) - - - , (continues flipping pages)

6.3.3.4 Example 3d

With some direction from M, typically with indicative pointing gestures (accompanying "this one goes here" and so on), C has been stacking nested cups ad nausea and has just terminated another series. Upon finishing each of the series of

four, M cheers and applauds. After the final series, C imitates the clapping act with some effort. And M affirms the correctness of this conventional act.

C: (G) stack - - - - - - - - - - - - , , cup - - - - - - - - - - - - - - - - , hands , , - - - - - - - -

(LH) on lap - - - - - - - - - - - - - , steady cup - - , , holding cup - , , drop , clap - -

(RH) stacks cup , , withdraw , Gr cup - - - - - - - - - - - - -, drop - - , begins clap -

M: (V) ya:eeyyy ayyyy:Rachel (3.2) hmm↑ (1.5) yup (.) yayyy

(G) C -

(begins to applaud)

6.3.3.5 Example 3e

C and M both treats C's outstretched arms towards M as a hands up gesture, that is as a gesture to signal C's desire to be picked up.

1 C: (G) tray - , , M - - - - - - - - - - - , ,

(LH) holding food in mouth - , takes out of mouth -

(RH) on tray - - - - - - - - - - - - - - - - - - , drop more food - - , , on tray - - - - - - - -

M: (V) hmm (1.0) dropped on the floor (.) yeah you did (1.5) let me get it (2.5)

(G) tray - , C's lap - - - - , floor -

---

2 C: (G) food in hand - - - - - - - - - , M - - - , food - , M -

(LH) holding food - , Re (M) - - - - - - - - - - - - - - - - - -

(RH) holding food - , Re (M) - - - - - - - - - - - - - - - - - -

M: (V) yeh you were d:roginem (1.5) your dropping em (1.7) you just got in there

(G) C -

---

3 C: (G) tray - - - - - - , forward (no clear referent) - - - - - - - - - - - - - - - , , M - - - - -

(hands behind head) (trying to pull bib off) (bib comes off) (holds up bib)

M: (V) in there (1.0) kay↑ (.) take a bite (.5) and then say (inaudible + laugh)

(G) C -

6.3.4 *Twelth Month Visit*

6.3.4.1 *Example 4a*

M has taken the phone out of toy bag and C is holding it in her mouth with her RH. M is trying to direct C's attention to other toys in the bag. C's pattern of gaze suggests that C treats M's gesture as an attempt to direct C's attention. While changing her attentional focus towards the bag, C also has a slip out point before reaching towards the bag.

1 C: (V) =aoh

(G) (forward no clear referent) - - - - - - - - - , , bag -

(LH) at side - , , PoS - , Re bag - - - - - - - - -

(RH) holding phone -

M: (V) °whats in there↑° (.) °whats in there↑°=

(G) bag -

(LH) on lap - - - - - - , Po bag - , Po (slips to lap) , (in lap)

---

2 C: (G) bag - - - - - - , , doll - - - - - - - - - - - - - - - - , bag - - - - , doll - - - - - - - - - - - -

(LH) in bag - - , Gr doll - - - - - - - - - - - - - - - - - , (floor to get up) , at side - - - -

(RH) holding phone -

M: (V) <<ooh (.5) °its a spon:gebob°

(G) bag -

(pulls doll out of bag while C reaches for it)

6.3.4.2 Example 4b

M and C are again playing with the stacking cups. And again M is pointing frequently during the activity. This time C is not sticking the cups in her mouth. M is giving them to C one at a time while using her index finger to indicate where they should be placed. It is not clear that C is following the points or whether the points are just a superfluous part of the activity.

1 C: (V) [pwa]

(G) stack - - - - - , , cup - - , stack - - - - - - - - , , cup - - - - , stack - - - , cup - - - - - - -

(LH) at side - - - , , leaning on stack - , at side - - - - - , holding up , Gr cup - - -

(RH) Gr cup - - -, placing on stack - - - - - , Gr cup - - - - , stacking , Gr cup - - -

M: (V) now try this one [(2.5)] up top (3.5) now this one (5.0) good job (.) this one

(G) bag -

(RH) at side , , PoT cup - - - , , (makes hand shape of placing cup) - - , at side - - - - - -

---

2 C: (V) pwae

(G) stack -

(LH) - - - -- - , hold up , hand down - - , Gr cup - - , hold up - , , Gr cup/stack - -

(RH) - - - - - , stacking , hand down - , Gr cup - - , Put stack , , Gr cup/stack - -

M: (V) now that's quicker (.) and last one (1.5) °up top (.) way high° (.) yay::yy

(G) stack - - - - - , C - - , stack - - - - - - - - , C - - - - - , stack - - - - - - - , C - - - -

(applauds while C grabs block in middle of stack) (top half stack falls over)

6.3.4.3 Example 4c

C gets up from M's lap and again takes steps towards the forbidden camera. While she gets up out of M's lap, her RH pointing finger slips out. C takes 2 steps towards the camera with her RH in a grasping gesture. Neither M nor C incorporate the slip out point into their prior or subsequent turns. In the context of this activity, some might interpret the slip out point as a primitive expression of desire (cf. Lock et al., 1990).

C: (G) camera -

(LH) getting up - - , at side -

(RH) getting up - - , PoS - , , Re not fully extended - - - - , drops to side - - - - - -

(walks towards camera) (stops) (walks back backwards)

M: (V) <<hh:hy>> (0.5) °get over here°

(G) C -

6.3.4.4 Example 4d

M again points while book reading and attempts to provide object labels. M treats the fact that C closes the book after a few pages as a lack of interest.

1 C: (G) book - - - - - - - - - - - - - - - - - - , lamb - - - - - , , book - - - - - - - - - - - - - -

(LH) holding truck -

(RH) holding truck - , Gr book (trying turn page)

M: (V) on the farm (1.5) look (.) there was a lamb (.) bah:hh (.) beh (.7)

(G) book -

(RH) on lap - - - - - - - - - - - - , Po (book/lamb) - , PoT -

* * *

2 C: (G) stack -

(LH) holding truck - , drops truck - , at side

(RH) (turns page) - - - - - , Gr (book) - - - - - - , ruffles pages , , (closes book) -

M: (V) there were some chickens (.) cluck cluck cluck cluck (.) some flowers

(G) book -

(RH) (turning page with C)

* * *

6.3.4.5 Example 4e

C holds out a piece of orange towards M out of which M pretends to take a bite. M points at C to indicate that C should eat it. When C does not respond, M points directly at C's mouth. C then eats the orange. (This occurs once more during this mealtime visit.)

C: (G) M - , tray - , own hand - - , M - - - -

(LH) on tray - , Gr orange - - , on tray - - - - - -

(RH) holding out orange - - - - - - - - - - - - - - - - , withdraw - - , to mouth - - - - - -

M: (V) haoup> (1.5) (pretend bite) °no you do it°

(G) C -

(RH) Po (C) - - - , in lap (1.0) , , Po (C's mouth) - - - , lap - - - - - - - - - - - - - - - -

* * *

6.3.4.6 Example 4f

C is putting sections of orange left on her tray in a bowl, which is treated by M as a signal that C may be finished her meal.

1 C: (G) bowl -

(LH) in lap -

(RH) Gr orange, put bowl - -, Gr orange - , put bowl - - , Gr orange , put bowl -

M: (V) rachel all done↑ (.) no more (3.5) do you want more or are you all done↑

(G) C -

(M also begins to put food in bowl)

---

2 C: (V) ahhu

(G) bowl - , , M - - - - - - - - - - - - - - - , bowl - - - - - - - - - -, , M - - - - - - - - - - - - - - - - -

(LH) - - - - - - - Re (towards M) - - - - - , hand behind head, Re (M) - - - - - - - - - -

(RH) - - - - - Re (towards M) - - - - - - - , hand behind head - - - - - - - - - - - - - - -

M: (V) rachel (0.7) rachel (0.5) out (1.0) all done (.) out (1.5) you want more↑ No

(G) C -

6.4 Discussion

Three main aspects of the interaction detailed above strike me as central to the goals of this dissertation. First, consistent with Shinn's explanation of the origins of the pointing gesture, there is an abundance of evidence of C tipping at the first, and only the first, visit. Consistent with my criticism of Shinn, however, with two exceptions there is

no uptake on the part of the mother or the infant. And although it also seemed to be non-randomly paired with contralateral grasping, neither party treated it as meaningful *activity* and as such I would call it nonmeaningful *behaviour.* Now, on two occasions M provided object labels in accompaniment to the tipping but the infant did not treat M's turn as relevant to C's turn. This also seems to question Werner and Kaplan's account of the origins of pointing because the extensions of the fingertip that were observed were not treated by the infant as a means with which to orient their own attention. And if Werner and Kaplan meant more fluid fully-formed gestures then in this case there is still no support for this view in the one canonical pointing gesture that we do observe in this corpus because the infant looks to the mother after the point.

Second, the mother points on many occasions in these transcripts. In a sense this is trivial and it would be shocking to find a mother who did not use indicative gestures in her interactions with her child. However, despite the fact that I have argued that pointing emerges spontaneously, given the amount of pointing that infants are exposed to in early ontogenesis perhaps modelling has some role in its development (but see Butterworth, 2003). But even if hand shape were acquired by modelling to some extent, the use of pointing could not be. More to the point though is that there is no consistent turn structure evidence that the infant and mother attribute roughly the same meaning to pointing gestures. On one occasion the infant does gaze in the direction of the mother's point and seems to find the correct referent, but C does not re-check the direction of the point but rather fixates on the bag of toys -- a stimulus likely to engage attention on its own (Butterworth & Jarrett, 1991). However, it seems likely that

these sorts of pointing frames (Kaye, 1982) will enable the development of the child's mastery of this practice.

Third, although we do not overly see social shaping of pointing gestures in this corpus, we do clearly see the social shaping of a hands up gesture. In it M treats C as having a desire to be picked up and C does not repair this treatment. I would argue that this gesture could also be said to display C's intention that her mother pick her up/expectation that her mother will pick her up/thought that her mother will pick her up/perhaps even belief that her mother will pick her up, and so on. Obviously some of these mental predicates fit better than others, but they are clearly not differentiated by some co-occurring state of mind on the part of the infant. We also witness the emergence of a conventional clapping gesture in the same manner. In both cases M clearly treats these as acts and there can be no doubt that the baby intends to imitate an adult model in the latter case. We also see the acquisition of another conventional gesture in this corpus, namely the word for truck. And it is treated by both infant and mother as an "object name." But here we have to keep in mind the distinction between what Bates (1976) called meaning as act and meaning as object. The problem here again is that reference is a matter of a word or gesture having a role in a language-game, not of a bare referential relationship being established between a word and referent (Goldberg, 1991; Montgomery, 2002). In either case, shared meaning is observable in the way that M and C treat the acquisition of these conventional forms.

Many studies, however, have concluded that infants have an understanding of attention, be it ever so tenuous. And experimental evidence shows that infants can in fact interact with others in ways that suggest shared meaning. For example, infants as

young as 10 months have been reported to engage in gaze alternation, meaning that they alter their gaze between object and other thereby showing some sensitivity to others' psychological states (Franco & Butterworth, 1996). If such interaction were transcribed, it might indeed show an orderly sequential structure. Liszkowski and colleagues (in press) have shown that 12- (and 18-) month-olds will point at an object for which an experimenter is searching, thereby demonstrating an informative function of pointing. Again, these sorts of interactions are likely orderly. And although it would be good to see the videotapes first hand, I agree that these sorts of activities will likely show shared meaning. Perhaps experimental manipulations produce stronger situation contingencies to "pull this understanding out of infants" (i.e., to lead them to act in ways where criteria for shared understanding can be more easily satisfied) than the more natural situations I investigated. However, not all infants in these studies show gaze alteration or an "ability to point informatively," so there is no obvious reason to be suspicious of the data in Study Two. I should also note that Bates and colleagues (1975) argued that some infants become sufficiently accustomed to or as they put it "confident in" their mother's attention that they do not display the same degree of overt visual checking. Similarly, Bates et al. (1979) argued that visual checking might very well decrease as a function of age. But infants do show increases in gaze alternation in the second year (e.g., Morissette et al., 1995) so it is unclear what to make of Bates et al.'s position as it currently stands.

In sum, there is little evidence of shared meaning in pointing gestures, but there is evidence of shared meaning in conventional gestures and acts. Further although M deploys pointing gestures with frequency and in a manner that is well-sequenced with

the ongoing flow of activity, C for the most part does not. C, however, does engage in a multitude of behaviour at the first visit that has been theorized to catalyze the pointing gesture. However, we see little evidence of the mother or child linking this up with their ongoing activities. And we have to square all of this against the fact that it is commonly assumed that when parent and child participate in the joint activities requiring attention that were detailed above it is appropriate to claim that shared meaning exists. These are often described as episodes of intersubjectivity, and in particular, as episodes of secondary intersubjectivity by Trevarthen and colleagues (1979; Trevarthen & Hubley, 1978). Some theorists take the even stronger position that infants can only participate in such activities when they have some grasp of other persons as agents with a mind (Tomasello, 1995, 1999a). But when we look in detail at what parent and infant are up to, it is not at all clear that such rich views are tenable. In the first instance, there are few instances of anything resembling what we would standardly think of as two persons sharing meaning. In the second instance, when this infant is able to make her desires known, even by a pointing gesture, she does it without demonstrating any clear sense of another person as an agent with an attentional focus.

Now granted I am reporting on the younger end of the spectrum, but this was precisely the reason why I looked at the end of the first year of life. Babies are fairly competent at these forms of interaction in the second year, and I wanted to look at interaction when babies were first afforded the opportunity to direct another's attention. Despite the fact that I find mixed evidence that infants engage in joint activities requiring attention with a clear sense of "intersubjectivity," this is of course not to say that we cannot step into an observer frame of reference and correctly describe infants' attention,

intentions and desires and even their responses to the attention, intentions and desires of others, but this particular 9- to 12-month-old infant seems to show a quite tenuous grasp of the meaning of her actions to others.

CHAPTER 7 SUMMARY AND CONCLUSIONS

In this dissertation, I have argued that theories of joint attention are largely predicated on a causal-psychological-representational picture of mind that is embedded in our languages and which repeats itself inexorably, and self-confirmingly, in our explanations of meaning and mind. As a result, despite the fact that theorists may have richer or leaner explanations of the origins of the various joint activities that require attention, developmental theorists debate the timing of mind not its ostensibly Cartesian nature. Although the CPR view of mind may seem to some to be most apparent in Tomasello's cognitive theory of joint attention and tests thereof (Carpenter et al., 1998; Liszkowski et al., 2004; Tomasello, 1995, 1999a, 1999b; Tomasello et al., in press), it is equally apparent in biological theories (Butterworth, 1998b, 2003; Butterworth & Jarrett, 1991; Gopnik & Meltzoff, 1997; Meltzoff & Brooks, 2001; Meltzoff et al., 1999; Meltzoff & Moore, 1995; Trevarthen, 1979; Trevathen & Aitken, 2001; Trevarthen & Hubley, 1978). And although social theories may have a different starting point, as currently formulated they seem to have an identical destination (Barresi & Moore, 1996; Corkum & Moore, 1995, 1998; Hobson, 1993, 2002; Moore, 1996b, 1998; Moore & Corkum, 1994; Moore & D'Entremont, 2001) -- namely a CPR picture of meaning and mind.

My argument then is that most social cognitive theorists of human ontogenesis have put the cart before the horse and have reified and interiorized the cart in so doing. The horse in this case is the emergence of particular activities such as pointing or social referencing that are taken, in a roughly Cartesian manner, to *reveal* some insight or pre-existing one-to-one biological unfolding. However, these practices were here all along in both an ontogenetic and a phylogenetic sense as proto-language games. And typically

developing humans go on to acquire a capacity to understand the concepts that are tied up with such forms of interaction. However, this is not because they develop a capacity to experience particular "internal" mental states. Rather, they come to understand the multiplicity of circumstances in and through which particular sets of criteria can be utilized. Now, this is of course not to say that humans are mindless S-R automatons, but rather that the mind is defined in and though such activities and what we think of as mental is inseparable from activity. And for the same reason, the development of a capacity for reflection is not constitutive of an understanding of mind but is rather parasitic on such particular activities and the concepts with which they, for possessors of sufficiently complex language, are tied up. Thus, we should not take mental states to be, well, mental states. That is, we should not take them to be mental representations that cause activity and render it intelligible. Mind is a non-local and emergent property of particular forms of interaction. And many animals have minds in the way that they are commonly understood. But humans alone seem to understand that they have them.

I have suggested instead that whereas joint activities that require attention naturally emerge between conspecifics of a variety of species, the social cognitive differences between species, and indeed between a younger human infant and an older child, are deeply linguistic. I do not mean this in a Sapir-Whorf sense, but rather at a much more fundamental level that is tied up with a social naturalistic notion of form of life as both a natural selection of a propensity towards interest in conspecifics and a cultural selection of particular linguistic forms that are grounded on natural reactions. This is a process whereby humans replace or augment natural expressions of attention, desire and intention with mental state terms for attention, desire or intention. This human

ability, I contend, is the source of social cognitive differences across the primate line. Other animals attend to, want and intend particular states of affair, the difference is that we humans come to understand that we attend, desire and intend. And although biological, social and cognitive factors are causally and externally related to the child's mastery of such concepts, it is form of life, criteria and the circumstances in and through which criteria are expressed that ultimate define and delimit the mind that human children learn about.

7.1 Social Interaction in Human and Non-Human Primates

It is also true that the affective interpenetration of interaction for humans seems quite different from what is natural for other primate species. This could be a simple reflection of the fact that great ape infants are not primed with the same degree of social interest as their human counterparts. However, there may be something powerfully formative to the very context of human social interaction that facilitates social development (Carpendale & Lewis, 2004; Carpendale et al., in press; Hobson, 1993, 2004; Tomasello, 1999b; Tomasello et al, in press). This may go some of the distance towards accounting for the effects of enculturation on social cognition in the great apes to the extent that there are any such effects. And it may help to explain why the bonobo Kanzi has been argued to declaratively point on occasion even though the protodeclarative is often argued to be a uniquely human act (Savage-Rumbaugh et al., 1998). However, not only are particular forms of interaction required to understand the mind but a sufficiently complex language is also required. The fact that Kanzi has developed a lexicon tantamount to that of a two-year-old could explain why we might say

that he does not have some of the more complex, linguistically-saturated, mental state concepts such as belief. However, perhaps Kanzi has simpler concepts like attention, desire or intention (Savage-Rumbaugh et al., 1998).

Carpendale and colleagues (in press) have argued that Portmann's (1944/1990) notion of an "extrauterine year" may be important for understanding the social cognitive differences between human and non-human primates. Portmann argued that because human infants are born a year earlier than other comparable mammals, their period of protracted dependence must have clear survival value because it comes with such a high cost. A by-product of such a contingency is that the first year of human life takes place in the context of human social interaction rather than in the womb. And an interest in, for example, faces and also some capacity for engagement with others might be all that is required for infants to benefit from this period of "social gestation." This raises an interesting possibility. Namely, that the social cognitive differences between humans and other primates are a function of genetic and social differences, which is what I have argued. Therefore, the degree and extent to which a capacity for joint attention must be specified as a biological primitive need not be very great to explain why humans might predisposed towards interest in conspecifics in a manner that is more diminished in other primates.

7.2 Empirical Work on Joint Attention

In the addition to the review and analysis presented in this dissertation, I also described two empirical studies of the development of pointing. These were not intended as "critical tests" of the approach I take, but rather as putatively consistent with,

and to a certain extent a demonstration of, that approach. In fact, I assume a philosophy of science where critical tests are replaced in provenance by entire programmes of work, which can then be judged more or less adequate (Lakatos, 1970). As such, although these two empirical studies are consistent with what I have perhaps dubiously called a Wittgensteinian psychology, they cannot be critical tests of it. This approach also calls for an extensive longitudinal study that could link up children's joint activity with their eventual linguistic practice. What I did in Study One though was to account for the manner in which pointing emerges within shared activity. I did this because I take social shaping to be a way of talking about the emergence of the pointing gesture. Clearly the other hypotheses that I considered were at odds with such an account. In Study Two I instead investigated how infant and parent treated the pointing gesture and other cognate activities. I did this to give some sort of paraphrase to the term intersubjectivity.

In the first study, I reported data that were inconsistent with the particular social shaping hypothesis that has been attributed to Vygotsky (1978). I reported additional data that was inconsistent with both Butterworth's (1998b, 2003; Franco & Butterworth, 1996) more biological and Tomasello's (1995, 1999a) more cognitive hypotheses. Although the results of this study did not support any particular approach that I reviewed, I argued that they favour a more gradual, non-insight nor biologically predetermined account of social development. These data also not only point the way towards future work in this sort of a tradition but also seem to naturally lead to an analysis along the lines of Study Two.

In the second study, I used the methods of social pragmatics (Turnbull, 2003; Turnbull & Carpendale, 1999) to show that both parent and child did not for the most part

orient to either index finger "tipping" or "slip out" points in their prior or subsequent interactional turns. Thus, we have participant (as opposed to observer) evidence that Lock et al.'s notion of "pointless points" may be a propos. The perhaps surprising side of pointing gestures in Study Two was the fact that in the only unambiguous case of "a pointing gesture" both the mother and child participants did not treat it as an attempt to direct the mother's attention -- and in either case they did not treat it in a manner that seemed to warrant the attribution of shared meaning. However, there were other sequences of interaction where parent and child seemed to attribute roughly identical meaning to their activities and as such the results of this study were mixed with respect to the degree to which this young infant shares meaning with her mother.

As I see it, Liszkowski et al.'s (2004, in press) and similar studies have an interesting, and roughly continuous, interactional structure that is an improvement over the more discrete measures that are often used where a researcher simply counts up the number of points or the number of points with gaze alteration. Ironically, the ESCS used in Study One is one of these more discrete sorts of measures.[41] And in the arguably weaker situational contingencies of natural interaction, we only witnessed one sequence of interaction that contained a point and even then it may have been a "pointless point" because as mentioned neither mother nor child treated it in this manner. Although it does raise the question how many of the points coded in the ESCS were actually meaningful points from an interactional point of view, I chose to include a

[41] For example, a point is coded as declarative if it occurs when a toy is activated, whereas it is coded as imperative if the point occurs before or after the activation. This aspect of the coding does not take the structure of the interaction into account at all. Interestingly, the dyad described in Study Two featured an infant who was in fact coded as pointing (26 times imperatively, 7 times imperatively), including pointing with eye contact to the experimenter (once on a declarative point, 4 times on an imperative point) across all visits on the ESCS.

standard measure like the ESCS in Study One because although lacking an experimental design I wanted to be able to make comparisons back to the field.

The fact that experimental studies like Liszkowski and colleagues' may have the quasi-interactional structure that they do though does not seem to signify some concern for participant meaning, or a belief that psychological meaning is a negotiated processing involving some shared practice and interactional resources. Also, experimental studies, like the ESCS, can still be somewhat of a hodgepodge of discrete non-interactional notions and continuous roughly interactional ones. For example, Liszkowski et al. (in press, p. 6) define successful imperative pointing canonically *or* by "open hand, palm down, in the direction of the object." If none of the signs of imperative pointing are present, then given the appropriate social situation it is coded as an informative point (Liszkowski et al., in press) *or* a declarative point given (Liszkowski et al., 2004). Thus, these decisions are based on the structure of the interaction in one sense, but not in another. And in either case, the logic of these designs is to test for the presence of an ostensibly internal capacity in the child. Such work is of scientific value to the extent that it is a correct description of what is possible for a child of a given age. However, its scientific value is also constrained by the extent to which the results of such empirical work can be properly understood.

Although young infants are clearly capable of interacting in a way that is suggestive of shared meaning towards the end of the first year of life, what are we to make of statements such as "infants understand something about attention" at such an age (Liszkowski et al., 2004, p. 306) and a host of similar claims? I have argued that this way of putting it is problematic because non-language using agents do not

understand attention and our claims are based on our knowledge and not theirs. I take it that to understand something about attention is to have a "partial understanding" of this psychological concept. But for non-language using agents what could this possibly mean? Only the fact that infants of this age act in a way that warrants the ascription of particular psychological predicates. They do not understand more about attention than other agents who act in ways that justify the ascription of such mental predicates. This is not to say that such agents might not differ in other ways, but not in what they can be sensibly said to understand. And in contrast to this commonsense interpretation of what the infant is capable of at 12 months, infants can react to others' attention much earlier in life (e.g., Reddy, 2001, 2003). Thus, these mentalistic concepts are applicable earlier in life as well. But most in the field would not say that infants understand something about attention at 3 months of age. But on this commonsense logic, why do they not?

7.3 Full Circle to the Meaning of Joint Attention

In this dissertation I have laid out in considerable detail what I take to be entailed in an understanding of attention and, a fortiori, an understanding of joint attention, and I can now hear a chorus of social developmentalists exclaiming "well, we didn't mean that." Fair enough, but what exactly did they mean? This self-same chorus is also probably wondering where the psychology is in all of this. After all, is my job not supposed to be one of describing the causal psychological representational mechanisms involved in joint attention? Should a complete account of joint attention not contain such details? But details of what? What sort of psychological enterprise are we considering

here? Is the question "Where is the psychology in all of this?" itself not emblematic of a CPR picture of meaning and mind?

I have argued instead that the mind is defined in and through activity. And to that end I have analyzed the concept attention, considered how agents might understand such a concept and conducted two empirical studies of an act thought to show such an understanding. Other researchers have contributed to the literature by identifying a variety of plausible structural preconditions for understanding that particular aspect of agents that we call their mind. However, what is notably absent and indeed presupposed in these accounts as I see it is the monolithic role of linguistically referenced human social practice in the understanding of other minds. To the extent that the present work is of any use, it might be to remind the field of this simple fact.

REFERENCES

Anisfeld, M. (1991). Neonatal imitation. *Developmental Review, 11,* 60-97.

Anisfeld, M. (1996). Only tongue protrusion modelling is matched by neonates. *Developmental Review, 16,* 149-161.

Antaki, C. (2004). Reading minds or dealing with interactional implications. *Theory and Psychology*, *14*, 667-683.

Atkinson, J. M., & Heritage, J. (Eds.) (1984). *Structures of social action: Studies in conversation analysis.* Cambridge: Cambridge University Press.

Bakeman, R., & Adamson, L. (1984). Coordinating attention to people and objects in mother-infant and peer-infant interactions. *Child Development, 55,* 1278-1289.

Baker, G. P., Hacker, P. M. S. (1982). The grammar of psychology: Wittgenstein's bemerkungen uber die philosophie der psychologie. *Language and Communication*, *2*, 227-244.

Baker, G. P., Hacker, P. M. S. (1985). *Wittgenstein, rules, grammar and necessity* (an analytic commentary on the *Philosophical Investigations*; vol. 2). Cambridge: Blackwell.

Baldwin, D., & Moses, L. J. (1994). Early understanding of referential intent and attentional focus: Evidence from language and emotion. In C. Lewis & P. Mitchell (Eds.), *Origins of an understanding of mind* (pp. 133-156). Hillsdale, NJ: Erlbaum.

Baldwin, D., & Moses, L. J. (1996). The ontogeny of social information gathering. *Child Development*, *67*, 1915-1969.

Baron-Cohen, S. (1989). Perceptual role-taking and protodeclarative pointing in autism. *British Journal of Developmental Psychology*, *7*, 113-127.

Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind.* Cambridge: Cambridge University Press.

Baron-Cohen, S. (1999). The evolution of a theory of mind. IN M. C. Corballis & S. E. G. Lea (Eds.), *The descent of mind: Psychological perspectives on hominid evolution* (pp. 261-277). Oxford: Oxford University Press.

Bates, E. (1976). *Language and context.* New York: Academic Press.

Bates, E., Benigni, L., Bretherton, I., Camaioni, L., & Volterra, V. (1979). Cognition and communication from 9-13 months: Correlational findings. In E. Bates (Ed.), *The emergence of symbols: Cognition and communication in infancy.* New York: Academic Press.

Bates, E., Camaioni, L., & Volterra, V. (1975). The acquisition of performatives prior to speech. *Merrill-Palmer Quarterly*, *21*, 205-226.

Bates, E., O'Connell, B., & Shore, C. (1987). Language and communication in infancy. In J. Osofsky (Ed.), *Handbook of infant development* (pp. 149-203). New York: Wiley.

Bickhard, M. H. (1987). The social nature of the functional nature of language. In M. Hickmann (Ed.), *Social and functional approaches to language and thought* (pp. 39-65). New York: Academic Press.

Bickhard, M. H. (1993). Representational content in humans and machines. *Journal of Experimental and Theoretical Artificial Intelligence, 5,* 285-333.

Bickhard, M. H. (2001). Why children don't have to solve the frame problems: Cognitive representations are not encodings. *Developmental Review, 21,* 224-262.

Bickhard, M. H. (2004). Why believe in beliefs? *Behavioral and Brain Sciences*, *27*, 100-101.

Bigelow, A. E. (2003). The development of joint attention in blind infants. *Development and Psychopathology*, *15*, 259-275.

Bilmes, J. (1992). Referring to internal occurrences: A reply to Coulter. *Journal for the Theory of Social Behaviour*, *22*, 253-263.

Bretherton, I. (1991). Intentional communication and the development of an understanding of mind. In D. Frye & C. Moore (Eds.), *Children's theories of mind: Mental state and social understanding* (pp. 49-75). Hillsdale, NJ: Erlbaum.

Bretherton, I., McNew, S., & Beeghly-Smith, M. (1981). Early person knowledge as expressed in gestural and verbal communication: When do infants acquire a theory of mind? In M. E. Lamb & L. R. Sherrod (Eds.), *Infant social cognition: Empirical and theoretical considerations* (pp. 333-373). Hillsdale, NJ: Erlbaum.

Brooks, L. R. (1978). Non-analytic concept formation and memory for instances. In E. H. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 169-211). Hillsdale, NJ: Erlbaum.

Brooks, L. R. (1987). Decentralized control of categorization: The role of prior processing episodes. In U. Neisser (Ed.), *Concepts and conceptual development: Ecological and intellectual factors in categorization* (pp. 141-174). New York: Cambridge University Press.

Bruner, J. (1983). *Child's talk: Learning to use language.* New York: Norton.

Butterworth, G. (1998a). Origins of joint visual attention in infancy. In M. Carpenter, K. Nagell, & M. Tomasello, Social cognition, joint attention and communicative competence from 9 to 15 months of age. *Monographs of the Society For Research in Child Development, 63,* 144-166.

Butterworth, G. (1998b). What is special about pointing in babies? In F. Simion & G. Butterworth (Eds.), *The development of sensory, motor and cognitive capacities in early infancy: From perception to cognition* (pp. 29-40). Hove, UK: Psychology Press.

Butterworth, G. (2003). Pointing is the royal road to language for babies. In S. Kita (Ed.), *Pointing: Where language, culture, and cognition meet* (pp. 9-33). Mahwah, NJ: Erlbaum.

Butterworth, G., & Jarrett, N. (1991). What minds have in common is space: Spatial mechanisms serving joint visual attention in infancy. *British Journal of Developmental Psychology, 9,* 55-72.

Butterworth, G., & Morissette, P. (1996). Onset of pointing and the acquisition of language. *Journal of Reproductive and Infant Psychology, 14,* 219-231.

Button, G., Coulter, J., Lee, J. R. E., & Sharrock, W. (1995). *Computers, minds, and conduct.* Cambridge, UK: Polity Press.

Call, J., Hare, B., Carpenter, M., & Tomasello, M. (2004). 'Unwilling' versus 'unable': Chimpanzees' understanding of human intentional action. *Developmental Science, 7,* 488-498.

Camaioni, L. (2001). Early language. In G. Bremner & A. Fogel (Eds.), *Blackwell handbook of infant development* (pp. 404-426). Oxford: Blackwell.

Canfield, J. V. (1981). *Wittgenstein, language and world.* Amherst: University of Massachusetts Press.

Canfield, J. V. (1993). The living language: Wittgenstein and the empirical study of communication. *Language Sciences, 15,* 165-193.

Canfield, J. V. (1995). The rudiments of language. *Language and Communication, 15,* 195-211.

Canfield, J. V. (1999). Folk psychology versus philosophical anthropology. *Idealistic Studies, 29,* 153-172.

Carpendale, J. I. M., & Lewis, C. (2004). Constructing an understanding of mind: The development of children's social understanding within social interaction. *Behavioral and Brain Sciences*, *27*, 79-96.

Carpendale, J. I. M., Lewis, C., Müller, U., & Racine, T. P. (in press). Constructing perspectives in the social making of minds. *Interaction Studies.*

Carpendale, J. I. M., & Racine, T. P. (2005). *Mead and meaning: Implications of views of mean for developmental theories.* Paper presented at the 35th Annual Meeting of the Jean Piaget Society, Vancouver, BC.

Carpenter, M., Nagell, K. & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development, 63,* (Serial No. 255).

Carpenter, M., Pennington, B. F., & Rogers, S. J. (2002). Interrelations among social-cognitive skills in young children with autism. *Journal of Autism and Developmental Disorders*, *32*, 91-106.

Chapman, M. (1987). Inner processes and outward criteria: Wittgenstein's importance for psychology. In Chapman & R. A. Dixon (Eds.), *Meaning and the growth of understanding: Wittgenstein's significance for developmental psychology* (pp. 103-127). Berlin: Springer-Verlag.

Chapman, M. (1988). *Constructive evolution: The origins and development of Piaget's thought.* Cambridge: MA, Cambridge University Press.

Chomsky, N. (1995). *The minimalist program.* Cambridge, MA: MIT Press.

Corkum, V., & Moore, C. (1995). Development of joint visual attention in infants. In C. Moore & P. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 61-83). Hillsdale, NJ: Erlbaum.

Corkum, V., & Moore, C. (1998). The origins of joint visual attention in infants. *Developmental Psychology*, *34*, 28-38.

Coulter, J. (1991). Cognition: 'Cognition' in an ethnomethodological mode. In G. Button (Ed.), *Ethnomethodology and the human sciences* (pp. 177-195). Cambridge: Cambridge University Press.

Coulter, J. (1999). Discourse and mind. *Human Studies*, *22*, 163-181.

Coulter, J. (2004). What is "discursive psychology"? *Human Studies*, *27*, 335-340.

Darwin, C. (1964). *On the origins of species.* Cambridge, MA: Harvard University Press. (Original work published 1859)

Darwin, C. (1965). *The expression of the emotions in men and animals.* Chicago: University of Chicago Press. (Original work published 1904)

Dennett, D. (1987). *The intentional stance.* Cambridge, MA: MIT Press.

Dewey, J. (1958). *Experience and nature.* La Salle, IL: Open Court. (Original work published 1929)

D'Entremont, B. (2000). A perceptual-attentional explanation of gaze following in 3- and 6-month-olds. *Developmental Science, 3,* 302-311.

D'Entremont, B., Hains, S. M. J., & Muir, D. W. (1997). A demonstration of gaze following in 3- to 6-month-olds. *Infant Behaviour and Development, 20,* 569-572.

Fodor, J. (1983). *The modularity of mind.* Cambridge, MA: MIT Press.

Fogel, A. (1981). The ontogeny of gestural communication: The first six months. In R. E. Stark (Ed.), *Language behaviour in infancy and early childhood* (pp. 17-44). Amsterdam: Elsevier.

Fogel, A. (1993). *Developing through relationships: Origins of communication, self and culture.* Exeter, UK: Harvester Wheatsheaf Press.

Fogel, A., & Hannan, T. E. (1985). Manual actions of nine- to fifteen-week-old human infants during face-to-face interaction with their mothers. *Child Development, 56,* 1271-1279.

Fogel, A., & Thelen, E. (1987). Development of early expressive and communicative action: Reinterpreting the evidence from a dynamical systems perspective. *Developmental Psychology, 23,* 747-761.

Franco, F., & Butterworth, G. (1996). Pointing and social awareness: Declaring and requesting in the second year. *Journal of Child Language, 23,* 307-336.

Frye, D. (1991). The origins of intention in infancy. In D. Frye & C. Moore (Eds.), *Children's theories of mind: Mental states and social understanding* (pp. 15-38). Hillsdale, NJ: Erlbaum.

Gallese, V. (2005). From mirror neurons to the shared manifold hypothesis: A neurophysiological account of intersubjectivity. In S. Taylor Parker, J. Langer, & C. Milbrath, *Biology and knowledge revisited: From neurogenesis to psychogenesis* (pp. 179-204). Mahwah, NJ: Erlbaum.

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of min-reading. *Trends in Cognitive Science*, *12*, 493-501.

Garfinkel, H. (1967). *Studies in ethnomethodology*. Englewood Cliffs, NJ: Prentice-Hall.

Garfinkel, H. (1988). Evidence for locally produced, naturally accountable phenomena of order, logic, reason, meaning, method, etc. in and as of the essential quiddity of immortal ordinary society, (I of IV): An announcement of studies. *Sociological Theory*, *6*, 103-109.

German, T. P., & Leslie, A. M. (2001). Attending to and learning about mental states. In P. Mitchell & K. J. Riggs (Eds.), *Children's reasoning and the mind* (pp. 229-252). Hove, UK: Psychology Press.

Goldberg, B. (1991). Mechanism and meaning. In J. Hyman (Ed.), *Investigating Psychology: Sciences of the mind after Wittgenstein* (pp. 48-66). London: Routledge.

Goldin-Meadow, S., & Butcher, C. (2003). Pointing towards two-word speech in young children. In S. Kita (Ed.), *Pointing: Where language, culture, and cognition meet* (pp. 85-108). Mahwah, NJ: Erlbaum.

Goodhart, F., & Baron-Cohen, S. (1993). How many ways can the point be made? Evidence from children with and without autism. *First Language, 13*, 225-232.

Gopnik, A., & Meltzoff, A. N. (1997). *Words, thoughts, and theories.* Cambridge, MA: MIT Press.

Hacker, P. M. S. (1990). *Wittgenstein, meaning and mind* (an analytic commentary on the *Philosophical Investigations*; vol. 3). Cambridge: Blackwell.

Hacker, P. M. S. (1996). *Wittgenstein, mind and will* (an analytic commentary on the *Philosophical Investigations*; vol. 4). Cambridge: Blackwell.

Hacker, P. M. S. (1997). *Wittgenstein: On human nature.* London: Phoenix.

Hannan, T. E., & Fogel, A. (1986). A case assessment of "pointing" during the first three months of life. *Perceptual and Motor Skills*, *65*, 187-194.

Hark, M. ter (1990). *Beyond the outer and the inner: Wittgenstein's philosophy of psychology.* Dordrecht, The Netherlands: Kluwer.

Hobson, R. P. (1991). Against the theory "theory of mind." *British Journal of Developmental Psychology*, *9*, 33-51.

Hobson, R. P. (1994). Perceiving minds, conceiving minds. In C. Lewis & P. Mitchell (Eds.), *Origins of an understanding of mind* (pp. 71-93). Hillsdale, NJ: Erlbaum.

Hobson, R. P. (1993). *Autism and the development of mind.* Hove, UK: Erlbaum.

Hobson, R. P. (2002). *The cradle of thought.* London: Macmillan.

Jacobson, S. (1979). Matching behaviour in the young infant. *Child Development, 50,* 425-430.

Jayyusi, L. (1991). Values and moral judgement: Communicative praxis as a moral order. In G. Button (Ed.), *Ethnomethodology and the human sciences* (pp. 227-251). Cambridge: Cambridge University Press.

Jones, S. S. (1996). Imitation or exploration? Young infants' matching of adults' oral gestures. *Child Development, 67,* 1952-1969.

Kaye, K. (1982). *The mental and social life of babies: How parents create persons.* Chicago: University of Chicago Press.

Kita, S. (2003) (Ed.). *Pointing: Where language, cognition and culture meet.* Mahwah, NJ: Erlbaum.

Krause, M. A., Fouts, R. S. (1997). Chimpanzee (Pan troglodytes) pointing: Hand shapes, accuracy and the role of eye gaze. *Journal of Comparative Psychology, 111,* 330-336.

Kusch, M. (1997). The sociophilosophy of folk psychology. *Studies in the History and Philsophy of Science, 28,* 1-25.

Lakatos, I. (1970). Falsification and the methodology of scientific research programmes. In I. Lakatos and A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 231-278). Cambridge: Cambridge University Press.

Leavens, D. A. (2004). Review of *Pointing: Where Language, Culture and Cognition Meet,* S. Kita (Ed.); Lawrence Erlbaum Asssociates, 2003; ISBN 0-8058-4014-1. *Cognitive Systems Research, 5,* 157-165.

Leavens, D. A., Hopkins, W. D., (1999). The whole-hand point: The structure and function of pointing from a comparative perspective. *Journal of Comparative Psychology, 113,* 417-425.

Leavens, D. A., Bard, K. A., & Hopkins, W. D. (2005a, April). *The affective envelope of joint attention.* Paper presented at the 71st Biennial Convention of the Society for Research in Child Development, Atlanta, GA.

Leavens, D. A., Hopkins, W. D., & Thomas, R. K. (2004). Referential communications of chimpanzees (Pan troglodytes). *Journal of Comparative Psychology, 118,* 48-57.

Leavens, D. A., Russell, J. L., & Hopkins, W. D. (2005b). Intentionality measured in the persistence and elaboration of communication by chimpanzees (pan troglodytes). *Child Development, 76*, 291-306.

Leslie, A. M. (1987). Pretense and representation: The origins of "theory of mind." *Psychological Review, 94*, 412-426.

Leslie, A. M. (1994). ToMM, ToBy and agency: Core architecture and domain specificity. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 119-148). New York: Cambridge University Press.

Leung, E., & Rheingold, H. (1981). The development of pointing as a social gesture. *Developmental Psychology, 17*, 215-220.

Leudar, I., & Costall, A. (2004). On the persistence of the "problem of other minds" in psychology. *Theory and Psychology, 14*, 601-621.

Liszkowski, U., Carpenter, M., Henning, A., Striano, T., & Tomasello, M. (2004). Twelve-month-olds point to share attention and interest. *Developmental Science, 7*, 297-307.

Liszkowski, U., Carpenter, M., Striano, T., & Tomasello, M. (in press). Twelve- and 18-month-olds point to provide information for others. *Journal of Cognition and Development.*

Lock, A., Young, A., Service, V., & Chandler, P. (1990). Some observations on the origins of pointing. In V. Volterra & C. J. Erting (Eds.), *From gesture to language in hearing and deaf children* (pp. 43-55). Berlin: Springer-Verlag.

Malcolm, N. (1991). Wittgenstein: The relation of language to instinctive behaviour. In J. Hyman (Ed.), *Investigating psychology: Sciences of the mind after Wittgenstein* (pp. 27-47). London: Routledge.

Martin, J., & Sugarman, J. (1999). *The psychology of human possibility and constraint.* Albany, NY: State University of New York Press.

Martin, J., Sugarman, J., & Thompson, J. (2003). *Psychology and the question of agency.* Albany, NY: State University of New York Press.

Masataka, N. (2003). From index-finger extension to index-finger pointing: Ontogenesis of pointing in preverbal infants. In S. Kita (Ed.), *Pointing: Where language, culture, and cognition meet* (pp. 69-84). Mahwah, NJ: Erlbaum.

Mead, G. H. (1964). *The social psychology of George Herbert Mead: Selected papers* (Ed. A. Strauss). Chicago: University of Chicago Press.

Meltzoff, A. N. (1995). Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology*, *31*, 838-850.

Meltzoff, A. N., & Brooks, R. (2001). "Like me" as a building block for understanding other minds: Bodily acts, attention, and intention. In B. F. Malle, L. J. Moses, & D. A. Baldwin (Eds.), *Intentions and intentionality* (pp. 171-191). Cambridge, MA: MIT Press.

Meltzoff, A. N., Gopnik, A., & Repacholi, B. M. (1999). Toddlers' understanding of intentions, desires, and emotions: Explorations of the dark ages. In P. D. Zelazo, J. W. Astington, & D. R. Olson (Eds.), *Developing theories of intention: Social understanding and self-control* (pp. 17-41). Mahwah, NJ: Erlbaum.

Meltzoff, A. N., & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science, 198,* 75-78.

Meltzoff, A. N., & Moore, M. K. (1995). Infants' understanding of people and things: From body imitation to folk psychology. In J. L. Bermudez & A. J. Marcel (Eds.), *Body and the self* (pp. 43-69). Cambridge, MA: MIT Press.

Moore, C. (1996a). Evolution and the modularity of mindreading. *Cognitive Development*, *11*, 605-621.

Moore, C. (1996b). Theories of mind in infancy. *British Journal of Developmental Psychology, 14,* 19-40.

Moore, C. (1998). Social cognition in infancy. In M. Carpenter, K. Nagell, & M. Tomasello, Social cognition, joint attention and communicative competence from 9 to 15 months of age. *Monographs of the Society For Research in Child Development*, *63*, 167-174.

Moore, C. (1999). Gaze following and the control of attention. In P. Rochat (Ed.), *Early social cognition: Understanding others in the first months of life* (pp. 241-256). Mahwah, NJ: Erlbaum.

Moore, C. (2005, April). Discussion of 'Imitation, Identification and Self-other Awareness'. In R. P. Hopson (Chair), *Imitation, identification and self-other awareness.* Symposium conducted at the 71st Biennial Convention of the Society for Research in Child Development, Atlanta, GA.

Moore, C., & Corkum, V. (1994). Social understanding at the end of the first year of life. Developmental Review, 14, 349-372.

Moore, C., & D'Entremont, B. (2001). Developmental changes in pointing as a function of attentional focus. *Journal of Cognition and Development, 2,* 109-129.

Moore, C., & Dunham, P. J. (Eds.) (1995). *Joint attention: Its origins and role in development.* Hillsdale, NJ: Erlbaum.

Montgomery, D. E. (1997). Wittgenstein's private language argument and children's understanding of mind. *Developmental Review, 17,* 291-320.

Montgomery, D. E. (2002). Mental verbs and semantic development. *Journal of Cognition and Development, 3*, 357-384.

Morissette, P., Ricard, M., & Decaré, T. G. (1995). Joint visual attention and pointing in infancy: A longitudinal study of comprehension. *British Journal of Developmental Psychology, 13,* 163-175.

Müller, U., & Carpendale, J. I. M. (2004). From joint activity to joint attention: A relational approach to social development in infancy. In J. I. M. Carpendale & U. Müller (Eds.), *Social interaction and the development of knowledge* (pp. 215-238). Mahwah, NJ: Erlbaum.

Müller, U., & Runions, K. (2003). The understanding of self and other: James Mark Baldwin's theory. *Developmental Review*, *23*, 29-54.

Mundy, P., Delgado, C., Block, J., Venezia, M., Hogan, A., Seibert, J. (2003). *A manual for the abridged early social communication scales (ESCS).* Unpublished manuscript, University of Miami, Miami, FL.

Myowa-Yamakoshi, M., Tomonaga, M., Tanaka, M., & Matsuzawa, T. (2004). Imitation in neonatal chimpanzees (Pan troglodytes). *Developmental Science*, *7*, 437-442.

Perner, J. (1991). *Understanding the representational mind.* Cambridge, MA: MIT Press.

Piaget, J. (1952). *The origins of intelligence in children.* New York: International Universities Press. (Original work published 1936)

Piaget, J. (1954). *The construction of reality in the child.* New York: Ballentine. (Original work published 1937)

Povinelli, D. J., & Davis, D. R. (1994). Differences between chimpanzees (Pan troglodytes) and humans (Homo sapiens) in the resting state of the index finger. *Journal of Comparative Psychology*, *108*, 134-139.

Povinelli, D. J., & Eddy, T. J. (1996). What young chimpanzees know about seeing. *Monographs of the Society for Research in Child Development, 61,* (Serial No. 247).

Povinelli, D. J., Bering, J. M., & Giambrone, S. (2003). Chimpanzees' "pointing": Another error of the argument by analogy. In S. Kita (Ed.), *Pointing: Where language, culture, and cognition meet* (pp. 35-68). Mahwah, NJ: Erlbaum.

Proudfoot, D., & Copeland, B. J. (2002). Wittgenstein's deflationary account of reference. *Language and Communication*, *22*, 331-351.

Racine, T. P. (2002). Computation, meaning and artificial intelligence: Some old problems, some new models. *Canadian Artificial Intelligence, 50,* 8-19.

Racine, T. P. (2004). Wittgenstein's internalistic logic and children's theories of mind. In J. I. M. Carpendale & U. Müller (Eds.), *Social interaction and the development of knowledge.* Mahwah, NJ: Erlbaum.

Racine, T. P., & Carpendale, J. I. M. (in press). The embodiment of mental states. In W. F. Overton, U. Müller, & J. Newman (Eds.), *Body in mind, mind in body: Developmental perspectives on embodiment and consciousness.* Mahwah, NJ: Erlbaum.

Reddy, V. (2001). Mind knowledge in the first year: Understanding attention and intention. In G. Bremner & A. Fogel (Eds.), *Blackwell handbook of infant development* (pp. 241-264). Malden, MA: Blackwell.

Reddy, V. (2003). On being the object of attention: Implications for self-other consciousness. *Trends in Cognitive Sciences*, *7*, 397-402.

Reddy, V., & Morris, P. (2004). Participants don't need theories: Knowing minds in engagement. *Theory & Psychology*, *14*, 647-665.

Reimers, M., & Fogel, A. (1992). The evolutions of joint attention of objects between infants and mothers: Diversity and convergence. *Analise Psicologica, 1*, 81-89.

Russell, B. (1921). *The analysis of mind*. London: Allen & Unwin.

Savage-Rumbaugh, E. S., Shanker, S. G., & Talbot, J. T. (1998). *Apes, language, and the human mind.* New York: Oxford University Press.

Scaife, M., & Bruner, J. S. (1975). The capacity for joint visual attention in the infant. *Nature, 253,* 265-266.

Schaffer, H. R. (1984). *The child's entry into a social world*. New York: Academic Press.

Schegloff, E. A. (1984). On some gestures' relation to talk. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action: Studies in conversation analysis* (pp. 266-296). Cambridge: Cambridge University Press.

Searle, J. R. (1983). *Intentionality: An essay on the philosophy of mind*. Cambridge: Cambridge University Press.

Shanker, S. (1991). The enduring relevance of Wittgenstein's remarks concerning intentions. In J. Hyman (Ed.), *Investigating psychology: Sciences of the mind after Wittgenstein* (pp. 67-94). London: Routledge.

Shanker, S. (1998). *Wittgenstein's remarks on the foundations of AI.* New York: Routledge.

Shanker, S. (2004). The roots of mindblindness. *Theory & Psychology, 14*, 685-703.

Sharrock, W., & Coulter, J. (2004). ToM: A critical commentary. *Theory & Psychology, 14*, 579-600.

Shotter, J. (2004). Wittgensteinian developmental investigations. *Behaviour and Brain Sciences, 27*, 121-122.

Slaughter, V., & McConnell, D. (2003). Emergence of joint attention: Relationships between gaze following, social referencing, imitation and naming in infancy. *Journal of Genetic Psychology, 164*, 54-71.

Sidnell, J. (2003). An ethnographic consideration of rule-following. *Journal of the Royal Anthropological Institute, 9*, 429-445.

Sinha, C. (1999). Grounding, mapping and acts of meaning. In T. Janssen & G. Redeker (Eds.), *Cognitive linguistics: Foundations, scope and methodology* (pp. 223-255). New York: De Gruyter.

Tomasello, M. (1995). Joint attention as social cognition. In C. Moore & P. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 103-130). Hillsdale, NJ: Erlbaum.

Tomasello, M. (1999a). Having intentions, understanding intentions, and understanding communicative intentions. In P. D. Zelazo, J. W. Astington & D. R. Olson (Eds.), *Developing theories of intention: Social understanding and self-control* (pp. 63-75). Mahwah, NJ: Erlbaum.

Tomasello, M. (1999b). *The cultural origins of human cognition.* Cambridge, MA: Harvard University Press.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition.* Cambridge, MA: Harvard University Press.

Tomasello, M., Call, J., & Hare, B. (2003). Chimpanzees understand psychological states -- the question is which ones and to what extent. *Trends in Cognitive Sciences, 7*, 153-156.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (in press). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences.*

Trevarthen, C. (1979). Communication and cooperation in early infancy. A description of primary intersubjectivity. In M. Bullowa (Ed.), *Before speech: The beginning of human communication* (pp. 99-136). London: Cambridge University Press.

Trevarthen, C., & Hubley, P. (1978). Secondary subjectivity: Confidence, confiding, and acts of meaning in the first year. In A. Lock (Ed.), *Action, gesture and symbol: The emergence of language* (pp. 183-229). London: Academic Press.

Trevarthen, C., & Aitken, K. J. (2001). Infant intersubjectivity: Research, theory, and clinical applications. *Journal of Child Psychology and Psychiatry, 42,* 3-48.

Turnbull, W. (2003). *Language and action: Psychological models of talk.* New York: Psychology Press.

Turnbull, W., & Carpendale, J. I. M. (1999). A social pragmatic model of talk: Implications for research on children's social understanding. *Human Development, 42,* 328-355.

Turnbull, W., & Carpendale, J. I. M. (2001). Talk and the development of social understanding. *Early Education and Development, 12,* 455-477.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Cambridge: Harvard University Press.

Walden, T., & Ogan, T. (1988). The development of social referencing. *Child Development, 59,* 1230-1340.

Wellman, H. M., Phillips, A. T., Dunphy-Lelii, S., & LaLonde, N. (2004). Infant social attention predicts preschool social cognition. *Developmental Science, 7,* 283-288.

Werner, H., & Kaplan, B. (1963). *Symbol formation: An organismic-developmental approach to language and the expression of thought.* New York: Wiley.

Whittlesea, B. W. A. (1997). The representation of general and particular knowledge. In K. Lamberts & D. R. Shanks (Eds.), *Knowledge, concepts and categories* (pp. 335-370). Cambridge, MA: MIT Press.

Whittlesea, B. W. A. (2003). On the construction of behaviour and subjective experience: The production and evaluation of performance. In J. S. Bowers & C. J. Marsolek (Eds.), *Rethinking implicit memory* (pp. 239-260). New York: Oxford University Press.

Wilkins, D. (2003). Why pointing with the index finger is not a universal (in sociocultural and semiotic terms). In S. Kita (Ed.), *Pointing: Where language, culture, and cognition meet* (pp. 171-216). Mahwah, NJ: Erlbaum.

Wittgenstein, L. (1958). *Philosophical investigations* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Wittgenstein, L. (1969). *On certainty* (Eds. G. E. M. Anscombe and G. H. von Wright, trans D. Paul and G. E. M. Anscombe). New York: Harper & Row.

Wittgenstein, L. (1980a). *Remarks on the philosophy of psychology* (vol. 1, Eds. G. E. M. Anscombe and G. H. von Wright, trans G. E. M. Anscombe). Oxford: Blackwell.

Wittgenstein, L. (1980b). *Remarks on the philosophy of psychology* (vol. 2, Eds. G. E. M. Anscombe and H.Nyman, trans C. G. Luckhardt and M. A. E. Aue). Oxford: Blackwell.

Wittgenstein, L. (1981). *Zettel* (2nd ed., Eds. G. E. M. Anscombe and G. H. von Wright, trans G. E. M. Anscombe). Oxford: Blackwell.

Wittgenstein, L. (1992). *Last writings on the philosophy of psychology* (vol.2, Eds. G. H. von Wright and H. Nyman, trans C. G. Luckhardt and M. A. E. Aue). Oxford: Blackwell.

Wootton, A. J. (1990). Pointing and interaction initiation: The behaviour of young children with Down's syndrome when looking at books. *Journal of Child Language, 17,* 565-589.

Wootton, A. J. (1994). Object transfer, intersubjectivity and third position repair: Early developmental observations of one child. *Journal of Child Language, 21,* 543-564.

Wootton, A. J. (1997). *Interaction and the development of mind.* Cambridge, MA: Cambridge University Press.