INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI®

YOUNG CHILDREN'S DEVELOPING THEORY OF MIND: PERSON REFERENCE, PSYCHOLOGICAL UNDERSTANDING AND NARRATIVE SKILL

by

SARAH E. HENSELER

A dissertation submitted to the Graduate Faculty in Psychology in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2000

This manuscript has been read and accepted for the Graduate Faculty in Psychology in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

December 13, 1999 Kathleen Lebon
Date **Chair of Examining Committee**

December 13, 1999
Date **Executive Officer**

Professor David Bearison

Professor Joseph Glick

Professor Herbert Saltzstein

Professor Anna Stetsenko
Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

ABSTRACT

YOUNG CHILDREN'S DEVELOPING THEORY OF MIND:
PERSON REFERENCE, PSYCHOLOGICAL UNDERSTANDING
AND NARRATIVE SKILL

by

SARAH E. HENSELER

Advisor: Professor Katherine Nelson

Research on young children's "theory of mind" concerns the process by which children come to understand that people are thinkers, as well as actors. This study of the relationships among person reference, psychological understanding, and narrative skill in early childhood tests two premises of the theory-construction theory (Gopnik & Wellman, 1992): 1) understandings of "self" and of "other" as thinker develop simultaneously, and 2) language reflects, but does not play a constitutive role in that development.

Fifty children (26 boys; 24 girls) in three age groups (15 3.6 to 4-year-olds; 17 4 to 4.6; 18 4.6 to 5-year-olds) from urban and suburban preschools, paired with peers, played a game in which partners had different objects, affective states and beliefs. Twenty-four hours later each child recounted the game. Recountings were coded for frequency of "self," "other," and "us" reference, psychological content of each, and demonstrated narrative skill. Children also answered questions about experience of "self" and "other" child, retold "The Dragon's Tears" story, took the TELD-2, and answered false belief task questions.

Findings were: 1) lowest language scorers referred only to "self" in game recounting, 2) measures of "self" and "other" reference were significantly correlated on frequency, psychological content, and false belief, 3) language scores were negatively correlated with percentage of self reference, and positively correlated with percentage of "we" reference and ability to answer false belief questions, and 4) children's articulated understanding of mental states was task-dependent.

Findings point to a process of mutual construction of "self" and "other" as thinkers (Nelson, 1996) in which linguistic skill and context play critical roles. Dynamic systems theory (Thelen & Smith, 1991) provides a basis for interpretation of the finding that "self" understanding is slightly more advanced than "other," and that children's mental state use in story retelling does not correlate with performance on a false belief task. The development of "we" use in early childhood has not been studied previously and raises important issues for further research on person reference.

ACKNOWLEDGEMENTS

The value to a doctoral candidate of an inspiring and trustworthy mentor cannot be overestimated. In Professor Katherine Nelson, I have had both. From my first days of graduate study, I have listened and learned as Katherine reframed longstanding questions and with elegant clarity, drew together aspects of development that have too long been kept apart in the field. Her integrative theory will continue to inspire my thinking and yield new insights and interests for many years to come. Dissertations, though, are not completed on the strength of inspiration alone. It has been my good fortune to have in Katherine a mentor whose intellectual depth is matched by the depth of her commitment to her students. I am delighted to have this opportunity to thank Katherine for the consistent and wise guidance she has given me throughout my doctoral study, and especially as I have worked on this project. It is with pride and deep appreciation that I count myself one of her students.

To the faculty of the City University of New York developmental psychology program, I owe a special word of thanks. I am grateful to Professors David Bearison and Joseph Glick not only for their service as members of my dissertation committee, but also for the important roles they have played in shaping my thinking about human development. Professor Herbert Saltzstein, valued teacher, and Professor Anna Stetsenko, who it was my pleasure to meet in this process, graciously agreed to read and comment upon this study. For the thoughtful comments of all, I say "thank you."

Dr. Matthew Schall contributed his rare and valuable expertise as a psychometrician to this project. In the midst of a very busy schedule, he responded with promptness and patience to my questions, offering statistical advice that was

both substantive and relevant to the questions I was raising. His gifts as a teacher were graciously put to use on my behalf, and for that I am deeply grateful.

This study owes its existence to the directors, staff, and children of the Lehigh Valley Child Care Centers, the Weekday Preschool, and Old First Nursery School. In each place, I observed dedication to children and a commitment to furthering our understanding of their development. Directors, teachers and children patiently tolerated the disruption in their routine that research inevitably causes. Working with the children was a joy. I thank them for allowing me to enter their world and to learn from them.

I am pleased, as well, to have this chance to put my gratitude to my family in print. My father, Warren John Henseler, has instilled in me, through word and example, a deep respect for, and enjoyment of, the life of the mind. It is unlikely that I would have attempted this work if not for his influence. My mother, Barbara Williams Henseler, demonstrates a care for and confidence in me that is the foundation for whatever success I achieve in life. I am grateful beyond words to them both. I am also blessed with two sisters, Phyllis Henseler O'Connell and Lois Henseler Albright. Their pride and joy in each success of my life is a treasured gift and a tribute to the quality of our shared family life. To Lois, I owe a special debt of gratitude. The respect that she has earned among her colleagues for her work on behalf of children opened the door that led me to those whose voices are heard in this study. I am enormously proud of her work and grateful for her contribution to mine.

Finally, I want to thank my husband, Alan Gerard Crudden. In accepting the considerable cost and challenge of being newly wed to a dissertation writer, he

demonstrated the depth of commitment that marks his character. His provision of a beautiful and an especially happy home in which to write these pages is a gift beyond value. I am delighted to have this opportunity to thank him, and all of those who have so generously contributed to making this work possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

No single question in developmental psychology has received more attention in the last decade than this one: When and how do children "discover" the mind? Researchers of widely divergent theoretical perspectives have agreed to focus their empirical attention on the 3½ to 5-year-old's growing understanding of mental life. Their research has documented a striking change in psychological understanding that occurs, for most children studied, between the ages of 4 and 5.

In their fifth year most children begin to be able when asked, to articulate that "my" beliefs and "yours" may differ: "*I know* there are crayons in that raisin box [which has a raisin pictured on it], but when my friend comes in and sees it [before she opens the box] *she'll think* there are raisins in it." The child who can say that, recognizes that thoughts do not necessarily reflect reality, and has grasped the possibility of "false belief." That conceptual advance, philosophers (see, for example, Dennett, 1978) have claimed, is essential to an understanding of the nature of mental life. In the current parlance of cognitive developmental psychology, then, the child who passes the tests of false belief is said to "have" a "theory of mind." So, this field of research received its name.

The name, however, does more than label a domain of interest. It, in fact, makes a striking and controversial claim: The young child, in the manner of the scientist, constructs an implicit theory of the workings of the mind. The name most accurately represents the claim of a single group of theorists (the so-called "theory theorists") who describe the child as a data gatherer. Her theory construction is disinterested,

they say, and her theory, once constructed, is applicable across persons and situations (Gopnik, 1993; Gopnik & Wellman, 1992; Gopnik & Wellman, 1994; Wellman, 1990). Once the child "has" a "theory of mind", they say, he can as readily talk about "your" thoughts as "mine" and could, it follows, do so despite differences in discourse context. Because of its dominance in the field, this "theory theory" stance has provoked much discussion. Others in the field have criticized its description of the child as scientist (Harris, 1994; Nelson, 1996; Nelson, Plesa, & Henseler, 1998) and its underlying claims continue to be called into question.

The current study joins in that questioning. It poses a challenge to the "theory theorists" claims. Specifically, it questions whether in fact the child's "theory" is constructed in a disinterested manner, free from the "biases" of situation or person. To do that, two issues are addressed. The first is the assumption that children's ability to articulate an understanding of "the mind" is task-neutral. Is a "theory of mind" or more broadly, psychological understanding of action and thought, something that children simply either have or lack? Is it an understanding that is independent of social support and activity context? It will be of interest to find out whether the children who participated in this study demonstrate different degrees of psychological understanding depending upon the requirements of, and support provided within, distinct discourse contexts. Exploring the role of context in children's developing cognitive competence does not set this study apart from others, though. To the contrary that exploration, inspired by the work of Vygotsky, has been of longstanding interest in the field and its relevance to all aspects and domains of cognition has long been acknowledged.

The second issue addressed in this study is the one that has yet gone largely unexplored, and is distinctly important to the domain of psychological understanding. This study asks: Is it warranted to claim that children's "theory of mind" is person-neutral? Is there any evidence that articulating his own experience and thought poses a different level of cognitive challenge to the child from that posed when children talk about the other? A third question will bring a long overdue dimension to the literature on person use: What level of challenge, relative to self- and other-reference, does the task of talking about "us" pose to the child? This study is of "I," "s/he" and "we" use in early childhood.

It is an exploratory and correlational study. The data were gathered in four discourse contexts: 1) the child's recounting of "yesterday's" game played with a partner, 2) the answering of standardized questions about the experiences of self and other in that game, 3) the retelling of a story read to the children by the investigator, and 4) responding to the questions posed in the standard "theory of mind" deceptive box task. The data will provide the opportunity to look at children's ability to talk about actions and psychological states both across persons within the game-recounting task, and across situations. Inclusion of the standard false belief (deceptive box) task, widely used in the field, makes it possible to look more specifically for correlations among context, person use, and children's false belief understanding. A pattern of person use across development in early childhood will be suggested.

RESEARCH BACKGROUND

The development of the understanding of "us"

The study of the use of "we" in the discourse of early childhood is unprecedented in the literature, and even references to the use of the first person plural in the toddler years are scarce. Most explicit and detailed in its focus on the transition from "I/my" to "we" is Nelson's description of the shift that occurs in the pronoun use recorded in the bedtime monologues of a precocious 23-month-old named Emily.

> ...[Between 23 and 24 months] 'we' emerges in [Emily's] discourse [in] contexts that were previously occupied by singular reference... There is some indication that 'we' is used first in normative or group contexts ['my have get my checkup, so we take my jamas off' and later, 'So after 'we' were finished with the store, we went...'] ... and by 32 months 'I' and 'we' appear to be used in different discourse contexts." (Nelson, 1989, p. 288-289).

Attentive to the social context in which this development takes place, Nelson notes that as Emily begins this transition, Emily's father "used 'we' as an alternative to 'you' quite frequently..." (Nelson, p. 288).

A study by Laks, Beckwith, and Cohen focused on the relationship between mothers' pronoun use in verbal interactions with their two-year-old children, and children's later intelligence test scores: "[Though all pronouns were considered in this study, it was only] mothers' use of the personal pronoun 'we' [that] was significantly related to their children's performance on the Stanford-Binet at age 5 and the Wechsler Intelligence Scale for Children at age 8" (Laks, Beckwith, and Cohen, 1990, p. 25). Mother's use of "we" was not found in the study to be related to social class

or education, nor to the amount and sophistication of the child's own talk. These two rare and theoretically divergent references to the use of "we" in an earlier period in development from the one discussed here nevertheless pique one's curiosity about the role of the first person plural pronoun in child development and suggest the importance of further research.

The development of self and other understanding

Interest in the relationship between self- and other-understanding, on the other hand, has a long history in the field. At the end of the nineteenth century, William James dared to tackle the challenge of defining the "self" in such a way that it could begin to be explored empirically. By distinguishing between the "me" and the "I", he opened the door to the exploration of the self-concept, of children's developing understanding of "me." James wrote

> I shall...treat....A) the self as known, or *me* ... *In its widest possible sense...a man's Me is the sum total of that he CAN call his*...[The *me* includes] 'the material me,' my body and possessions... 'the social me', the recognition I get from others... and 'the spiritual me,' the entire collection of my states of consciousness...and B) the self as knower, or the *I* ...It is that which at any given moment *is* conscious, whereas the Me is only one of the things which it is conscious *of* (James, 1985, p. 43-63).

Asking children to talk about "me" proved to be a fruitful beginning exploration of children's developing self-understandings and through it psychologists (Damon & Hart, 1988; Harter, 1996) discovered that distinct patterns of change were evident in children's self-descriptions across childhood. Physical descriptions and actions

dominated accounts in early childhood: "I have brown hair and I go to preschool... and I can lift this chair, watch me!"; psychological generalizations ("I'm pretty popular... because I'm nice and helpful") began to emerge by age seven or eight (Harter, 1996, p. 208). Though the knottier questions of the "I" remained difficult to ask and answer empirically, the data from interviews about "me" demonstrated the fruitfulness of asking children to talk about the self. This study shares the assumption that while children's talk does not tell us everything we want to know about the construction of children's self-understanding, it is a most valuable resource for looking for patterns of change in the articulation of that understanding across childhood.

There are those who have dared to do more, though, than describe the change in self-description across childhood. They have theorized about the construction of the self, and have reflected upon possible answers to some fundamental and challenging questions. Are human beings born "selves" or are they "made selves" through interaction with others? And of special interest here is the question: Which comes first in development, knowledge of self (as thinker, for example) or knowledge of the other? Or, do both emerge simultaneously, out of the interaction between them? What follows is a summary of the different positions currently taken by researchers and theorists in the field.

The priority of self understanding

For some theorists, psychological understanding is constructed by the individual who builds it upon the foundation of the direct experience of the self. Though Jean Piaget does not use "self" language, his concept of perspective taking describes an "I" who constructs the child's "reality" in progressively complex ways. For Piaget, the individual "I" is at first aware only of her own point of view, and only much later, comes to recognize that other points of view are possible. The child, thus, believes that what she sees or knows is simply "the way it is." That is the "relative nondifferentiation between the self and the external world" that Piaget called "realism" (Chapman, 1988, p. 48). She is "centered" in the self; her perspective is undifferentiated and not understood to be "a perspective" at all. Piaget describes that state of "egocentrism" in his 1923 book, The Language and Thought of the Child. The child's speech, like his intellectual life, is initially "egocentric" not because he is indifferent to others but "...because he does not attempt to place himself at the point of view of his hearer" (p. 32). This "primacy of his own point of view," Piaget says, is "perfectly natural" since one's own is "the only point of view that is possible at first" (p. 161).

The child does not move beyond egocentrism, according to Piaget, until about the age of 7 or 8 when she is able to coordinate her thought and action with peers. Children, as they engage in arguing their distinct points of view, in justifying their positions and negotiating with one another to arrive at solutions, recognize that different points of view are possible and construct the ability to differentiate and

coordinate their perspectives with those of others. For Piaget, that awareness of the other as a distinct perceiver and thinker does not begin to become possible until the child is able to engage in such exchanges with peers. The "I" has awareness first solely of the point of view of the self, and only later of others. Though Piaget has repeatedly been criticized for overlooking the perspective-taking ability of the young child (in communication, for example, Shatz & Gelman, 1973; Menig-Peterson, 1975; Fivush & Hamond, 1990), and for defining perspective-taking ability in such a way that it is, by definition, a late achievement (see Flavell's 1988 distinction between "Level I" and "Level II" perspective taking), his notion that development occurs as egocentrism diminishes and perspective-taking ability increases has been accepted by many, and has become the foundation of influential theories of moral and interpersonal development (e.g., Kohlberg, 1969; Selman, 1980).

Flavell (1985), in his textbook on cognitive development, made explicit, not only his commitment to Piaget's basic notion of egocentrism, but also his shared assumption that we have privileged, and prior, access to our own mental states:

> ... I believe we are "at risk" (almost in the medical sense) for egocentric thinking all our lives, just as we are for certain logical errors. The reason is... we experience our own points of view more or less drectly, whereas we must always attain the other person's in more indirect manners. Our own points of view are more cognitively "available" to us than another person's [Tversky & Kahneman, 1973]...Furthermore, we are usually unable to turn our own viewpoints off completely when trying to infer another's. Our own perspectives produce clear signals that are much louder to us than the other's... (Flavell, 1992, p. 108).

Flavell's statement of the Cartesian view that we have privileged access to our own mental states has been hotly debated among those interested in children's developing psychological understanding. Those who hold to that view have, however, had a

strong voice in the "theory of mind" debate. Among them are Harris (1991) and Johnson (1988) who have each put forth a "simulation" (information processing) theory of children's understanding of the mind.

> ...Mental simulation depends on the capacity to engage in two successive steps: (1) to imagine having a particular desire or belief, and (2) to imagine the actions, thoughts, or emotions that would ensue if one were to have those desires or beliefs. The products of such a simulation can then be attributed to other people who do have the simulated desires or beliefs (Harris, 1991, p. 283).

Harris' theory rests upon the assertion that children's ability to run simulations (i.e., to use their own minds as a working model of others') rests upon the foundation of two 'default settings': 1) states of reality, and 2) the Intentional states of the self toward that Reality (Harris, 1991, p. 290). The work of the imagination, as Harris puts it, is that of overriding these default settings in order to simulate alternatives. The child might need, for example, to imagine past or future states of the self and would thus need to override one default setting (i.e., overriding the current reality, but maintaining her own intentional state toward it). She might also want to imagine the intentional state of another person toward a current reality (i.e., "She wants that cookie.") In this case, too, the child needs to override only one default setting, since she can maintain the current reality (the real, currently available cookie), but she needs to override her own intentional state toward it in order to simulate the other's. A more difficult situation is that set up in the false belief task in "theory of mind" research. That scenario requires that the child put aside both current reality ("There are really crayons in the box") and his own intentional state toward that reality ("I know that there are crayons in the box"). The information processing demands of such situations are

much more challenging, Harris says. That added challenge explains the relatively late ability of the child to pass the false belief task, correctly stating that "My friend will think there are raisins in the box."

Regardless of level of difficulty, though, simulation theorists claim that the process of coming to know the other's mental states (as well as one's own past or future or pretend states) is built upon the notion that the child knows her own mental states in a direct, unmediated way. Her understanding of the other depends upon her knowledge of herself and upon her ability to use that knowledge to understand what the other thinks or knows or believes. "I understand, at some point," the child might say, if he could, "that others are thinking people, and that they act on the basis of those thoughts in certain principled ways, because I myself am a thinking person."

The 'theory theory'

The most outspoken opponents of simulation theories are the ones whose assumptions are challenged in this study, those who hold to the 'theory theory' of mind (Gopnik & Wellman, 1992; Gopnik, 1993; Wellman, 1990; Bartsch & Wellman, 1995). They claim instead, with others (Murphy & Medin, 1985; Carey, 1985; Spelke, 1991; Hirschfeld & Gelman, 1994) that cognition proceeds as children build theories about distinct domains of human knowledge (e.g., biology, physical realities, etc.). Understanding of the mind, "theory theorists" say, is not a matter of the child's using her own mental experience as a model to run simulations, observing the output of her own mind, given certain inputs, and then applying the results to others. Rather,

understanding the mind requires theoretical constructs, i.e., "abstract entities [in this case, desires and beliefs] postulated to provide a separate causal-explanatory level of analysis to account for evidential phenomena [in this case, behavior]" (Gopnik & Wellman, 1992).

That understanding has, Wellman says, the three essential characteristics of theoretical knowledge: 1) Psychological knowledge is coherent (i.e., "theoretical terms get their meaning through their interconnections with other terms in the theory"), 2) It rests upon ontological distinctions ("[Theories] carve phenomena into different kinds of entities and processes"), and 3) It provides a causal-explanatory framework to "account for... and make predictable phenomena in its domain" (Wellman, 1990, p. 6-7).

In standing for the 'theory theory,' its proponents oppose, not only the simulation view of children's psychological understanding, but also its underlying philosophical assumption of prior and unmediated knowledge of one's own mind. Gopnik (1993) writes,

> ...[though] we as adults believe that first-person knowledge of our psychological states comes directly from experience and third-person knowledge involves inference, developmental evidence suggests otherwise. Many 3-year-olds are consistently wrong in reporting some of their own immediately past psychological states and show similar difficulties in reporting others'... (Gopnik, 1993, p. 1).

Gopnik relies here on the paradigmatic false belief tasks of the "theory of mind" research, one of which (the so-called "deceptive box" task, that here will be a raisin box) is performed as follows: The researcher shows the child a box with a picture on it. In this study the box has raisins pictured on it. The researcher asks the child, "What

do you think is in this box?" The child answers, "raisins." Then the researcher opens the box and shows the child its contents; here those contents are crayons. The researcher asks the child "What's in the box?" and the child says, "crayons!" The researcher then closes the box and asks the child, "Before I opened the box, when it was closed just like this, what did ***you think*** was in the box?" The correct answer, given by most 5-year-olds is "raisins;" the common incorrect answer is "crayons." The second question, of interest in this study, is then posed: "If your teacher [or any other person] came in right now and saw the box closed just like this, what would ***s/he think*** was in the box?" Gopnik's claim then is that there is no significant difference between children's ability to answer the "self think" question and their ability to answer the question about what the "other" would think.

Though Gopnik's claim has been contradicted in other studies, she has remained insistent that children's reports of the prior knowledge states of the self are as equally likely to be wrong as are their reports of the knowledge states of others. That "fact" indicates, she says, that children's understanding of their own mind is no more direct, and therefore, no more advanced than is their understanding of the mind of others.

Bartsch & Wellman (1995) believing that word use reveals children's underlying conceptual understanding of mind, undertook to analyze records of children's developing word use in naturalistic settings. Using data available through the Child Language Data Exchange System (CHILDES), Bartsch and Wellman set out to discover, among other things, whether children's first uses of mental state terms were used to refer to the self, to the other, or equally and simultaneously, to both self and other. Their hypothesis was that since mental state terms are theoretical constructs,

once constructed by the child, they would, if they were "genuine references" to thinking, be applied at about the same time in development, to self and other. Bartsch and Wellman claim that their hypothesis is supported by the data.

Several reasons for skepticism, however, exist. For one, the authors declare that references to the mental states of the self clearly predominate for all seven subjects (whose ages at first "genuine reference" to mental states range from 2.4 to 3.6). The possible explanations are, they say, just two: 1) children simply prefer to talk about themselves, or b) they at first conceive of only their own beliefs (Bartsch & Wellman, 1995, p. 62). One could wonder, though, what that "simple preference" indicates and whether, in fact, there might be a relationship between their "preference" for (which might also be conceived as "interest" or "attentiveness to") themselves and their cognitive development. Perhaps children learn mental state words as others talk to them about their own minds (Nelson, 1996). Bartsch and Wellman do not suggest such a possibility or any others. They do, though, go on to explicitly dismiss the "self first" position by saying

> ...It is telling...that reference to others' belief follows so quickly behind the first references to belief ...The average delay between referring to one's own beliefs and thoughts and referring to someone else's is one month (Bartsch & Wellman, 1995, p. 62-63).

A look at the data, however, reveals a good deal of variation (*sd* = 2.75 months) in the time lag between self and other reference, with two children's other references lagging five months behind their self reference, and another two in fact referring to other before self. Given the small sample, possibility of sampling error in the data recorded and transcribed, the variation within it, and the degree of individual differences in

language acquisition in general, a claim of nearly simultaneous use of self and other reference does not seem warranted. Considerably more work needs to be done if this question about the relative priority of self or other in mental state understanding is to be answered.

For "theory theorists," though, these pieces of evidence lead to the conclusion that children do not think about "my mind" and "your mind," but rather, construct a theory of mind that is person-neutral. The child posits psychological states, first of desire, and then, between four and five years, of knowledge and belief, as causal entities that make possible the explanation and prediction of behavior ("She'll think there are raisins in the box, so she might want to eat one.") The child's theory is not socially derived. It

> ...[the'theory theorists' claim,] develops individually under its own creative power, responding to encounters in the world with new constructions derived somehow from universal principles (Nelson, 1996, p. 302)

The notion, however, that the simultaneous demonstration of self and other knowledge necessitates a 'theory theory' of mind is a mistaken one.

The representational change theory

Alternatives exist. Leslie (1987), Perner (1991) and Wimmer (Wimmer & Hartl, 1991), and Karmiloff-Smith (1992) are among those who advocate the position that the ability to pass the false belief task at about age 4 is a result of a newfound understanding of the mind as representational. Perner in his 1991 book entitled Understanding the Representational Mind describes three levels of representation: 1)

primary, in which the infant has only her own direct experience of reality, 2) secondary, in which the child can re-represent reality, considering past or future states of the self, for example, in addition to the present, and 3) metarepresentation, at which level the child knows that her representations are, in fact, representations. At about four, Perner claims, children can understand that other's thoughts are re-presentations and thus, can be false beliefs. Children's understanding of representation is equally applicable to self and other.

More general processing changes

Several current theories of cognitive development point to changes in children's ability to process information in general, as the underlying cause of their difficulties in many tasks, including false belief. A number of theorists point to changes in children's executive functioning skills (Leslie & Roth, 1993; Russell, 1996; Frye, Zelazo, & Palfai, 1995) as responsible for the dramatic changes we see in preschool children's behavior along many dimensions. Robbie Case's neo-Piagetian theory (1991) makes a case for the presence of central conceptual structures that apply to children's thinking across domains and variations in social and cultural context. Specifically, limits in working memory (i.e., how many units, systems of representation, can be processed and coordinated at once) play a central role. Case (1985) argues then that children's success on the false belief task comes when they are able to include two mental representations to one another in an integrating scheme.

Carlson, Moses, and Hix point specifically to the centrality of "inhibitory control" which they define as "the ability to suppress potentially interfering thought processes or actions" (Carlson, Moses, & Hix, 1998, p. 672) in a child's ability to correctly respond to theory of mind tasks. The salience of the actual state of affairs (in the case of this study, seeing crayons in the box) may make it difficult for children in this age range to inhibit referring to it, even if they have some understanding that false belief is possible. Their series of experiments designed to test that theory led them to the conclusion that "...an inhibitory problem most likely interacts with conceptual difficulties in the theory of mind domain." All of these theories are asocial. They relate to the functioning of brain and mind and do not give any, or much, explanatory power to the social world in which the child develops.

Priority of "other" understanding

Simultaneity of self- and other understanding is, however, not the only alternative to a "self first" position on the understanding of people as thinkers as well as doers. A second possibility is that the understanding of the other may precede the understanding of self as a thinker whose mental states affect his action. It could be that infants and young children do not introspect, nor build personal-neutral theories, but rather that they attend most closely to, and build their understanding of persons upon their knowledge of, the important others who sustain their survival and structure their experience of the world. Tomasello (1993), for example, writes:

> ... I do not see how the major changes in the concept of person and self could be determined by a child observing his own thoughts and behavior Cartesian style; I see no mechanism for development there, no reason for the child's concept to change over time. On the other hand, people are always doing things that confound young children, and it thus seems reasonable to suppose that intentional agents with perceptions and intentions, mental agents with thoughts and beliefs, and reflective agents with reflective thoughts and beliefs are concepts constructed by children *to explain the behavior of others.* Only after the child has constructed one of these concepts of person does his simulation of the point of view of others allow him to assimilate himself to that concept (Tomasello, 1993, p. 179).

This idea is an intriguing one, in which the others in the child's world serve as a kind of figure that stands out against the ground of the child's own, unexamined experience of himself. Infants and young children might well, in the service of enhancing their ability to participate in events that are directed by others, focus their attention on the important others in their lives. Person understanding might be constructed as infants and children seek to make sense of other's actions. The possibility remains open that the data collected in the proposed study will suggest that, on some dimensions, other-knowledge is more accessible to the young child, or more complex, than self-knowledge is.

The mutual, social construction of self and other understanding

Most theorists, though, who might be said to posit the priority of the other over the self, in fact claim that the origin of the child's understanding of both self and other can be found in some form of interaction between the two. Tomasello himself writes, "my inclination is to follow Vygotsky's (1978) lead in positing the primacy of the interpersonal over the intrapersonal" (Tomasello, 1993, p. 179). Not unlike the

psychoanalytic and attachment theorists who talk about the self developing out of "working models" of the relationship between self and other (Bowlby, 1980; Bretherton, 1985), other theorists of self construction refer to the social, interactive construction of the child's understanding of both self and other.

James Mark Baldwin, with his emphasis on the role of imitation in the construction of the self, is one who might at first be thought to posit the construction of the self out of the other. Through imitation, "something of the other is taken over by the self" (Baldwin, 1906). That incorporation of certain aspects of the other is not automatic or rote, however. Imitation, Baldwin says, is the child's first act of volition, and so can clearly be said to be interpersonal, an act carried out between self and other. "Both ego and alter are essentially social; each is a socius, and each is an imitative creation" (Baldwin, 1906, page 9).

Werner and Kaplan (Werner & Kaplan, 1963) describe a child who, out of the "primordial sharing experience" with the mother, develops to the degree that the child moves out of the state of undifferentiation to one of increasing differentiation (of self and other) and hierarchic integration. Though Piaget's construction is not social, but individual, one could describe his theory of perspective taking as involving that kind of mutual construction of self and other out of a prior state of undifferentiation. He writes,

> ... As soon as the schema of substantial and permanent objects is acquired, and especially at the level of intuitive intelligence, persons become other 'egos' at the same time as the ego itself is being constituted and becoming a person (Piaget, 1962 , p. 207).

The mutuality of self and other construction, while posited by theorists of various perspectives, is most often claimed by those who say with Baldwin that person knowledge is social in origin.

Language and the social construction of self and other understanding

Like Baldwin, George Herbert Mead offered a view of the self and of mind itself that is radically social.

> Our contention is that mind can never find expression, and could never have come into existence at all, except in terms of a social environment; that an organized set or pattern of social relations and interactions (especially those of communication by means of gestures functioning as significant symbols and thus creating a universe of discourse)...(Mead, 1934, p. 223).

"Self-consciousness," he goes on to say, "involves the individual's becoming an object to himself by taking the attitude of other individuals toward himself within an organized setting of social relationships..." (Mead, 1934, p. 225). As a developmental theorist, though, Mead was not content to stop there. He considered the process through which self-construction takes place. What is the means, Mead went on to wonder, by which the individual becomes able to objectify herself and to think of herself in the same way that others think of her? The answer, he said, is language. Unlike Piaget, for example, who believed that even language in a sense, belonged to the self who first used it egocentrically and only later began to communicate with others (Piaget, 1955), Mead believed that the self does not exist apart from language. Language is, he said, fundamental to selfhood. It is

> ... the means of social interaction that allows the individual to take the perspective of others, to perceive the self as an object in its own right, and to differentiate between self and others (Cicchetti & Beeghly, 1990, p. 2).

Language allows the child to think of "self" and "other" at all.

Vygotsky's thought about the relationship between others and the self is strikingly consonant with Mead's:

> ... we could say that humans' psychological nature represents the aggregate of internalized social relations that have become functions for the individual and forms of his/her structure...(Vygotsky, 1981, p. 164).

"We are aware of ourselves," he said, "for we are aware of others...[This is] because in relation to ourselves we are in the same [position] as others are to us." (Vygotsky, 1979, p. 29-30). Vygotsky's position was not limited, however, to consciousness of self; consciousness itself, he said, "is built from the outside through relations with others" (Kozulin, 1986, p. xxiv).

> Each function in the child's cultural development appears twice: first, on the social level and later on the individual level; first, between people(interpsychological) and then inside the child (intrapsychological) (Vygotsky, 1978, p. 57).

Central to Vygotsky's theory is the notion that essential to the construction of the individual's consciousness are the interpersonal relations and the psychological tools provided in the social world. Foremost among those tools, for Vygotsky, is human language. He writes,

> ...This [is an]... indisputable fact of great importance: Thought development is determined by language, i.e., by the linguistic tools of thought and by the sociocultural experience of the child...The child's intellectual growth is contingent on his mastering the social means of thought, that is, language (Vygotsky, 1986, p. 94).

The understanding, then, of the self as a thinking person, Vygotsky would say is, like all other intellectual development, reliant upon the child's linguistic interaction with others. In contrast to Piaget, who claims that knowledge of self and other as distinct perspective takers is constructed out of conflict with peers, Vygotsky assigns the greatest role in the child's knowledge construction to the adults in his world who are able, from their position of more expert knowledge and greater linguistic sophistication, to guide the child to more complex levels of thought.

The collaborative construction of self and other through language

The role of language as mediator of human cognition has been further developed by Nelson (1996) in her book, Language in Cognitive Development: The Emergence of the Mediated Mind. Unlike Piaget who claims that it is the individual who constructs knowledge and also in distinction from Vygotsky who says that it is the social world in the form of language and more expert others, that plays the active, constructive role, Nelson asserts that the individual and the social world together, in the context of linguistically-mediated activity, construct the child's "social-cultural mind" (Nelson, 1996, p. 21) and thus, the understanding of self and other. Several aspects of Nelson's theory will be of particular relevance to the study proposed here.

First, Nelson claims that the "the primary cognitive task of the human child is to make sense of his or her situated place in the world in order to take a skillful part in its activities" (1996, p. 5). The importance of the specific activity contexts of a child's life cannot be overestimated, since human cognition itself, Nelson says, is built upon

the foundation of the child's mental representations (called MERS, mental event representations, or models by Nelson) of those activities and her own action within them. The centrality of MERS to cognition in early childhood has been demonstrated in numbers of studies of very young children's event knowledge (Nelson, 1986; Hudson, 1986). Building upon this theoretical foundation, the proposed study will not ask children to observe and reflect on the experience of puppets on a stage or characters on a page. Here, children will be invited to participate in, and make sense of, the actions and thoughts of self and other in the context of activity. This method reflects most accurately, it is claimed, the way in which children, in the course of their everyday lives, construct psychological understanding "in action."

In the context of shared activity, young children build representations of specific activities (mealtime or bathtime routines) and of people as agents with goals. They are able to remember and report the sequence of activities that constitute mealtime or bathtime routines and the roles people play within them ("Mommy makes a sandwich and I eat it"). But, Nelson says, the understanding of human beings (both self and other) as *thinking* people requires more than shared activity. It requires language.

> ...Language alone enables the child to gain insight into the knowledge/belief states of others...We only know what other people think, if it is different from what we think, because they tell us about it... (Nelson, 1996, p. 321).

Nelson does not stand, though, with those who claim that the process of "being told about it" leads directly to the child's sharing the adult meaning of mental state words. Nor does she assume, as others do, that using words like "think" or "know" is a simple reflection of already-existing concepts of mental states. Nelson outlines a

much more complex process by which the child moves from an action-focus, a notion of people as actors, to a consciousness-focus with its understanding of people as thinking actors. She describes a developmental process in which children, at first immersed in action, begin between the ages of 2 and 4 to use language as part of, and in conjunction with, activities. It is not yet a means of conveying information or knowledge from one person to another. It is only later, at about the time they are able to pass the standard false belief tests, that children begin to use language as a mediator of knowledge about the world.

After several years of hearing and using language in context, children begin to be able to construct new or alternate models of the world based on language alone. The claim is that it is only after extensive experience with language in context that the child comes to be able to use linguistic input to form new internal representations of reality. Only then can children use verbal representations as alternates to current, experienced reality ("I know it's crayons, but she'll think it's raisins").

The construction, then, of an understanding of self and other as thinking people, whose thoughts are not visible, shared actions, nor simply linguistic accompaniments to action (since they often contradict the state of the world) requires a good deal of linguistic sophistication. Children must first acquire an understanding of the rich meanings of mental state words through repeated exposure to them in context (e.g., "I thought you wanted this," looking puzzled as child refuses the object). There is now much evidence that the learning of adult-level word meaning of mental state terms is a long and a gradual process (Shatz, Wellman, & Silber, 1983; Moore, Bryant, & Furrow, 1989; Plesa, Goldman, & Edmondson, 1995). Children under four or so, fail,

for example, to differentiate "know" and "think" (Johnson & Maratsos, 1977). Specifically, while two-year-olds refer to mental states ("think," "know," "remember,") (Bretherton & Beeghly, 1982), it is not until they are about 4 that they understand the distinction between factive verbs (like "know") and nonfactive verbs (like "think") (Abbeduto, 1985). Children's performance on word meaning tasks has been found to correlate with their performance on the standard false belief tasks (Moore, Pure, & Furrow, 1990). As Nelson (1996) suggests, the process of coming to know what it means to say that others are "thinking people" and to both think and talk about them as such, is a long process spanning the early childhood years and beyond.

The role of narrative in psychological understanding

Context, for Nelson, is critical in the learning of language itself and in coming to understand and use language as a mode of representation of the peopled world. Context is not limited for her, to activity. In talking about word learning, Nelson writes

> ...the general solution to the puzzle of how words are learned at any age lies in the fact that people - children as well as adults – use *discourse context* to interpret language. They make inferences about what is *relevant* within the context of the utterance. The context they use for this purpose is *cognitive context* - in the child's case, the dynamic model of the world of events... (Nelson, 1996, p. 139-140).

So the relevant linguistic competence is not just, as Bartsch and Wellman (1995) suggest, the ability to use mental state words, but rather to understand their

discourse context. To understand people as thinking people, the language of human intentions and consciousness must be embedded in action and discourse contexts that support or "scaffold," to use Bruner's phrase describing Vygotsky's understanding of the adult role in the child's cognitive development, the child's understanding.

It is narrative, Bruner and Nelson and others claim, that is the mode of discourse most appropriate for, and fundamental to, the construction of psychological understanding. For Bruner, narrative is one of the two "natural kinds" of human thought. Unlike "paradigmatic or logico-scientific" thought with its focus on causation, "narrative" thought focuses on intention. "Narrative," he says, "deals with the vicissitudes of human intentions... It involves [as Kenneth Burke argues] characters in action with intentions or goals in settings using particular means" (Bruner, 1986, p. 16-20).

Nelson defines narrative as the "unfolding of events through time as told from a particular perspective of time, person, and situation" (Nelson, 1996, p. 189). In so doing, she joins Donald (1991) in claiming that "... narrative is the 'natural product' of language; it precedes and is the source of theoretical thinking" (Nelson, 1996, p. 184). So Nelson prefers not to talk about children's "theory of mind," a phrase that implies that children's understanding of self and other "is constructed ... in terms of abstract principles rationally and coherently organized" (Nelson, 1996, p. 302). Instead she prefers to ask, "What do children know about others' motives, goals, and actional dispositions within situations?" Answers to those questions may be found by listening closely to children's narrative recounting of the experience of self and other in a specific action context.

From different theoretical perspectives, then, the thinking of Nelson and Bruner converges on an understanding of narrative as providing the structure and context for children's emerging understanding of others as those who think as well as act, whose actions are motivated by particular goals and intentions. Bruner and Nelson have a shared focus as well. Both emphasize the need to integrate two elements of experience in narrative, what Bruner calls the "landscape of action" and the "landscape of consciousness."

> ... a story must construct two landscapes simultaneously. One is the landscape of action, where the constituents are the arguments of action: agent, intention or goal, situation, instrument, something corresponding to a 'story grammar.' The other landscape is the landscape of consciousness: what those involved in the action know, think, or feel, or do not know, think or feel. The two landscapes are essential and distinct... (Bruner, 1986, p. 14).

Nelson's developmental theory gives sequence to those two landscapes in early childhood. The child, she claims, first inhabits the landscape of action; his developmental task is to figure out what he and others are to do in certain routine events. Only later, as his language skills develop in the context of an expanding experience of the world of action and discourse, does the child come to recognize and inhabit the landscape of consciousness. One interpretation of the difficulty 3-year-olds have with the Wimmer and Perner (1983) false belief task in which the child is asked where Maxi (who believes wrongly that the chocolate is in the drawer) will look for it, is that

> [Up to about the age of 4, children]... cannot simultaneously tell the story in the landscape of action "The chocolate is in the cupboard" and in the landscape of consciousness "[Maxi] thinks the chocolate is in the drawer" (Astington, 1990, p. 155).

Similarly, Astington claims, in the appearance-reality tasks, younger children seem unable to tell a story

> ...on two landscapes simultaneously, the landscape of reality ('it's really a sponge') and the landscape of consciousness ('It looks like a rock,' which is equivalent to 'someone might think it's a rock') (Astington, 1990, p. 155).

Children need, it is claimed here, experience with the interpretive language of narrative, used by adults to explain children's experience, in order to know that they are to pay attention not only to action, but to the state of consciousness that is said to "lie behind" it (e.g., "Yes, your sister took your toy, but she didn't know that you wanted to play with it"). It is through narrative that adults make connections, for children, between actions and thoughts.

Narrative discourse also assists children in constructing an understanding of self and other as subjective "selves" who are continuous in time. Comprehending a story depends upon the hearer's ability to take the point of view of a single character and maintain it through changes in action and setting. The child, exposed to such narrative continuities, is assisted in recognizing that, though events may be bounded ("Dinner is over when we leave the table"), consciousness is not. For example, a parent may say, "Yesterday, when you played this game, Bobby wasn't here, so he [still] doesn't know how to play the game. Why don't you show him how?" or "Remember that [two pages ago] Susan said that she wanted a doll for her birthday? Now, when she opens her present, she finds a game. She's so disappointed!" Through repeated experiences with narrative, children, it is claimed here, begin to attend to the "through-line" of consciousness that is essential to the construction of an "extended" (Neisser, 1988)

self and other, and to an understanding that people's past experiences, or lack thereof, affect their current state of mind. Narrative introduces children to the "landscape of consciousness" and its continuity. It is claimed here that experience with narratives, both fictional and personal, is an essential component of children's construction of an understanding of self and other as thinking people.

Evidence that children's competence to understand and use narrative forms is related to their success on false belief tasks, and thus, it is claimed here, to their understanding of self and other as thinking people, has begun to be gathered. Lewis writes:

> ... we contend that three-year-olds [in general, fail the false belief tasks] because they fail to grasp the causal texture of the false belief narrative. Ordinarily events pass children by and they cannot retrace vital links between the landscapes of action and consciousness when asked to retrace the protagonist's mental state...(Lewis, 1994, p. 462).

His studies reveal "a linear relationship between narrative fluency and success in the false belief question [we asked]: Poor narrators were largely unsuccessful, average children largely successful, and good narrators were at ceiling" (Lewis, 1994, p. 475). Further evidence of specific connections between narrative ability and the understanding of self and other as both actors and thinkers is needed. This study will provide correlational data that are expected to highlight the relationship between language, particularly narrative skill, and the child's developing understanding that "you" and "I" are thinkers as well as doers.

Theoretical foundations of the research design

Whether there are differences between children's self- and other-knowledge during this transitional period is a question that is, for various theoretical reasons, hotly debated. Whether children first understand themselves and then others, or first others and then themselves, or whether the understanding of both is simultaneous, could offer a clue to the process by which children make sense of human action and thought.

The entire debate about self- and other-understanding, though, especially as it is carried on among those interested in the 3 ½ to 5-year-old's developing "theory of mind," rests upon very little data, collected in the service of answering very specific theoretical questions. Strikingly, no one has systematically explored or described the relative change in self and other understanding across this period of dynamic change. This study has as its goal that exploration and description. It asks: Are there patterns of difference between self- and other understanding across this transitional period, as children begin to demonstrate their knowledge that people are thinkers, as well as actors? What does children's ability to talk about joint experience, using "us" and "we," tell us about their changing psychological understanding and narrative skill?

This study is built upon the theoretical foundations laid by Vygotsky, Bruner, and Nelson. Those foundations affect the research design in the following ways:

1) Children's psychological understanding is not derived from their ability to generalize or theorize about "the mind" as a cause of human action, as claimed by those who study children's "theory of mind." Children, rather, it is claimed here, want to know and talk about "where mommy thinks I am" or "what toy [my sister] Sue will

want to play with." So, the paradigmatic "theory of mind" method will not be used here. That is, rather than inviting children to a laboratory in which they are asked to watch a Maxi ("the other," who is a puppet or a storybook character and thus a pretend, or an abstracted "generalized other"), listen to his story, and infer his false belief about an event observed by "the self," two real children ("self" and "other") will be invited to participate together. To draw this study into conversation with "theory of mind" research, though, one of the experiences these real children will have, will be parallel to the deceptive box task used in theory of mind research.

2) Since an understanding of people as thinkers (believers, knowers, etc.) is thought to originate from children's experience of linguistically interpreted and supported activities, this study will provide such an activity for the children. Two children will be taught a novel game by an experimenter who will structure their activity through instruction and commentary on the children's action and thought: "What do you think will be in the box?," for example. Such a situation is common in children's lives. In early childhood, as children are just beginning to construct a linguistically supported understanding of self and other, they often interact with one another under the supervision of adults who intervene to interpret self to other and other to self. Children's attention to and use of adult's verbal input in constructing their own understanding of the experience are of interest here.

3) Given the claim here that narrative is the form we ordinarily give to our psychological explanation, children's recounting of the game playing event will be analyzed for the presence of certain narrative elements: protagonist (Whose story is it, self, other, both?), action sequence and connective use (Does children's ability to

sequence actions correctly relate to their ability to recount the experiences of self, or other, or "us"?), and so on.

4) Because children's representations of events and the people within them, are constructed by the child based upon their own interpretation of the event (the relative salience of particular features, their understanding of its goals, their role within it, etc.), children will not simply be questioned about the event, but will first be asked, a day or so after playing the game, to recount their experience of that game event to the experimenter. Giving children time and opportunity to reflect upon the event and then to spontaneously describe it, will give the experimenter a better idea of what elements of the experience have become the child's "own." This will not be a "memory test" in that the game materials will be set out as they were in the game playing session, to remind children of the event itself.

5) Because language and narrative skills are thought to relate so intimately to children's understanding of self and other, independent tests of those skills will be given, to see whether scores correlate with growing complexity of psychological understanding. For example, is their ability to recount their own story related to their ability to retell a fictional story? Do scores on a language test correlate more highly with demonstrated self-knowledge than with demonstrated other-knowledge?

6) a. Since a careful comparison of self and other understanding is the goal of this study, it is necessary to set up a common situation for all pairs of children. By asking each pair to play the same game with a fixed and known (to the experimenter) sequence of events and resulting mental states, it will be possible to both assess

children's accuracy in their recounting, and to know what elements children left out of their stories.

b. In order to assess children's ability to differentiate the experience of the self from that of the other, it is also essential to make the experience of the "self" different in some ways from that of the "other." As they play together, certain elements of the game (what each finds in a box, for one thing) will differ for self and other.

Building upon these theoretical foundations, this study will offer a systematic way of directly asking a question which has, so far, been asked only peripherally. It will give us a way to explore and to describe the relationships among children's use of "self," "other," and "us", their articulated psychological understanding, including mental states, and their narrative skill in this period of dramatic change. We will be able to address the questions of task- and person-neutrality in children's developing ability to understand themselves and others not just as actors, but as thinkers as well.

METHOD

Participants

Participants were fifty preschool children (26 boys; 24 girls) ranging in age from 3½ to 5 years (M = 51 months, SD = 5.6). Fifteen were between the ages of 3½ and 4. Seventeen were between 4 and 4½; 18 children were in the 4½ to 5-year age range.

These 50 children had been reported by their teachers to be English speakers with no known difficulties in comprehension or expression. Those reports were considered confirmed when the children spoke to their game-playing partners in session one, and attended and responded to, the verbal tasks in sessions two and three. Eight other children began the study, but are not included here either because their fluency in English was not confirmed or because they did not complete all three sessions.

Though children in the same age group (group 1 = 3.6 to 4; group 2 = 4 to 4.6; group 3 = 4.6 to 5) and of the same sex were paired to play the game in session one, the dependent variables in the study are derived from individual sessions two and three. Children's responses were used even if their game-playing partner's responses were excluded for the reasons mentioned above. There was one exception to this rule. One child in a particular pair, not fluent in English, was silent throughout the game-playing session. Because she did not give her partner any information about what she was thinking or feeling, both partners' responses were excluded from this analysis.

The 50 participants attended one of six daycare centers and preschools in metropolitan areas of Pennsylvania and New York and composed a group that was diverse in a number of ways. The racial composition as described by parents was: 10% Black, 22% Hispanic, 56% White, and 12% biracial. Socioeconomic status, as indicated by the source of daycare or preschool tuition, ranged from those who were fully subsidized (one family, or 2% of the sample) to those who were unsubsidized (52%), and included those (46%) who received partial subsidy.

Children's social experiences, both within and outside the family, also varied. 30% lived in single-parent households with their mothers; 66% lived in two-parent households; 4% lived with two parents and two grandparents. The children had a wide range of preschool experience. Some were newly-enrolled and others who had been in daycare since infancy (range = 0 to 59 months; M = 23 months, SD = 17). Hours per week spent in daycare or preschool ranged from 4 to 50 hours (M = 32 hours, SD = 17).

Materials

The investigator designed the game used in this study, and tested it on twelve 3½ to 5-year-old children in a suburb of New York City (ranging in age from 41 to 58 months, M = 49 months, SD = 5.8). They seemed to enjoy both the game objects and activities, and all, despite a wide range of language competence, were willing and able to talk about the game. As expected, some talked only about the actions while others made references to the internal states of the game players, mentioning what

someone had "liked" or "tried to figure out," for example. Because the game invited the kinds of responses of interest in this study it was used as the primary activity here.

In the current study, there were three sessions. Session one was the game playing session. Game materials were arranged in order of their use, on a table in a separate room in the center or school. Most centers provided large round tables, but in one school the table was rectangular. The figure below shows the most commonly used table arrangement as the children saw it when they first entered the room.

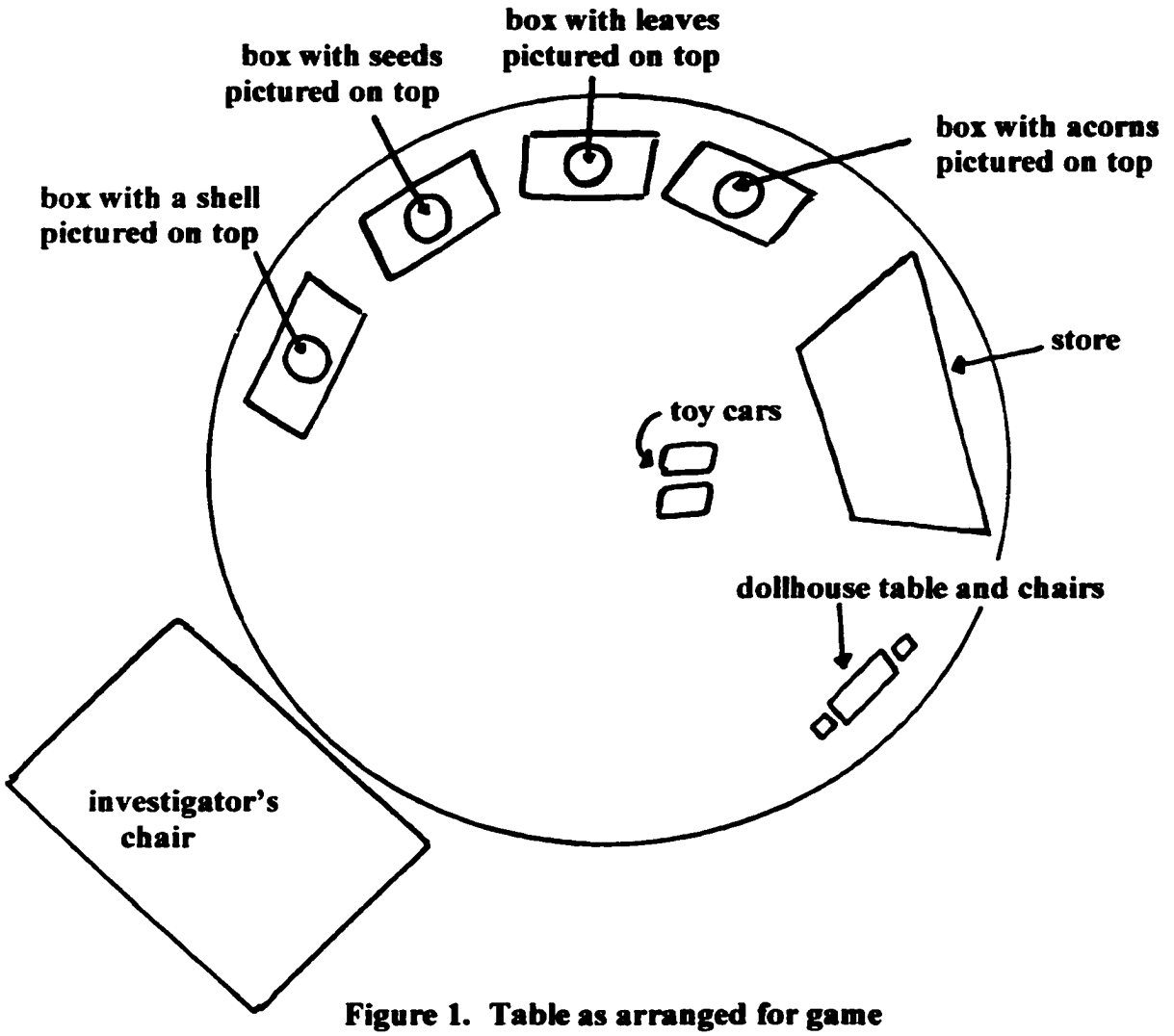


Figure 1. Table as arranged for game

The other objects used were shown to the children as they were needed for the game. Those objects included two 8-inch by 12-inch cards that were blank on one side. The other side had a number and pictures of two fingerpuppets. The card that had a "1" on it showed a picture of a parrot and an octopus; the card that had a "2" on it showed a picture of a goat and a chipmunk. There was also a flip book with a cover that pictured all four fingerpuppets and was titled, "What do they eat?" Inside were four pages, each with two pictures, one of a fingerpuppet and one of the food that that animal eats. The foods pictured are precisely the same as those pictured on the tops of the four boxes. Also out of the children's sight until it is introduced is a bag of dollhouse-size plates, silverware, and cups that the children will use to set the table at the end of the game.

For session two, all the game objects except the finger puppets were again set out on the table to help children remember the game objects and activities, as they recounted the game playing experience. Also used in this session was the text and black-and-white pictures of a twelve-page story. The story was a Japanese folktale, translated into English, called "The Dragon's Tears" (Brown & Smiley, 1977) and offered for use in the study by Lucia French (French, 1988; French and Pak,1995). The story, used in research on children's narrative skill, was selected because it contains references to characters' emotional and mental states. Its retelling provided an independent measure of children's ability to include psychological states in a narrative. See Appendix A for a sample page and the complete text of the story.

During the third session the investigator showed children a raisin box filled with crayons and posed the standard questions used in the theory of mind deceptive box

task (see session three below). She also used the materials provided to administer the Test of Early Language Development (Hresko, Reid, & Hammil, 1981), referred to in the following pages as the TELD-2.

Procedure

The associate executive director of a group of twenty daycare centers in urban and suburban eastern Pennsylvania gave permission for the study to proceed, and selected five centers for participation. The centers were selected because they had a sufficient number of children in the needed age range, and a room for use during the study. Two independent preschools, one in suburban Pennsylvania and one in New York City, were also contacted; the on-site directors gave permission for the study to proceed.

In each case, the on-site directors distributed informed consent letters and consent forms and collected the signed forms from parents. The directors also provided the following information: 1) the child's racial/ethnic background, as described by parents, 2) the source of payment of the child's tuition, as an indication of socioeconomic status, 3) the length of time the child had attended the daycare center or preschool, 4) the number of hours per week the child spent at the center or school, 5) whether or not the child had any known difficulties with English, and 6) what languages were spoken in the home.

Classroom teachers, because of their daily interaction with the children, were also able to provide important information. They made assessments, for example, of the

game-playing partners' degree of familiarity with each other. Degree of familiarity was operationalized as the number of hours per week the pair spent in the same room engaged in the same activity. Classroom teachers, having the most frequent contact with children and their families, were also most often the ones who provided information about the composition of children's households.

The time that elapsed between the sessions was held as constant as children's schedules allowed. Each of the three sessions was held on a different day. Session two, the game recounting, question answering and Dragon's Tears story retelling, most often twenty-four, but never more than 48, hours after the game playing Session one. The third session, the administration of the deceptive box task and the TELD-2, was always held within the week that the children participated in the other two sessions. One child's preschool schedule dictated that the session three tasks (the deceptive box questions, and the TELD-2) be administered before the two consecutive sessions (game playing and game recounting). Forty-seven of the fifty children met with the researcher on three consecutive days during a single week, with the order of sessions following the above description.

All sessions with all children took place in a separate room in their daycare center or preschool. Usually the same room was used for all three sessions for each child; the same room, set up in the same way, was always used for sessions one when the game was played, and session two when it was recounted.

Session One

General Procedures

Each session began when the investigator went to the child's classroom to greet the child. The investigator asked him whether he would be willing to go and play (or talk with) the investigator in another room, and, if he said 'yes,' to walk him from the classroom to the room where the session was held.[1] For session one, two children were greeted and invited to participate. Children were paired in consultation with classroom teachers on the basis of their being the same sex, in the same age group (i.e. both between 3 ½ and 4, or 4 and 4 ½, or 4 ½ and 5) and their having daycare schedules that permitted them to participate in session one together. Pairs were always made up of children from the same school; not all were members of the same class. If the children did not know each other, the classroom teacher or the researcher introduced them to each other and explained that if they were willing, they would play a game together. Any further explanations that the children requested would be given, and the two children were walked to the classroom by the investigator. During the brief walk, the three talked; children's curiosity about the game and expectation that the game would be fun was encouraged.

When they arrived at the designated room, the researcher explained the need for the video camera that was set up about eight feet from the table. She also explained, and attached the remote microphones to the children's collars or pockets, and gave children a few minutes to get used to wearing them. Throughout this time, usually less than five minutes, children and investigator engaged in natural conversation. The

only questions that the investigator did not answer were questions about the game objects on the table. Children who asked about the game were told that in just a minute the researcher would "tell you all about the game." After the microphones were attached securely to each child, the investigator took her seat at the game-playing table and asked the children to stand together next to her while she told them how to play the game.

All pairs of children were given exactly the same game-playing instructions (see below). In addition to explaining the game, the instructions were designed to encourage children to attend to the inner states of "self" and "other" child, and so included such words as "like" and "know" and "think." While each pair of children heard the same instructions, there were two main ways in which the talk during the game playing session varied from pair to pair: 1) amount of regulatory language required (e.g., "Please wait here until [partner's name] closes her box") and, 2) amount of talk the children engaged in, both with the investigator, and each other. Most pairs did most of their talking at the end of the game while they set the table for the animals' snacks, though some talked a bit to each other as the game progressed.

Nine activities, or steps, were included in the game. Eight required the use of the objects pictured above. For some steps, the two children acted simultaneously (looking through the book to find what their animals ate, for example, and setting the table for the animals' snack) and for others, they took turns. Player A, for example, looked in her box and then waited and watched while Player B looked in hers. The steps, labeled as simultaneous or turn taking, are listed below, with the investigator's script that accompanied them. For each step, there were common elements in the two

children's actions and experiences, e.g., both looked in a box. There were also distinct elements. Player A found money in his box while player B found nothing. Those differences were intended to give rise to distinctive expectations and reactions. Player A, who saw a picture of her animal's food on the top of the box was surprised, and usually pleased, to find money in the box. That surprise and pleasure was demonstrated most often in facial expression and tone of voice. Player B's expectation was then, that her box would contain either food or money, and was often surprised and either disappointed or confused when she opened the box and found it was empty. All of those who were Player B seemed to accept the investigator's explanation that "Maybe you don't need money to get food for your animal."

Game Instructions

The steps and standardized instructions for the game follow. Inner state words are italicized. The children are given general instructions about the whole game first, and then the step-by-step instructions begin.

"[To both children] Your job in this game is to feed two animals. [One child's name], you will feed one, and [other child's name], you will feed a different one. The game is over when the two animals both have some delicious food and they are sitting at that table [pointing to the dollhouse-sized table and chairs set out on the table] ready to eat. OK?"

Step one in the game requires that the children take turns picking cards to determine who will go first in the game. "I have two cards here. One says '1' and

one says '2.' The person who picks this card [pointing to the card with the '1' on it] will go first. [Investigator puts the cards behind her back]. Who *wants* to pick a card? [Some children talked about this, but in most cases one simply raised a hand and then picked. Then the child who did not raise his hand took the second card. Both children show the investigator which cards they have]. So [name of the child who will be player A in the game], you will go first, and [other child who is player B], you will go second."

In **step two**, the children take turns looking at the finger puppets pictured on their cards and choosing which of the two animal puppets they want to play with. The investigator says, "On your card, [player A's name], there was a parrot [experimenter brings the parrot puppet out from under her chair and shows it to the child. She does that in turn for each puppet mentioned] and an octopus. Which animal do you *want* to feed? [Child picks, usually by pointing, one of the two]. OK. Let's put the [chosen animal] on your finger. On your card [player B's name], there was a goat and a chipmunk. Which animal do you *want* to feed? [Child B picks one and investigator puts it on the child's finger]. [To both children] Do you *like* your animals? [Investigator gives the children opportunity to respond, saying anything they'd like to say about the animals. Most children simply said 'yes.'] Now you each have an animal to feed. Good!"

In **step three** the investigator says, "Would you like to name your animals?" [If one answered and the other didn't, investigator went on to say] "That's a good name! [Other child's name], would you like to name yours?" [then, if given] "That's a good

name too!" [If either or both children did not name their animals investigator said] "That's fine. You don't have to name them."

In **step four,** the two children look through the book provided for the picture of their puppet next to the food that animal likes. To begin, the investigator says, "But do you *know* what your animals like to eat? [Children who respond to this question before looking in the book are encouraged to talk. If they name a food that is correct, but not in the book and game, investigator says, 'Yes! [Chosen animals] do eat that, but in this game we have just one special food for each animal to eat.'] Why don't you [addressing both children] go over to that book and look inside to see what foods we have here that your animals like to eat? [Children flipped the pages of the book together, pointing things out to each other, and stopping at the pages that pictured their two chosen animals].

[Player A], do you see a picture of your [chosen animal] and its favorite food? [Investigator waits for the child to say what animal eats, giving encouragement and prompting if necessary]. [Player B], what does your animal *like* to eat? [waiting for an answer and helping if necessary]. Good. Now that you know what your animals eat, you can look for the food that they *need.*"

In **step five,** children take turns looking in one of four boxes. "[Player A], where do you *think* you could go to find your [chosen animal]'s food? [Child usually pointed to one of the four boxes that has the chosen animal's food picture on top. If the child started to walk toward it or reach out to open it, investigator asked them to 'stay here for just a minute']. What do you *think* is in that box [pointing to it]? [Player A answered, usually with the name of the food pictured]. Well, let's see.

Why don't you go over and open the box? [Player A opened the box, usually lifting the money out and announcing that "It's money!" or "It's dollars!"].

Now, [Player B], it's your turn to start looking for the food your [chosen animal] needs. Where do you *think* you should look? [Child suggested, or pointed to, the box with the appropriate picture on it]. What do you *think* is in that box? [Child said either 'money' or 'food']. Well, let's see. Why don't you go ahead and open it up? [Child said 'It's empty' or 'nothing.'] That's ok. Maybe you don't *need* money to get your animal's food."

In **step six**, the children take turns going to the store to "buy" the food their animals need. "[Player A,] now it's your turn again. You have money, but [chosen animal]s don't eat money, do they? Where do you *think* you could go next to find food? [If child didn't respond, researcher said] Where does your mother go when she has money and *needs* to buy food? [If child said 'the store,' investigator said] Yes. There's a store right there [pointing]. [If child said nothing, investigator said] She goes to the store, doesn't she? Well, there's a store right there [pointing]. Why don't you go, open it up, and look inside? [After child walked over, and opened the store sometimes with help from the investigator, she goes on to say] Do you see the kind of food your [chosen animal] *needs*? To get the food, put one dollar in the store. Then you can take the food out. Great! Now your animal has food to eat. But your animal can't eat until [Player B]'s animal has food, right?

So [Player B], now it's your turn to find food for your animal. [Player A], you can help [Player B] get the food [his/her] animal needs. OK? [Addressing Player B] Where do you *think* you could find food? [Child pointed to the store, or walked over

to it, joining Player A there]. Do you see the food that your animal *likes* to eat? [Child said 'yes.' Often Child B points to it]. What do you *need* to do to get the food? [If children couldn't remember, or if Player A didn't offer Player B the unused dollar bill still in his hand, investigator went on to say]. [Player A], do you have something [Player B] could use to buy the food? What does [Player B] *need* to put in the store? [Player A handed Player B the money, pointed to the place in the store where the money belonged and Player B put the money in and took the appropriate food out]. Great! Now both of your animals have the food they like to eat!"

The **seventh step** in the game is optional. Children are invited to use the toy cars to drive their animals home from the store. "Now you can take the animals home for their snack. If you *want* to drive them, you can use those two cars ["parked" behind the store]." One car is red and one is green, so some children expressed a preference and claimed one car or another. Then the children both drove to the place where the small table was set up.

The last game activity, **step eight**, is setting the table. The children do this together. "[When the children get their animals over to the table, investigator says] *remember* I told you that the game is over when the two animals are sitting at that table? Now it's time for you to set the table together. Here are the dishes [bringing a bag full of dollhouse-sized plates, knives, etc., and pouring them out on the table for the children to use]. You can talk about what you *want* to put on the table. Then set the table. When you're finished, and both of your animals are sitting at the table ready to eat, the game is over. Please tell me when your animals are ready to eat, OK?"

As soon as the children began setting the small table, the investigator walked away from the table where the game materials are set, and sat in another part of the room while the children talked and set the table. The length of time it took varied with the amount of talking that the children did or simply with their involvement in the task of putting different dishes on the table. Some pairs of children needed to be prompted to tell the investigator that they were finished. When they did that, the game was over. The investigator then thanked the children for playing the game, removed the remote microphones, and walked them back to their classroom. She told them that she would be back the next day to talk with each of them about the game.

All pairs of children maintained their interest in the game throughout the session, and many asked, when the game was over, to "play it again," choosing a different animal. Because a second playing would make an accurate recounting of the first game impossible, the investigator had to tell them that it could only be played once. The length of time it took to play the game varied depending on how long each pair spent setting up the table. Up to that point, the game took about ten minutes to play.

Session Two

General Procedures

Usually within twenty-four hours of session one, the researcher returned to each child's classroom to ask the child to go with her to the separate room in the school where the game had been played. She explained that "today we're going to talk about the game you played yesterday with [player's name]." When they arrived there, the

child saw that a game object representing each step in the game was set out on the table. The puppets were not set out. When children requested them or asked to play again, the investigator reminded the child that "Today we're just going to talk about the game, remember?" The video camera was again set up in the room to record the session. The investigator attached a remote microphone to the child's collar or pocket and put the second microphone on her own collar. For this session, two chairs were at the table. The chairs were both far enough from the game materials that though the children could see the boxes, store, and the small table and chairs, they could not easily reach out and touch the objects while they talked. Some children, though, did need to be shown the book to prompt their memory, and pointed to the pictures in it saying, "I picked this" rather than saying "I picked the goat."

The session consisted of three parts: 1) the child's recounting of the game playing experience, 2) the investigator asking some standardized questions about the game to determine whether the child remembered aspects of the self and other's experience that had not been spontaneously reported in the recounting, and 3) the investigator reading the Dragon's Tears story twice (see Appendix A) and asking the child to turn the pages, look at the pictures, and retell it. Sessions lasted between twenty and thirty minutes, depending upon how elaborate the child's recounting and story retelling were.

Game Recounting

First, the child was asked to recount the game playing experience. The investigator began by saying: "Yesterday, you and [player's name] played a game together, remember? Your mother wrote her name on a piece of paper telling me that

it was ok for you to play the game, but she wasn't here to watch you play, was she? [waits for the child to acknowledge that that was the case]. She wasn't here when you played the game, so she doesn't know anything about the game. Let's pretend that you're telling your mother all about how you and [partner's name] played the game, ok? Let's pretend you're telling your mother all about the game from the beginning to the end. You're telling her all about how you and [partner's name] played the game, ok? I'll start and then you can tell the rest, ok? ... First you and [partner's name] walked into the room with me and I put the microphones on you. Remember? Then it was time to start the game. I hid these two cards [showing them to the child] behind my back [putting them there] and then [showing the child the cards]... Now it's your turn."

Children were given opportunity to recount the game spontaneously, including whatever aspects they chose. If a child spontaneously gave a recounting that included all the steps and objects in the game, the investigator simply said, "Thank you very much!" when the recounting was over. Most children, however, required some prompting. Initially the prompts were general ones: "What happened next?" or "And then what?" All children recounted some aspects of the game experience without more specific prompts. When children came to the end of their spontaneous recounting, as indicated by long silence or an explicit comment like "that's all," the researcher pointed to any of the game objects that represented steps in the game that the child had not yet recounted, and said, "What happened there?" When she had finished pointing to those objects and asking the child about them, she said, "Is there

anything else you would tell your mother about the game?" If the child said, "no," part one of session two was over.

Standardized Questions

Then the researcher told children that she would like to ask them some questions about the game they had played with their partner. One set of eleven questions was asked twice, differing only in the person who was the subject of the question (i.e., "self" or "other"). The order of questions (whether the "self" set or "other" set was asked first) was counterbalanced across subjects to avoid order effects; the results of that analysis will be reported in the Results section below. The questions, as asked about the self were:

1) Which animal did you pick?
2) What did you name your animal?
3) Did you like your [chosen animal]?
4) You looked in this book and found out that your animal eats [appropriate food]. Which box did you go to look for your animal's food?
5) Before you opened the box, when it was closed just like that, what did you think was in the box?
6) What did you say would be in the box?
7) What was really in the box?
8) How did you feel when you saw what was in the box?
9) Were you surprised when you saw what was in the box?

10) Where did you find your animal's food?

11) Did you need help to get the food?

The "other" questions were "Which animal did [partner's name] pick?" and so on.

Answers took a variety of forms. Sometimes children simply pointed at the correct animal picture or box. Children often gave one-word answers or general, rather than more specific ones. Many children, rather than saying that they felt "happy," for example, said that they felt "good" when they found money in their box. Since children's language skill varied greatly, and the investigator wanted the questioning to be as comfortable as possible for the children, any level of answer that indicated an understanding of the question was accepted. All were attentive to the questions and all children were thanked for having answered them.

Story Retelling

The last part of session two was a story retelling, used to provide a measure of the child's narrative skill and psychological understanding in a context in which the words and accompanying pictures were provided. The story was introduced with "Listen very carefully. I'm going to read you a story. I'll read it twice. When I'm finished, I want you to tell me as much as you can about the story." During both readings, the investigator turned the pages and read the text exactly as written. Some children pointed things out and made comments as they listened; others listened in silence. The researcher tried to neither encourage nor discourage children's talk. She

listened with interest to what the child said, but did not give children additional information or repetitions of the story.

When the second reading was finished, the investigator said to the child, "Now you can turn the pages and tell me the story." If children hesitated, or got stuck on one page, or protested "But I can't read!" they were prompted: "You don't have to read the words." "Just tell me anything you can about this page." When children couldn't say anything about a particular page, they were encouraged to turn to the next one. The researcher expressed appreciation to all children when they were finished, and gave them stickers to thank them for their participation.

Session Three

General Procedures

The last session, most often held on the third consecutive day, consisted of a standard theory of mind deceptive box task and the administration of the TELD-2. Again, children were asked if they were willing to participate, and those who said "yes" were accompanied to a separate room. No video camera was used on that day, but an audiocassette recorder and microphone were set up on the table where the game materials had been for sessions one and two. Usually children were interested in the equipment on the table, so were invited to talk into the microphone and hear their voices on tape before they began the session. They then were able to turn their attention to the tasks at hand. The child and researcher sat in the same seats they had occupied during session two.

Raisin box task

First, the investigator brought out and showed the child a closed raisin box. She asked, "What do you think is in this box?" Most often the child said "raisins," but other answers ("grapes" or "these [pointing to the picture]") were also given. Then the researcher opened the box and said, "What's in the box?" "Crayons!" was the correct answer. Then she closed the box again and said, "Before I opened the box, when it was closed just like this, what did you think was in the box?" When the child answered, he was then asked, "What did you say was in the box?" and then, "What was really in the box?" Finally, this question was posed: "If your teacher came in right now, and saw the box closed just like this, what would she think was in the box?" Children's answers helped the investigator to assess their ability to pass this standard theory of mind task designed to measure their understanding of false belief.

Language testing

Lastly, the children were given the TELD-2, a widely used and available test of early language development. The starting point for each child's testing, designated in the TELD-2 examiner's manual (Hresko, Reid, & Hammill, 1981) depended upon the child's chronological age. Questioning continued until the child gave five incorrect answers in a row. When the child completed the TELD-2, the investigator thanked him, told him how much she had enjoyed her time with him, and gave him a choice of either more stickers or colorful erasers to show him appreciation. She then accompanied the child back to his classroom.

Coding

Session Two Coding

Coding the Game Recounting

Before determining the precise coding strategies and categories, it was necessary to transcribe sessions one and two in full, recording all of the talk of both children and the investigator. The investigator chose to use the CHAT system, "...a standardized format for computerized transcripts of face-to-face conversational interactions" (MacWhinney, 1995, page 1), the transcription system used for the Child Language Data Exchange System, known as "CHILDES." Transcripts and videotapes were available for reference as the coding proceeded.

Two abbreviated transcripts were made from the full transcript. The abbreviated transcripts enabled the coder to readily focus on the language the child used during the specific verbal exchanges of interest (i.e., the game recounting and the story retelling). They also eliminated irrelevant portions without denying the role of the investigator in the exchanges. In the Dragon's Tears retelling, the abbreviated transcript begins when, having read the story to the child twice, the investigator says: "Now I'd like you to tell me as much as you can about the story." The transcript ends when the investigator, seeing that the child has finished turning the pages and is now quiet, thanks the child for having participated. The investigator's prompts were designated on the transcripts using the same system described below for the game recounting. The prompts were intended to encourage the child and put him at ease:

"Anything you want to tell me is fine" or "When you're finished, go ahead and turn the page."

The abbreviated transcript of the game recounting (see Appendix B for a sample) begins when the investigator has completed the standardized instructions ("... let's pretend you're telling your mother all about the game from the beginning to the end..."). The investigator starts the retelling by shuffling the cards (see description above), showing them to the child and saying, "and then what happened?" A few children began the retelling before the researcher instructed them. In those cases, the abbreviated transcripts include the game-relevant statements made before the instructions. The abbreviated game transcript ends when the investigator asks the child whether there is anything else he would like to tell his mother about the game and he says "No." If a child continues talking about the game, those further comments are included in the transcript.

The researcher's contribution to the game retelling, though not appearing word for word in the abbreviated transcript was available for reference on the full transcript. Abbreviated transcripts for both the game recounting and story retelling cued the coder to the presence of investigator utterances, and indicated the degree of constraint which her comment put upon the child's next utterance. "S" means "support" and indicates that the investigator simply repeated the child's last utterance or said "uhhuh" or "ok" or the like. "GP" means "general prompt" and refers to comments like "What happened next?" and "And then?" "SP," rare in the story retelling transcripts but common in the game recounting, stands for "specific prompt" in which the investigator points the child's attention to a specific object or location and asks

the child to comment on it. "SP/boxes," for example, on the game recounting transcript tells the coder that the researcher has pointed to the boxes on the game table and said, "What happened there?" Two other types of investigator utterances while rare, are noted in the transcripts. "U" indicates that the child was asked a question because a previous comment had been unclear. The child might have been asked to repeat something ("What did you say?") or to specify further the meaning of the previous utterance ("'playing' with what?"). Sometimes the researcher elaborated on something the child had said. An example is when the child says, "She picked this one [pointing to the picture of the parrot]" and the investigator says, "the parrot." Those kinds of utterances are marked with an "E" for "elaboration." As few of those comments as possible, were made.

In abbreviated game recountings, like the full transcriptions, each utterance (defined as a meaningful unit of speech marked by pauses both preceding and following it) is typed on a separate line. Utterances can be as simple as a single word (i.e., "parrot [pointing to the picture]") or as complex as a grammatical sentence with two verbs (i.e., "We took the money and bought the food.") Utterances were the units considered when making assessments of topic.

The first important question asked was: Which utterances are references to yesterday's game and thus coded further, and which should be excluded? Reasons and rules for exclusion follow:

1) Off-task utterances were considered to be of interest, as a measure of the degree of the child's attentiveness to the recounting task, so were counted. A variable, calculated as the percentage of total utterances that were "on" the task of game

recounting, is included in the study; off-task utterances were not further coded as part of the game retelling. Examples of such excluded utterances are "Tomorrow's Halloween and I'm gonna be a clown" and "What are those blocks [in another area of the room] for?" Perhaps most common and important to the study were children's requests to "play the game [again] today." Those, too, were considered to be off-task and excluded from further study. It was important to the findings of this study, though, that they could reliably be distinguished from those utterances that were about "yesterday's" game, but expressed in the present tense. Two independent coders, counting the off-task comments in a sample of 16% of the abbreviated transcripts, achieved an interrater reliability rate of .96. Rates of agreement for each transcript were calculated by taking the number of off-task utterances chosen by both coders and dividing it by the sum of that shared number and the total number of unshared utterances chosen. A percentage of agreement for each transcript was derived; the percentages for all eight transcripts were summed and divided by 8 to arrive at the reliability rate reported above.

2) Two characteristics of natural conversation were not of special interest in this study, and so were excluded. False starts and word-for-word repetitions were not coded further.

3) Since it was children's spontaneous recountings of the game that were of interest, utterances made in response to a direct question were excluded from the count. If, for example, the investigator did not hear what the child said and asked, "Your partner picked the what?" the child's "He picked the octopus" was not included in the count. Most often those comments were also repetitions.

4) Many children made metanarrative comments, like "I don't know" in response to "What happened there?" Those utterances were excluded from further count; their inclusion would have inflated by 100% the number of mental state verbs used "in the game retelling." No child referred to the mental state of self or other during the game playing; that is, no child said, for example, "He thought there would be money in the box, but there wasn't" or "I was surprised to find money in my box." The only use of mental state terms in the game recounting was formulaic and in metanarrative comments. The investigator noted the number of metanarrative comments a child made, but the use of such comments did not relate significantly to any other variable, so were excluded from further analysis.

Having made these exclusions, the distinctive (as over against repetitive) utterances that related to "yesterday's" game were coded. The number of on-task utterances ranged from 4 to 29 ($M = 13.10$; $SD = 5.85$).

Further coding was done to derive variables measuring: 1) frequency and complexity of the use of self, other, and both; 2) psychological understanding as demonstrated in children's narratives; and 3) narrative skill. The relationships among these three are at the heart of this study.

Of primary interest is the question of children's relative understanding of self and other as people who participate in shared activities, but have different inner lives. Therefore the coding of person in the game recounting narratives was done with some specificity. It was not considered enough to code the topic of each utterance, asking "Who is the actor here?" or "Who is this utterance about?" Instead, each reference to person was coded. That included, of course, references to the children themselves as

agents: "*I* picked the chipmunk" and "*She* wanted the pink one." It also included the game player as patient, the one affected by the agent's action ("He gave *me* a dollar"), and as possessor ("*Her* money was in there.") Other person references (to the investigator or the indefinite "you," to the puppet as "he...") were coded, but proved to be too few to be used in statistical analyses.

Whole utterances could not be the units of analysis because, since many utterances included more than one person reference, and much valuable information would have been lost. The coder instead was attentive both to meaning and grammatical form when coding for children's use of person in the game recounting. The variables, then, were called "number of references [to self], [to other], [to 'we'/us'] in game recounting." The total number of references to person were used to derive percentage variables: "% of [total number of person references that were] self reference," etc.

In order to find out how attentive children were to the differences between the experience of the "self" and the "other", transcripts were coded for the number of distinctive experiences reported and whose experiences they were. An example of the reporting of such differences is when children said, "He went first" or "I found money in the box, but [Partner's name]'s box was empty." Virtually all children showed evidence that they had distinguished the objects and actions of each child from the other, so further analysis of children's awareness of difference was not done.

Psychological content was coded separately for the three often-referred to grammatical persons: the "self," the "other," and the pair ("we" and "us"). The goal was to compare the level of psychological description of the three persons to determine whether children described them in equally detailed and complex ways or

not. For example, were the intentions, goals, emotions and mental states of the "self" described more fully than those of the "other" or of "us"? Each person reference (above) was further coded for its psychological content.

First, the transcripts were examined to determine what elements of psychological life children included in the game recounting. The various aspects they mentioned as part of "my," "her/his" or "our" game playing experience are listed below: 1) Being and possession were referred to by many children as they described a person's relationship to objects in the game; "I was the chipmunk" or "I had the chipmunk" are examples. Use of the possessive pronoun, as in "her parrot ate seeds," was included in this category, but only 3 of the 50 children (6%) talked only about these aspects of experience. Preliminary analyses did not indicate that the use of these verbs was any easier than action verbs, so the use of being and having was not considered a separate category for analysis, but was included in the "action only" category. 2) Action was the most commonly-referred to aspect of the children's game playing experience, including such utterances as "I picked the parrot" or "We drived [*sic*] the cars." Nineteen children (38%) mentioned action only when recounting the game, with another 14 (28%) mentioning both action and being or possession 3) Psychological states, broadly defined, were mentioned by 14 of the children (28%), all of whom also mentioned action. Though no child mentioned either emotional or mental states, a number of other psychological states were included. Ten children (20%) mentioned goals ("We looked in the store *to find food*") and rules ("Then we *had to look for food*). Two (4%) talked about perception ("[Partner's name] *looked* over there and he *saw* seeds"), and six (12%) referred to desires ("I *wanted* the purple animal").

Permission ("*You let us* pick an animal") and need ("[Partner's name] *didn't need* money") were mentioned by one child (2%) each. Because the numbers of references to each psychological state was small, children were scored as either having used any one of them, or not.

Four variables were constructed to describe the psychological content of each child's recounting: 1) psychological content used to describe the experience of the self; 2) psychological content used to describe the experience of the other; 3) psychological content used to describe shared experience ("we" or "us); 4) the highest level of psychological content used by the child. The first three variables have three levels:

0 = that person was not mentioned at all by the child

1= the person's being and/or possession and/or action were mentioned

2 = the action and psychological state of the person were mentioned

The variable indicating the highest level of psychological content used by the child includes only 1 and 2 above. Because every child made reference to at least one of the three persons of interest, no one had a "0" as the highest level of psychological content used in the game retelling.

The content of children's game recountings was of interest not only because it provided a view of children's understandings of persons and their experiences, but because it was an indicator of their narrative skill. How sketchy or complete was their recounting? How well sequenced were the events?

The most global assessment of the narrative's comprehensiveness was the number of total game actions included in the recounting. To aid that assessment, an outline of the game actions was made (see Appendix C). The outline includes the nine main shared actions of the game and the specific actions taken by each partner.

Completeness of the story told was measured by the number of main actions included in the child's recounting. To be counted the utterance in which the action was mentioned had to include a verb. So "we looked there [pointing to the box]" was sufficient to count as a reference to looking in the box. Saying "box" and pointing to the box in response to the investigator's question, "What happened next?" did not count. The child could receive a maximum of one point for each main action, regardless of how many separate references were made to it. The highest score possible was 9, since a number of children described "step eight" above as including two events, the setting of the table and the animals eating. Because no child mentioned naming the animals, the highest score actually received was 8. The range was 1 to 8 ($M = 4.68$; $SD = 1.79$).

Children's ability to tell an ordered and coherent narrative was assessed by finding in each recounting the longest string of in-sequence actions not interrupted by a specific prompt from the investigator. Specific prompts, designated "SP" on the abbreviated transcripts, occur when the child's attention is directed to game locations and objects the child left out of the spontaneous recounting: "What happened here [as investigator points to the box]?" Responding to a specific prompt requires that the child, who may for example have already said that she set the table and her animal ate, now answer the question "What happened here [pointing to the cards used at the

beginning of the game]?" Such prompts disrupt the sequence, and so are considered breaks in the narrative time line. General prompts only served to move that timeline along, so were not considered interruptions for the purposes of coding this variable.

The longest string of uninterrupted, in-sequence actions could include several references to the same main action. In such cases, a child who said: "We picked the animals. [Player A's name] went first. He picked the parrot" was scored as having mentioned three in-sequence actions in this case, though he mentioned only one main action. Intervening actions could be skipped as long as the two that were mentioned were recounted in order. A child could say, "We picked animals. We looked in the box. Then we set the table" and get credit for three in-sequence actions, though the narrative is incomplete. The range of actions mentioned in the children's longest strings was from 1 to 10. Those who mentioned only one action could be said not to have provided a narrative at all if narrative is defined as "the oral sequencing of temporally successive events" (McCabe, 1991, ix). However, all children mentioned action, though some did not put any two in sequence. The mean number of utterances the children put in sequence was 3.92 ($SD = 2.29$).

Children's verb use provided a means to assess their level of narrative skill. Many children used the past tense consistently and accurately: "She took the food out of the store. Then she drove home in the car." Others used the wrong form, but clearly communicated that they were talking about "yesterday": "He drived [*sic*] the red car." Some children made extensive use of the timeless present tense, making it hard to distinguish their recountings of the one-time event of playing a novel game from a description of a game they play regularly. "We open the book. Then we look in the

box." When used with the indefinite "you," ("then you go to the store and buy food"), such recountings closely resemble the "scripts" (for definition see, for example, Nelson, 1991, p. 112) children use to describe the everyday routines of their lives, such as eating lunch or taking a bath. Many children were inconsistent in their use of tense, using a combination of the timeless present and the past tense. In many cases, it was impossible to determine from the context whether the child intended the timeless present ("Squirrels eat acorns") or were referring to the past, but in the present tense ("chipmunk eat acorn," for example). In the case of such uncertainty, children's tense use was coded as "ambiguous." Each verb used was coded for tense.

Children's varying ability to use the past tense and to communicate clearly that the game they were recounting as a one-time event was considered to be an important aspect of their recountings to code. Because of this study's focus on the use of person, it seemed important to determine how many of the verbs that were clearly in the past tense had clear subjects. Many children, for example, included a number of "personless" verbs in their story retellings. Two types were differentiated from each other. One was the case in which no subject is explicitly mentioned, but the actor is not in doubt. "We went to the store and *bought* food." In other cases, it was not possible to tell to whom the child referred. One child begins the recounting by saying, "Then put it under there and take the animals out," leaving the listener to wonder "Who put it under there?" The former type was, of course, not counted as a mistake.

Since many children used verbs without subjects ("then put it under there and take the animals out"), there were differences among children in the degree to which the

narratives could be judged to be "autobiographical" (i.e., particular one time experiences of the self and specific others). An autobiography variable was constructed for this purpose and defined as the percentage of total verbs used that are in the past tense and which have a designated subject. Those sentences without understandable subjects were not considered autobiographical.

A final, widely used measure of narrative coherence is the number, and variety of connectives children use to join one clause to another. For these purposes, the "and" in "we put dishes and cups on the table" is not counted because it simply connects words to each other, but "We looked in the box and we went to the store," is counted. Most children used connectives (41 subjects or 82%), with the number used ranging from 0 to 28 (M= 5.66; SD = 5.36).

The most striking difference among children, however, was in the variety of connectives used. An empirically derived Guttman scale was constructed. The most frequently used connectives were at the lowest level on the scale, and the rarely used were at the highest level as follows. The assumption that higher levels subsume lower ones is accurate in this case; the literature is full of references to this progression (e.g., Peterson & McCabe, 1991).

0 = no connectives used

1 = the child uses "and" only

2 = the child uses ("and" +) "then"

3 = the child uses either of the above +

"but" and/or "so" and/or "because"

4 = the child uses any of the above + "when"

Coding of Answers to Standardized Questions

Children's answers to the standardized questions posed about the game played "yesterday" were coded as correct or incorrect. An independent coder watched 14% of the videotapes and read the related transcripts and scored the children's answers to the standardized questions as correct or incorrect. Interrater reliability was .95.

The following rules were applied to the coding: 1) for any one animal, a number of names are acceptable; for example, "squirrel" and "chipmunk" were often used interchangeably by the children, but clearly referred to the same animal. "Bird" and "parrot" were also both frequently used. 2) When a child, without prompting, changed her answer before the next question was asked, the first answer was considered a false start, and the second or last answer was counted. 3) A child's answer does not need to be verbal. Pointing to the correct box, when asked, "What box did you look in?" was counted as correct. 4) A child's answer does not need to be specific or given in particular, prescribed words. "I felt good" or "I was happy" were both accepted as correct answers to the question, "How did you feel when you found money in the box?" In order to compare children's facility to answer questions about the self with their ability to answer questions about the other, a "total number correct on self" variable as well as a "total number correct on other" variable were constructed.

Two questions were excluded from that count: "Before you/s/he opened the box, what did you/s/he think was in the box?" and "Did you/s/he need help [to buy the food]?" The first question was omitted because it was counted separately as one of

the mental state understanding variables. The second question was not counted because it created considerable confusion in the children, and was answered in many different ways: "The octopus helped me." "You [the investigator] helped me." "[Partner's name] gave me money." Many children gave simply a "yes" or "no" answer and it was difficult to tell whether the question had been understood. So the "Did you need help?" question was eliminated from further analysis.

Standardized questions about the "self" and the "other" were each also coded for their psychological level. If the child answered correctly only the questions about action (e.g., "Which box did you go to to look for your animal's food?"), she received a 1 for "action only." All children answered at least one action question correctly, accurately identifying the animal they had picked or the box in which they had looked. A child received a score of 2 if he correctly answered at least one action and one emotion question (e.g., "How did you feel when you saw what was in the box?"). A correct response to at least one action, emotion, and mental state question (e.g., "Before you opened the box, when it was closed just like that, what did you think was in the box?") resulted in a score of 3. These scores made it possible to compare the psychological level of "self" and "other" understanding as demonstrated in the answers to standardized questions. Variables derived from the standardized questioning are designated as "SQ" (example, "SQ # self correct").

Coding of Dragon's Tears story retelling

A number of aspects of the children's retelling of the Dragon's Tears story were analyzed to measure children's narrative ability. Comprehensiveness was measured by the number of sentences to which the child made reference. Sentences varied in grammatical complexity, but each one focused on one main action and most were accompanied by a picture of that action. There were 27 sentences, so children could score from 0 to 27. To get credit for reference to a particular sentence, children needed to make both the subject and verb clear. Pointing to dragon and saying, "monster," was insufficient for credit, but "dragon lived cave" did suffice for "And the dragon's home was in a deep mountain cave." Story sequencing did not emerge as an important aspect of narrative skill here, since children used the pictures to remind them of the order.

Grammatical accuracy was considered to be a sign of narrative sophistication, and a separate count was made of the number of verbs that agreed with the subject and were in the appropriate tense. Unlike in the game retelling, the past tense was usually, but not always the correct tense. The present tense was the appropriate form for the reporting of characters' speech: "When it *was* almost time for this funny little boy's birthday, his mother *asked* him, "Whom *would you like to invite* for your birthday party?... '...I *want to invite* the dragon' [*said* the little boy very seriously]." As in the game retelling "autobiography" variable, the number of verbs were counted, and the percentage of them that included a subject and a verb which agreed with it, was calculated.

For the Dragon's Tears story retelling, a second Guttman scale of connective use was constructed, using the same procedure described above. That scale is as follows:

0 = the child uses no connectives

1 = the child uses "and" only

2 = ("and" +) "then"

3 = either of above + "so" and/or "because" and/or "but."

4 = any of above + "when"

5 = any of above + "where"

In addition to looking at narrative complexity, the Dragon's Tears story retellings were analyzed to determine the psychological content used by the child. Coding required attention to the content of the story the child told, particularly the choice of words. Once again, action was by far the most frequently mentioned category (i.e., "The boy walked and walked," "The dragon roared at the boy," "The mother asked..." and "The dragon cried"). Forty-nine out of the 50 children (98%) referred to characters' actions; 15 of them (30%) referred to action only. To count as action, the actor needed to be explicitly mentioned. Stating simply that something happened ("tears turned into a river") was not considered an indication of a psychological reference.

A variety of other psychological references were made by the children. Twenty-six children (52%) referred to emotion ("The children were frightened, but the little boy wasn't frightened"), 21 (44%) to the motivational state of desire ("I want to

invite the dragon") and 1 (2%) to "trying" ("he's trying to wipe up his eyes"), 14 (28%) to psychological traits ("Isn't he a funny little boy?" or "The dragon said, 'How kind you are"), and 5 (10%) to perception ("He saw the dragon"). Especially important in this study is the use of epistemic states. Those included in this story are "believe," "know," "think," "surprised," "understand," "joking," "serious," and "kidding." 12 children (24%) used epistemic states in their story retelling. The variable constructed to indicate the levels of psychological reference in the Dragon's Tears story retelling is:

1 = action only

2 = action plus psychological reference

3 = action + psychological reference + epistemic state

In addition to this global assessment of psychological use, variables called "emotion word use" and "mental state use" were constructed, on a scale of 0 – 2.

0 = no such word is used in Dragon Tears retelling

1 = one such word is used (or the same one word is repeated several times)

2 = more than one different word is used

Coding was straightforward. The one decision made was to count two different words referring to the same state (e.g., "scared" and "frightened") as two emotion words. A child who said, "He was scared. He was really, really scared," for

example, was put in the "one word use" category. All variables derived from the Dragon's Tears retelling task begin with the letters "DT."

Session Three Coding

Scoring the TELD-2

Children's responses had been noted during the session on the scoring sheets provided with the test materials, as well as on audiocassette tape. If confusion arose during coding, the investigator listened to the tape to verify answers. For each child, a TELD-2 raw score as well as four subscores (receptive, expressive, syntactic, and semantic) were calculated in the way prescribed in the TELD-2 examiner's manual. [2] The TELD-2 was scored according to its instructions. Its standardized scores are reported in the Results section below.

Because the relative importance of language competence and chronological age to children's responses was an issue, the investigator wanted to have a "language group" variable that paralleled the "age group" variable. Children's language scores were divided into three groups analogous to age groups, with the goal of having each language group composed of approximately the same number of children. Equality in number was impossible to achieve because often more than one child shared a score. Seventeen children had "low" TELD-2 raw scores, between 30 and 46. Sixteen children had "medium" scores of 47 to 52, and seventeen children had "high" scores of 53 to 57.

Scoring the Raisin Box Task

Children's responses to the deceptive raisin box task questions were recorded in writing and on audiotape. Coding followed the procedure used in the "theory of mind" literature. The child, having seen the closed raisin box with raisins pictured on it and then been shown that there are crayons inside, is considered "correct" when she answers the following questions in this way: 1) To the question "Before I opened the box, what did you think was in it?" the answer is "raisins," 2) When asked "If your teacher came in right now and saw the box closed like this, what would she think was in the box?" correct responders say "raisins," and 3) To the last question, "What's really in the box?" the answer is "crayons." Children who answer incorrectly say that they thought, and the other would think, there were crayons in the box.

For this study, question 1 about the past mental state of the self, and question 2 about the mental state of the naïve other are of primary interest. So children were coded as having one of the following four patterns of answers: incorrect on both questions ("failure") right on self and wrong on other ("in transition"), right on other and wrong on self ("in transition"), and correct on both questions ("pass"). The variables derived from children's performance on the raisin box task are referred to as "RB" variables in the variables list provided in Appendix D and throughout the report of this study.

[1] No exceptions were made to this procedure, except in one of the independent preschools. There, the director of the school went to the classroom, asked the children if they were willing to participate, and escorted them to the room where the researcher was waiting to greet them.

[2] One question, though, was found to be very difficult to score reliably. Because children's scores would be compared only with the scores of others in this study, and standardized scores would not be used, the investigator decided to eliminate the hard-to-score question from the calculation of all children's scores.

RESULTS

Overview of chapter

Findings are presented in several sections. The first section is an overview of the analyses and includes the rationale for each step taken. Preliminary findings are reported. The second section details the findings of the first run of correlations; the findings of the factor analysis and the composition of task variables are presented in the third section. A fourth section presents the findings of initial regression analyses and a fifth describes group comparison results.

The group comparison results section includes subsections. Each subsection deals with the relationships between two of the three main aspects of interest in the study: person use, psychological understanding (defined here as children's ability to talk about people's actions, desires and goals, emotions, and mental states), and narrative skill. The last section of the chapter is a summary of the results.

Analyses: Order and rationale

The current study explores possible relationships among three coded aspects of children's social cognition and linguistic skill. Before analysis could begin, the data was explored to determine what kinds of analyses would be appropriate. Appendix E is the table of means, standard deviations, and range for all variables in the study. In general, data distributions are non-normal; therefore, most tests throughout the study

are non-parametric. In those instances when the assumptions of normality are met, parametric tests are performed.

First, though, testing was done for systemic bias in the data. Given the centrality of the self-other question in the study, it was essential to find out whether the different experiences of Player A and Player B in the game affected children's game recounting. Because Player A found money in the "food" box and Player B's box was empty, it was hypothesized that Player A's experience had greater salience for all children, whether for them Player A was "self" or "other." Such differential salience would have resulted in the Player A group's game recountings having significantly more references to self, both in raw numbers and in percentage of total number of person references, than did the recountings of those who were Player B. Children who were Player B would, then, have more and greater percentage of references to the experience of the "other" than did those who were Player A. Such a difference in salience might also result in differences in children's ability to answer direct questions about the experience of self and other.

To test the hypothesis of a difference in salience and recounting, the non-parametric Mann-Whitney U test was performed, comparing the mean rank of Player A group ($N = 25$) with Player B group ($N = 25$) on a number of dependent variables. Note that "GR" means that the variable is derived from the game recounting; "SQ" variables are measures of children's ability to answer standardized questions about the game. "DT" indicates "Dragon's Tears retelling" and "RB" stands for "raisin (deceptive) box task."

The variables investigated are: the number of "self" references in game recounting (GR # self), number of "other" references in game recounting (GR # other), percentage of total person references that were references to the self (GR % self), percentage of total person references that were references to the other child (GR % other), number of standardized questions about the experience of the self in the game that were answered correctly (SQ # self ?s) and the number standardized questions about the other's experience in the game that were answered correctly (SQ other ?s). The table of results follows:

Table 1. Mann-Whitney U test comparisons of GR person use and SQ questions correct by Players A and B

N=50

PLAYER		*N*	Mean Rank	Sum of Ranks	Mann-Whitney U	Asymp.Sig (2-tailed)
GR # self	A	25	27.36	684.00	266.00	.36, *ns*
	B	25	23.64	591.00		
GR # other	A	25	22.88	572.00	247.00	.20, *ns*
	B	25	28.12	703.00		
GR % self	A	25	27.02	675.50	274.50	.46, *ns*
	B	25	23.98	599.50		
GR % other	A	25	23.60	590.00	265.00	.27, *ns*
	B	25	27.40	685.00		
SQ # self ?s	A	25	27.34	683.50	266.50	.36, *ns*
	B	25	23.66	591.50		
SQ # other ?s	A	25	25.96	649.00	301.00	.81, *ns*
	B	25	25.04	626.00		

No significant differences in the mention of self, other, or shared experience in the game recounting nor in ability to answer direct questions about the self and other were found between those who were Player A and those who were Player B.

Also of interest was whether the order in which standard questions were posed to children (i.e., either "self" questions first or "other" questions first) made a difference in their ability to answer those questions. Did children demonstrate practice effects, so that they were better at answering the second set of questions, regardless of whether it was "self" or the other child that was the subject of those questions? Did children get tired, and find it more difficult to answer the second set correctly? Again, group comparisons using the Mann-Whitney U test to compare the "self questions first" group with the "other questions first" group revealed no significant differences in the groups' number of correctly answered questions about the experience of the self or of the other.

Table 2. Mann-Whitney U test comparisons of SQ performance by SQ question order groups

N = 50

SQ Order Group		*N*	Mean Rank	Sum of Ranks	Mann-Whitney U	Asymp.Sig (2-tailed)
SQ # self correct	self first	24	22.81	547.50	247.50	.20, *ns*
	other first	26	27.98	727.50		
SQ # other correct	self first	24	25.50	612.00	312.00	1.00, *ns*
	other first	26	25.50	663.00		

A third difference among children was also considered: Would boys' (*n* = 26; mean age = 51 months, *SD* = 5.56) and girls' (*n* = 24, mean age = 51.46 months, *SD* = 5.75) responses differ from each other in systematic ways? Because the groups are even in number and scores normally distributed, analyses of variance were run, comparing the groups on TELD-2 raw and subscores, factor scores, and the task-related composite variables constructed as described below. Girls outperformed boys

on: TELD-2 expressive subscore (F (1, 48) = 1.70, $p < .05$), TELD-2 syntactic subscore (F (1, 48) = 8.59, $p < .01$), Factor 3 (Self-Other answers) (F (1, 48) = 5.34, $p < .05$), and SO composite score (F (1, 48) = 5.35, $p < .05$). TELD raw score did not reach significance (F (1, 48) = 3.35, $p < .07$).

The table below shows the results of the comparison of boys and girls. The nonparametric Mann-Whitney U test was used because the variables included here cannot be assumed to have normal distributions.

Table 3. Group comparisons (girls vs. boys) using Mann-Whitney U test

* $p < .05$

	sex	*N*	Mean Rank	Mann-Whitney U	Asymp. Sig. (2-tailed)
GR # self references	female	24	26.69	283.50	.58, *ns*
	male	26	24.40		
GR # other refs	female	24	23.69	268.50	.39, *ns*
	male	26	27.17		
GR # 'we' refs	female	24	27.69	259.50	.30, *ns*
	male	26	23.48		
GR % self refs	female	24	25.33	308.00	.94, *ns*
	male	26	25.64		
GR % other refs	female	24	22.63	243.00	.18, *ns*
	male	26	28.15		
GR % 'we' refs	female	24	27.79	257.00	.28, *ns*
	male	26	23.38		
GR psych content	female	24	27.00	276.00	.42, *ns*
	male	26	24.12		
DT psych content	female	24	26.90	278.50	.49, *ns*
	male	26	24.21		
DT emotion word use	female	24	21.42	214.00	.04*
	male	26	29.27		
DT mental state use	female	24	28.35	243.50	.08, *ns*
	male	26	22.87		
SQ # self correct	female	24	27.58	262.00	.32, *ns*
	male	26	23.58		

SQ # other correct	female male	24 26	30.33 21.04	196.00	.02*
RB deceptive box task performance	female male	24 26	29.67 21.65	212.00	.04*
RB self think	female male	24 26	28.13 23.08	249.00	.16, *ns*
RB other think	female male	24 26	30.13 21.23	201.00	.01*
SQ deceptive box Task performance	female male	24 26	28.19 23.02	247.50	.16, *ns*

On one variable, boys scored higher than girls did; they used more emotion state words in their retelling of the Dragon's Tears story than girls did. On the rest of the variables on which boys' and girls' performance differed significantly, it was the girls who performed at a higher level. Girls correctly answered significantly more of the standardized questions posed to them about the game playing experience of their partner and were more likely to pass the Raisin Box task.

The fact that girls' performance on the Raisin Box task was significantly better than boys' prompted further exploration of person understanding in that context. The variable score, 0 – 2, is based upon children's ability to answer these two questions correctly. "Before I opened the box [with raisins pictured on it, showing you that there are crayons inside], when it was closed like this, what did you think was in the box?" is the so-called "RB self think" question. The second, called the "RB other think" question, is "If your teacher came in right now and saw the box closed, just like this, what would she think is in the box?" Mann-Whitney U tests comparing boys and girls on their ability to correctly answer each of those questions reveals that the difference between the groups lies in their answers to the question about what the

teacher (i.e., the "other" in this test) would think was in the box. Girls were significantly more likely to answer that question correctly, than were boys. There is no significant difference in boys' and girls' ability to correctly answer the "self" question.

The reason for boys' greater use of emotion words in their Dragon's Tears retelling is not readily apparent from the correlational data in this study. We do know that the variable is not significantly correlated, for example, with language skill, but further investigation of emotion use across tasks would be required to investigate the cause of this difference, and to see whether this finding can be generalized.

There is, however, a discernable pattern in the girls' better performance. All the differences seen here can be explained by girls' higher language scores. Regression analysis, entering the four TELD subscores as possible predictors of performance on answers to direct questions (both SO composite and SO factor scores), reveals that it is expressive language skill that most highly predicts children's performance. That is an aspect of language on which the girls in this study receive significantly higher scores than the boys. Also, factor analysis indicates that knowing the number of correct answers a child gave to questions about the "other" is a more discriminating marker of verbal maturity than is the number of correct answers given to questions about the self (details are given in the discussion of self-other competence). In other words, if the "gender" differences on these variables were masking or standing in for a difference in language competence, the above pattern of differences is exactly what we would expect to see. So gender per se will not be discussed further.

Having established that these sources of possible bias were not affecting the data, analyses were performed on the entire sample of children ($N = 50$). The first step in analysis was to evaluate the degree of association between variables derived from children's performances on four tasks: game recounting (GR), answers to standardized questions about the game (SQ), Dragon's Tears story retelling (DT), and the deceptive box task (the "raisin box [RB]" task). The non-parametric Spearman's rho correlations between all variables included in this report are presented in Appendix F (table 1). Because the factor analysis to follow stems from analyses performed on Pearson product moment correlation matrices, those parametric correlations were also performed, and are presented in Appendix F (table 2). Unless otherwise noted, the correlations reported here retained significance across the non-parametric and parametric tests. The Pearson product moment correlations will be reported in section 2 (below) because of their further use in the factor analysis described below.

While the correlations provided immediately interesting information relevant to the main questions of the study (see section 2 below), it was also clear that the large number of significant correlations presented a problem. Reporting each, one by one, would be unwieldy and relatively uninformative, since each variable entered represented just one isolated aspect of children's performance. To reduce the number of variables in a statistically meaningful way, a series of factor analyses were carried out.

Factor analysis offered several advantages. One is that it clusters the large number of variables into a smaller number of statistically related groups, thus showing

relationships among three or more variables. Researchers then have a statistical rather than simply intuitive or conceptual rationale for the formation of composite variables if they are needed.

A second advantage is that the "factor" around which the related variables cluster is considered to be causal, though hypothetical. The assumption of factor analysis is that the variables cluster because they have the same causal foundation. While it is up to the researcher to name the factors (i.e., to hypothesize what the underlying cause of the given interrelationships might be), factor analysis allows us to move beyond correlations to raise useful questions about causes of the behavior we observe.

Third, factor analysis provides a factor correlation matrix that answers the further question of whether the hypothesized underlying causal factors have any statistical relationship to one another. This is especially appropriate in a study like the current one that seeks to determine whether there are relationships among a number of different aspects of cognition.

After factor analysis was completed and found in general to cluster variables according to task, new composite variables were formed (see below) to summarize children's performance. Please note that "GR composite" is the composite score for the game recounting task; "DT composite" is the score for children's Dragon's Tears story retelling, and "SO composite" is the score for children's skill in answering direct questions about the self and other.

Composite variables were used in several ways. First, multiple regression analyses were performed using the composite scores. Entering age and language as possible predictor variables tested their relative impact on children's composite scores

and other target variables, and helped to determine the shared and unshared foundations of the target variables.

Group comparisons were also an important source of information in this study. While correlations made general relationships between variables apparent, group comparisons made possible a more explicit look at developmental change. Constructing and comparing groups of those who responded in different ways to a particular task (those who failed, those who are 'in transition' and those who passed the RB task, for example) yielded two observations. First, when such groups were compared on variables known to mark biological or cognitive change (e.g., age and language skill) it was possible to see which level of behavioral change tended to signal broader changes. Was it those in transition on the RB task who differed significantly in language skill from those who passed, or was language development more apparent when comparing those who failed with those who passed? Do those who refer to the other child when recounting the game playing experience differ significantly, in age or language, from those who did not mention the 'other'?

Secondly, comparing groups provided another opportunity to examine the relations between and among person use, psychological understanding, and narrative skill. Very specific questions can be asked: Do people who refer to self and other have significantly higher narrative skill scores than those who refer only to "us"? Do those who have low narrative skill scores perform differently on the RB task than those whose score is high? Group comparisons are a particularly rich source of data to help answer the questions posed in this study.

Membership in the groups (e.g., referrers to self only vs. those who referred to two or more persons in GR) was, in most cases, determined a posteriori, on the basis of children's behavior in the study. Given the inability to plan for equal membership in each level of a group, the number of children in each group and the distributions of their scores, vary. This is a second reason for the use of the non-parametric Mann-Whitney U and Kruskal-Wallis tests. Group membership and levels are reported at the beginning of each subsection of the Group Comparisons section, V, below.

Results of Preliminary Correlations

The first step in analysis was to perform correlational analyses on all pairs of numerical and ordinal variables (see Appendix E for descriptives) derived from the four main tasks in the study (GR, SQ, RB, and DT). An initial look at the table of Pearson product moment correlations produced (see Appendix F, Table 2) revealed a number of important results.

First, it was apparent that the variables describing children's demonstrated understanding of "self" and "other" were correlated. The Pearson product moment correlation coefficient of raw number of self-references and raw number of other references in the game recounting was .29, $p < .04$. The psychological content of the self-references and "other" references were likewise significantly correlated: $r = .47$, $p < .01$, and each, not surprisingly, was correlated with the number of person references made. More talk about the experiences of persons yielded more psychological references, and there was a tendency for children to talk about "self"

and "other" in parallel terms ("I wanted the parrot; he wanted the octopus" or "She looked in that box [pointing] but I looked in this box [touching the box]."

Table 4. Pearson product moment correlations of game recounting psychological content level and number of references
*p < .05
** p < .01

	GR # of 'self' references	GR # of 'other' references
GR 'self' psychological content level	.55**	.35*
GR 'other' psychological content level	.38**	.69**

The correlation between performances on "self" and "other" was found, as well, when considering the number of correctly answered questions. Children's correct answers to self-related questions and to those posed about their game-playing partner were correlated at Pearson r = .48 (p < .001). The psychological understanding demonstrated in the answers to the standardized questions about "self" and "other" was also significantly correlated: Pearson r =.53, p < .001. Individual question answers were also correlated. The question "How did you feel when you saw what was in the box?" if answered correctly tended also to be answered correctly when asked about the other child (Pearson r = .52**, p < .01). Because the two answers to be correct, needed to be different ("good" or "happy" for Partner A and "bad" or "sad" for Partner B), the significant correlation is especially telling.

In their relationships to other variables, however, "self" and "other" reference were distinct. It was not "'self" and "other" references, but those made to "we" and "us" that have significant relationships to the length of the game recounting:

Table 5. Pearson product moment correlations of game recounting number of acts recounted and number of person references

* $p < .05$
** $p < .01$

	# of self references	# of other references	# "we/us" references
GR # of acts recounted	.35**	.25	.53**
GR # in-sequence acts	.31*	.25	.59**

The number of references to "us" or "we" in the game recounting bears some unique relationships to other variables of interest, not shared by number of references to "self" and "other". The raw number of references to the joint activity and experience of both partners correlates significantly with: percentage of on-task utterances in game recounting (Pearson $r = .32, p < .05$), 'our' psychological content level (Pearson $r = .59, p < .001$), number of acts retold in the Dragon's Tears story (Pearson $r = .41$, $p < .01$), and Dragon's Tears variety of connectives used (Pearson $r = .35, p < .05$).

Considering the use of "self", other, and "we/us" as a percentage of total person use proved to be an important, second way to look at the issue of person use. Since the three numbers (percentage of "self", percentage of other, and percentage of "us/we") are orthogonal, relationships among them are negative and correlations are obviously highly significant. What is interesting, however, is the finding that more

exclusive reference to the "self" in the game recounting correlates negatively with the ability to respond correctly to direct questions about "self" and "other."

Table 6. Pearson product moment correlations

RB task: 0 = both self and other 'think' questions wrong; 1 = 'self' or 'other' correct; 2 = both persons correct

$N = 50$

* $p < .05$

** $p < .01$

	# correct on SQ self	# correct SQ other	Raisin Box Task
% of SELF reference	-.44 **	-.30 *	-.32*
% of OTHER reference	.24	.16	.07
% of "WE/US" reference	.25	.17	.25

The percentage of "other" reference and reference to the pair do not predict children's ability to respond to direct questions in this study.

In fact, the percentage of "other" reference in a child's game recounting does not correlate significantly with any variables, except the variable that describes the psychological content level of children's description of the "other." As the percentage of "other" reference increased, the likelihood that children would mention not just his game playing partner's action, but also such things as his partner's goals and desires, increased. It is again the percentages of reference to self and to 'us' that tell the striking, and opposite stories.

Table 7. Pearson product moment correlations r for those variables significantly related either to percentage of "I/me" or "we/us" reference

$N = 50$
$^{*}p < .05$
$^{**}p < .01$

	GR % of person use = reference to the SELF	GR % of person use = reference to 'US/WE'
TELD-2 raw score	-.43**	.34*
GR acts recounted (of 9)	-.39*	.27
GR # of in-sequence acts	-.30*	.28*
GR variety of connectives	-.42**	.27
DT acts retold (of 27)	-.20	.35*
SQ # 'self'-related questions right	-.44**	.25
SQ # 'other'-related questions right	-.30*	.17
RB & SQ 'self' think question right	-.37**	.33*

Given the current debate about children's "theory of mind" and their ability to think and talk about mental states, it is notable that some children made reference to mental states in their Dragon's Tears retelling: "[In response to the mother's asking her son whether he is really sure he wants to ask the dragon to his birthday party] 'No, I'm not kidding,' said the boy, 'I'm serious'"; "[After first getting angry at the boy's invitation] the dragon understood [that the invitation was meant to be kind]." Such uses of mental state terms in the story retelling correlate significantly with: DT number of acts retold (Pearson $r = .66$, $p < .001$); DT % of subject-verb agreement (Pearson $r = .59$, $p < .001$), DT Guttman scale of variety of connectives (Pearson $r = .60$, $p < .001$); correct answers to direct questions about the experience of the self in

the game playing session (r =.30, $p < .05$); correct answers to direct questions about the experience of the other in the game (r = .37, $p < .01$). A striking lack of significance is found in the correlation between children's mental state use in the story and their performance on the standard deceptive box task (in this study, RB = raisin box) test of false belief understanding (Pearson $r = .21, p = .14$, *ns*).

Some within-task correlations were obvious and expected. Examples are those between numbers of acts and variety of connectives used, both in Game Recounting (Pearson $r = .60, p < .001$) and Dragon's Tears retelling (Pearson $r = .64, p < .001$) as well as those between psychological content and numbers of person references in GR. Others, though not surprising, were not entirely self-evident. Dragon's Tears total number of utterances, for example, correlates with the percentage of subject-verb agreement (Pearson r = .44, $p < .01$). Longer stories were more grammatically correct stories.

The initial correlations reflected both expected and unexpected between-task relationships. As expected, the number of acts reported in the game retelling and in the Dragon's Tears retelling were significantly correlated: r = .40, $p < .01$. Other aspects of tasks, thought to be measures of the same competence, proved not to be correlated significantly. Autobiographical references in game retelling (percentage of verbs that were correctly put in the past tense, and attributed to a specific person or subject) were thought to parallel the percentage of subject-verb agreement in the Dragon's Tears story, though admittedly the latter was more complex in its tense requirements. That relationship did not reach significance (Pearson $r = .26, p = .067$). Unexpectedly, it was the number of acts reported in the game recounting that

correlated significantly with subject-verb agreement in the Dragon's Tears retelling (r = .31, p < .05). The data indicated that a priori assumptions about relationships among aspects of the tasks could not be relied on. A further exploration of the relationships among variables was indicated.

Results of factor analysis

The first step in factor analysis was to identify a group of variables within the entire set that could be introduced into the factor analysis. Ideally, this group would include variables relevant to each of the major aspects (person use, psychological, especially mental state, understanding, and narrative skill) under study here, so that their interrelationships could be observed.

Second, the group of chosen variables needed to have the mathematical properties necessary to perform the analysis. For example, overlapping or redundant variables had to be eliminated from the list as highly correlated variables invalidate the analysis. In this set of data the three "percentage of total number of person references that refer to self [other, and 'us']" variables were excluded because they are, by definition, orthogonal. To explore GR person use, the raw numbers of references were included.

Since the relative importance of age and language are of interest, those variables were introduced into the analysis. The introduction of children's age in months and TELD raw scores, perhaps because of their correlation, made the analysis impossible.

Therefore those two variables were reconstructed in terms of age groups and language groups (see methods section). These groups were then entered into the factor analysis successfully.

Several attempts to carry out the factor analysis with different variable composition were made. In addition to language and age group, other variables were included in the factor analysis solution that is reported here. The variables entered, their levels, range, means and standard deviations are shown in Table 8 (see next two pages).

Table 8. Descriptives of nineteen variables in factor analysis solution

GR = game recounting variable; DT = Dragon's Tears retelling variable
SQ = standardized questions about game; RB = Raisin (deceptive) Box task

Variable	Levels	*N*	Minimum	Maximum	Mean	*SD*
AGE GROUP	1 = 3.6 - 4 2 = 4 - 4.6 3 = 4.6 - 5	50	1	3	2.06	.82
LANGUAGE GROUP	1 = low 2 = medium 3 = high	50	1	3	2.00	.83
DT # acts retold		50	0	20	7.86	4.77
DT % subj-vb agreement		50	0	98	47.34	31.35
DT mental state use	0= none used 1 = one used 2 = 2 or more	50	0	2	.32	.59
DT variety of connectives	0 = none 1 = and 2 = 1 + then 3 = 2 + so/but/ because 4 = 3 + when/ until 5 = 4 + where	50	0	5	1.98	1.35
DT emotion state words	0 = none 1 = used one 2 = 2 or more	50	0	2	.72	.73
GR # main acts recounted		50	1	8	4.68	1.79
GR # in-sequence actions		50	1	10	3.92	2.29

Variable, continued	Levels, continued	*N*	Minimum	Maximum	Mean	*SD*
GR variety of connectives	0 = none 1 = and 2 = 1 + then 3 = 2 + so/but/ because 4 = 3 + when/ until	50	0	4	1.86	1.11
GR # self references		50	0	12	3.04	2.59
GR # other references		50	0	10	2.24	2.35
GR # references to 'we/us'		50	0	17	3.58	3.63
GR % on-task utterances		50	27	100	74.76	19.46
SQ & RB 'other think	0 = 0 right 1 = 1 right 2 = both right	50	0	2	.70	.81
SQ 'self think'	0 = wrong 1 = right	50	0	1	.28	.45
RB 'self think'	0 = wrong 1 = right	50	0	1	.52	.50
SQ # self ?s correct		50	2	9	6.58	1.46
SQ # other ?s correct		50	2	8	6.12	1.29

The following factor analysis solution was chosen, not only because the variables entered worked to provide a solution, but also because it: 1) included the maximum number and variety of variables of interest, 2) "passed" the Goodness-of-fit Test (chi square = 100.27 (101), p = .50), 3) met the accepted standard of providing at least three factors, all of which have at least three variables that loaded on the factor at a level of .3 and above, and 4) made sufficient conceptual sense that names could be given to the factors. The initial solution provided five factors, but only the first four

made enough conceptual sense to be named. Four solutions were then extracted, resulting in the solution reported here.

The factor analysis structure matrix (see table 9 below) shows the variables that loaded on each factor and the correlations of each with the factor on which it loads. Negative correlations in one factor (factor 2 for instance) indicate a negative correlation to the variables included in the prior factor (1, in this case). Only those correlations that are considered high enough to be useful for further analysis (i.e., those that are .3 or higher) are included in the table here; a complete structure matrix table is provided in Appendix G.

Table 9. Structure Matrix for Factor Analysis Solution

Extraction Method: Maximum Likelihood
Rotation Method: Oblimin with Kaiser Normalization

VARIABLES	FACTOR			
	1 = "DT factor"	2 = "IM factor"	3 = "SO factor"	4 = "GR factor"
DT # of acts	.889	-.434		
DT % subj-vb agree	.818		.307	
DT mental state use	.754	-.323	.302	
DT connectives	.678	-.325		
DT emotion states	.515			
GR # in-seq acts		-.840		.521
LANGUAGE GROUP		-.676	.531	
GR # acts recounted		-.649	.360	.632
GR # we references		-.644		
GR % on-task utts		-.524	.460	
AGE GROUP	.343	-.431		
SQ & RB 'other think'			.856	
SQ # other correct			.707	
RB self think			.703	
SQ # self correct		-.393	.624	
GR variety of connectives		-.491		.604
GR # self references				.546
SQ self think				-.452
GR # other refs				.426

The most notable feature of this solution is the interrelations among those variables that are derived from children's performance on the same task. The factors are

marked by a lack of common variables. Just two of the nineteen variables entered (DT % subject-verb agreement and DT mental state use) load on more than one of the task-related factors (factors 1, 3 and 4), and their loading on the second factor (in this case, factor 3, the SO factor) is so low as to just rate mention. Each of the task-related factors has a cluster of same-task variables that relate most highly to its hypothesized cause.

The distinctness of each of the three task-related factors is further indicated by the unique relationship it has to the overarching variables of "language group" and "age group." Age group alone loads on factor 1 (DT factor). Language group relates positively only to factor 3 (SO factor), and neither age nor language relate at the .3 level to factor 4 (GR factor). A third indicator of the independence of task-related factors 1, 3 and 5 from one another is seen in the factor correlations.

Table 10. Factor Correlation Matrix

FACTOR	2(IM)	3 (SO)	4 (GR)
1 (DT)	-.241	.196	0
2 (IM)		-.375	-.215
3 (SO)			0

Again, the guideline is that only those factors that are correlated at .3 level or above are reported. Among the four, only factors two and three are correlated to a degree that is worth noting: -.38. No two task-related factors relate significantly to each other.

Because of their task-specificity, and lack of overlap, it seemed legitimate to give factors 1, 3 and 4 the names of the tasks they represented: Dragon's Tears retelling (DT), self-other question answers (SO), and game recounting (GR). Factor 2 was named the "immaturity factor" (IM) for two reasons. One is that it is negatively related to children's ability to do such things as stay on task in GR, put a number of actions in sequence, use a variety of connectives and tell the story of what "we" did when we played the game. Secondly, unlike the other factors that at most, included either age or language, factor 2 included both age and language. Given those facts, it seemed reasonable to hypothesize that the cause of the correlations among factor 2 variables is a child's general immaturity level.

Next, composite variables were constructed to summarize each child's performance on the variables that were most central to each of the factors. In other words, four composite scores were constructed: DT composite (from factor 1), IM composite (from factor 2), SO composite (from factor 3), and GR composite (from factor 4).

The task-specific composites were derived in several steps. First, all variables, except language and age, that loaded on the factor at a .5 level or higher, and related to the specific task which was the focus of that factor, were included. Second, all scores that were not already on a 0 – 100 scale were converted to that scale in order to give each variable equal weight in the composite; for example, the number of DT acts retold ranged from 0 to 20. The number each child used was multiplied by 5 to get a normalized score, scaled 0 to 100. Third, the normalized scores for each included variable were summed and the sum was divided by the number of variables included.

A composite score from 0 to 100 for each child was the result. The following table lists the variables included in the three task-related composite scores:

Table 11. Composition of task-related composite scores as drawn from factor analysis.

Dragon's Tears composite (DT composite)	Game Recounting composite (GR composite)	Self-Other Answer composite (SO composite)
DT # of acts retold	GR # of acts recounted	**SQ & RB other**
DT mental state use	GR # in-sequence acts	**SQ # self ?s correct**
DT % of subject-verb agreement	GR variety of connectives	**SQ # other ?s correct**
DT variety of connectives	**GR # self references**	**RB self think**
DT emotion state use	**GR # other references**	

The immaturity composite score was, like those above, composed of those variables that loaded on the immaturity factor at an $r = .5$ level or higher. Using that criterion, all four variables were derived from children's performance on the GR task. Also, for this score the numbers needed to be such that high scorers, since their performances were "mature," had to have low "immaturity scores." After the scores were normalized and then summed (as above), each child's total score was subtracted from 100 in order to get an "immaturity" score from the variables listed below.

Table 12. Composition of Immaturity composite score, as drawn from factor analysis

Immaturity Composite Score (IM composite)
GR # of in-sequence acts
GR # of main acts recounted
GR # of references to "we/ us"
GR % of on-task utterances

A narrative skill composite score was constructed by eliminating content-related variables from the DT composite and the GR composite scores, leaving only those variables which related to the child's ability to structure their telling of the story. Three variables remained for each task/factor (see table below).

Table 13. Narrative skill composite score variables

Narrative Skill Composite (NA composite)
DT # of acts retold
DT % of subject-verb agreement
DT variety of connectives
GR # of acts recounted
GR # of in-sequence acts
GR variety of connectives

To test the legitimacy of simply adding the six together, first the three standardized variables for DT were summed and divided by three to get scores on a 0-100 scale;

the same was done with the GR scores. A Pearson product moment correlation was then performed on the new "DT narrative score" and the "GR narrative score" to determine if the correlation was such that the two could be summed to make a narrative composite score. The relationship was statistically significant ($r = .40$, $p < .01$), so the two scores were summed and divided by 2 to get a narrative (NA) composite score that, like the other composite scores, is on a 0 – 100 scale.

The derived composite scores were used to extract more information from the data in three ways. The first was through the performance of further correlational analysis. The initial analysis was of the relationships among the composite scores themselves to identity cross-task relationships that were not apparent in the factor correlation matrix. Correlations are reported in table 14 (below) only if the two composite scores have no variables in common.

Table 14. Pearson Product Moment correlations of composite scores

* $p < .05$

	IM	SO	GR	NA
DT	-.237	.254	.204	
IM		-.331*		
SO			.194	.318*

The negative correlation between the immaturity IM) composite and the self-other answer (SO) composite parallels the significance of the correlation between the IM factor and the SO factor. New information is provided here, though, about the

relationship between children's ability to answer direct questions and their narrative skill as demonstrated in the GR and DT tasks. Children answering more questions correctly tend to put more actions in sequence, use a variety of connectives, and produce DT sentences in which subject and verb agree. Composite scores proved useful in summarizing both tasks and skills.

The second way the composite scores were used was as target variables of interest in multiple regression analyses, entering both age and language to see whether each or both predicted the particular composite scores. Those analyses, with variables entered stepwise, were performed, and are reported in the next section. The third use of the composite variables was in group comparisons. Those findings will be detailed in the group comparison section below.

Results of Regression Analyses

The inclusion of language group and age group in the factor analysis provided useful information about the relationship of language and age to the tasks children were asked to carry out in this study. However, as noted above, the more precise and informative variables of age in months and TELD-2 raw scores could not be included there. Further information is made possible through the use of stepwise multiple regression analysis, entering age in months and TELD raw scores as the possible predictor variables and each composite score in turn as the target variable. The significant relationship between age in months and TELD raw scores (Pearson product moment $r = .58$, $p = .000$) means that only one will be chosen as the predictor

of the composite scores. The multiple regression analyses yielded the following results:

Table 15. Multiple regression analysis results for relationship of composite scores to age in months and TELD-2 raw scores

	Age in months			TELD raw scores		
	B	*SE B*	*β*	*B*	*SE B*	*β*
DT composite	1.932	.525	.469**			
IM composite				-1.256	.315	-.499***
SO composite				1.914	.436	.535***
GR composite				.818	.338	.330*
NA composite				1.203	.359	.435**

DT = Dragon's Tears; GR = Game Recounting
SO = answers to self and other questions; IM = immaturity
Blank cells = excluded variable
* = $p < .05$, ** = $p < .01$. *** = $p < .001$

A further question was then raised: Which particular TELD subscores, from among those of expressive, receptive, syntactic, and semantic, best predict the composite scores? The TELD subscores are interrelated (in all cases, $p < .01$) because they are derived from the same test and are all necessary components of language skill, so here again only one subscore will be entered as a predictor variable into the equation. The analysis will determine which of the four is most predictive of the composite score of interest.

Regression analysis using the stepwise method was performed, entering the four TELD-2 subscores as possible predictor. The following results were obtained.

Table 16. **Multiple regression analysis results entering four TELD-2 subscores as possible predictors of composite scores**

Target Variable	Best Predictor	*B*	*SE B*	β
DT composite	syntactic score	5.019	1.985	.343*
IM composite	semantic score	-1.55	.412	.477***
SO composite	expressive score	4.402	.990	.540***
GR composite	receptive score	1.346	.623	.298*
NA composite	receptive score	2.137	.658	.424**

Because specific variables within the composite scores might relate differently to age and language, further regression analyses of each variable were performed, entering age and language as the possible predictor variables. The table of results follows.

Table 17. Stepwise multiple regression analysis results entering age in months and TELD raw scores as possible predictors of variables that compose the three task-related composite scores

	Age in months			TELD raw scores		
	B	SE B	β	B	SE B	β
DT # acts	.351	.112	.413**			
DT mental state	0	.014	.436**			
DT % s-v agree	2.39	.729	.427**			

DT connectives	-- neither predicts --					
DT emotion use	0	.018	.323			
GR # acts				0	.035	.377**
GR # in-seq act				.138	.044	.412**
GR connectives				0	.022	.287*
GR # self refs	-- neither predicts --					
GR # other ref	-- neither predicts --					
GR # "we" ref				.161	.073	.302*
SQ + RB other				0	.015	.438**
SQ # self ?s				.125	.024	.602***
SQ # other ?s				.109	.024	.555***
RB self think				0	.010	.366**

Group Comparisons

Age and Language Groups on Task Competence

Given our interest in the role of task requirements on children's verbal expression of psychological understanding, it is important to determine the relative challenge that the tasks pose to children in this study. One way to assess that is to compare age and language groups on their performance on the three main tasks of the study. Because the composite scores are normally distributed and the groups nearly equal in number, analyses of variance are appropriately used for these comparisons.

Table 18. Analysis of Variance for Age Groups on Composite Scores

$^{*}p < .05$

		Sum of Squares	df	Mean Square	*F*	Sig.
DT composite	Between groups	3209.103	2	1604.552	3.288	.046*
	Within groups	22934.344	47	487.965		
	Total	26143.447	49			
SO composite	Between groups	964.030	2	482.015	.739	.482
	Within groups	30658.453	47	652.308		
	Total	31622.482	49			
GR composite	Between groups	356.991	2	178.495	.612	.547
	Within groups	13711.028	47	291.724		
	Total	14068.019	49			

Scheffe post hoc test determined that the significant difference among the groups on DT composite was between group 1 (3.6 to 4 year olds) and group 3 (4.6 to 5 year olds) group. Group 1's mean score (M=25.44) was 19.8 points lower than that of Group 3 (M=45.20). Group 2's mean score was 35.09. Age did not relate significantly to children's SO and GR composite scores.

Self-Other Answer composite scores differed significantly from each other when language groups' (high, medium, and low) mean scores were compared.

Table 19. Analysis of Variance for Language Groups on Composite Scores

** $p < .01$

		Sum of Squares	df	Mean Square	*F*	Sig.
DT composite	Between groups	1975.151	2	987.575	1.921	.158
	Within groups	24168.296	47	514.219		
	Total	26143.447	49			
SO composite	Between groups	6644.834	2	3322.417	6.252	.004**
	Within groups	24977.648	47	51.439		
	Total	31622.482	49			
GR composite	Between groups	810.048	2	405.024	1.436	.248
	Within groups	13257.971	47	282.084		
	Total		49			

High language scorers had SO composite scores that were an average of 27.92 points higher than those of low language scores. High scorers mean was 72.65 while low scorers' was 44.74. Medium language scorers had a mean SO composite score of 60.07.

The oldest children did best on the DT retelling; the most skilled language users were most able to answer the direct questions. The GR task, though, showed a distinct pattern that did not relate either to age or language. The comparison across age groups showed a slight, but not significant increase in performance in the expected direction. The youngest children's mean GR composite score was 35.61. Group 2 averaged 37.17 points, and the oldest group had a mean score of 41.85. The range of scores was considerably narrower than were the DT and SO ranges. When compared across language groups, the differences were again, not significant, and in this case, were highest for the medium language scores. Low TELD-2 scorers had a mean score of 33.03, *SD* = 19.52; medium scorers' mean was 42.66, *SD* = 13.98.

High scorers had an average of 39.72, *SD* = 16.26. The range here was again narrow. Children's performance on this task did not vary as dramatically across this study's age and language skill range.

Person Use group results

The groups of interest in this study are those that reflect children's range of performance on the three aspects of interest in the study: person use, psychological, especially mental state understanding, and narrative skill. The levels and composition of the person use groups are presented in the following table.

Table 20. Composition of person use groups as determined by children's reference to persons in the Game Recounting task

GROUP VARIABLE NAME	LEVELS OF GROUP	*n*	% of *N*
# OF PERSONS mentioned	1 = one person	8	16
	2 = two persons	13	26
	3 = three persons	29	58
'SELF ONLY' REFERENCE groups	0 = not self only	46	92
	1 = self only	4	8
'OTHER' REFERENCE groups	0 = no reference	14	28
	1 = refers to other child	36	72
'WE/US' REFERENCE groups	0 = no reference	9	18
	1 = refers to we/us	41	82

Tables 21 and 22 (below) report the means and standard deviations for the performances of each person group above on important variables of interest in the study. Means and standard deviations are listed for the purpose of comparison with other groups; however, when tests are carried out, they will be nonparametric because of the uneven group sizes.

Table 21. Means and standard deviations of 1-person, 2-person and 3-person reference groups

	1 PERSON REFERENCE *n*=8	*2 PERSON REFERENCE* *n*=13	*3 PERSON REFERENCE* *n*=29
Age in Mos. (42-60)	**49.15** *SD* = 6.25	**52.29** *SD* = 6.09	**51.37** *SD* = 5.25
TELD score (30-57)	**41.25** *SD* = 9.35	**50.46** *SD* =5.77	**49.83** *SD*=5.19
RB task score 0=fail;1=transition 2=pass	**.25** *SD* = .46	**1.31** *SD*= .95	**1.00** *SD* = .93
SQ task score 0=fail; 1=transition; 2=pass	**.25** *SD* = .46	**.77** *SD* = .83	**.45** *SD* = .69
GR # on-task utt. (4-29)	**8.63** *SD* = 5.48	**12.15** *SD* = 5.54	**14.76** *SD* = 5.50
GR # in-seq acts (1-10)	**2.38** *SD* = 1.85	**3.00** *SD* = 2.00	**4.76** *SD* = 2.20
GR # self refs (0-12)	**.50** *SD* = .53	**3.08** *SD* = 3.20	**3.72** *SD* = 2.23
GR # other refs (0-10)	**.00** *SD* = .00	**1.77** *SD* = 2.62	**3.07** *SD* = 2.12
GR # ref shared (0-17)	**2.75** *SD* = 4.03	**1.85** *SD* = 2.23	**4.59** *SD* = 3.78
GR psych level (1-2)	**1.50** *SD* = .53	**1.38** *SD* = .51	**1.45** *SD* = .51
DT emotion words (0-2)	**.63** *SD* = .74	**.85** *SD* = .90	**.69** *SD* = .66
DT mental states (0-2)	**.13** *SD* = .35	**.69** *SD* = .75	**.21** *SD* = .49

DT psych level (1-3)	**1.88** *SD* = .64	**2.31** *SD* = .85	**1.83** *SD* = .71
SQ # self correct (2-9)	**4.75** *SD* = 1.83	**7.08** *SD* = 1.44	**6.86** *SD* = .92
SQ # other correct (2-8)	**5.00** *SD* = 1.85	**6.85** *SD* = .99	**6.10** *SD* = 1.01
SQ self psych level (1-3)	**2.00** *SD* = .93	**2.23** *SD* = .73	**1.86** *SD* = .74
SQ other psych level (1-3)	**1.63** *SD* = .74	**2.08** *SD* = .86	**1.90** *SD* = .77
DT composite (0-100)	**25.40** *SD* = 21.51	**44.53** *SD* = 28.43	**34.82** *SD* = 20.10
Immaturity Composite (0-100)	**67.31** *SD* = 22.55	**56.80** *SD* = 10.50	**44.98** *SD* = 14.66
SO composite (0-100)	**35.94** *SD* = 17.62	**69.62** *SD* = 22.77	**60.09** *SD* = 23.20
GR Composite (0-100)	**18.40** *SD* = 13.23	**31.97** *SD* = 14.55	**46.78** *SD* = 12.72
Narrative composite (0-100)	**29.47** *SD* = 21.95	**43.51** *SD* = 20.04	**50.08** *SD* = 15.30

Table 22. Means and standard deviations of "self only," "other" and "we" reference groups

	"SELF ONLY" REFERENCE ***N*** **= 4**	**"OTHER" REFERENCE** ***N*** **= 36**	**"WE/ US" REFERENCE** ***N*** **= 41**
Age in Mos. (42-60)	**43.85** *SD* = 1.43	**51.28** *SD* = 5.20	**51.89** *SD* = 5.17
TELD score (30-57)	**34.50** *SD* = 4.65	**49.28** *SD* = 5.20	**50.17** *SD* = 5.50
RB task score 0=fail; 1=transition; 2=pass	**.00** *SD* = .00	**1.08** *SD* = .94	**1.02** *SD* = .91
SQ task score 0=fail 1=transition; 2=pass	**.00** *SD* = .00	**.47** *SD* = .70	**.56** *SD* = .71
GR # on-task utt. (4-29)	**6.50** *SD* = 2.08	**14.75** *SD* = 5.48	**13.27** *SD* = 5.71

GR # in-seq acts (1-10)	**1.00** *SD* = .00	**4.28** *SD* = 2.22	**4.39** *SD* = 2.23
GR # self refs (0-12)	**1.00** *SD* = .00	**3.64** *SD* = 2.68	**3.05** *SD* = 2.34
GR # other refs (0-10)	**.00** *SD* = .00	**3.11** *SD* = 2.23	**2.24** *SD* = 2.22
GR # ref shared (0-17)	**.00** *SD* = .00	**3.86** *SD* =3.74	**4.37** *SD* = 3.55
GR psych level (1-2)	**1.25** *SD* = .50	**1.42** *SD* = .50	**1.46** *SD* = .50
DT emotion words (0-2)	**.75** *SD* = .96	**.72** *SD* = .74	**.71** *SD* = .68
DT mental states (0-2)	**.00** *SD* = .00	**.28** *SD* = .57	**.32** *SD* = .57
DT psych level (1-3)	**1.75** *SD* = .50	**1.92** *SD* = .73	**1.95** *SD* = .77
SQ # self correct (2-9)	**5.00** *SD* = 1.83	**7.89** *SD* = .98	**7.88** *SD* = 1.14
SQ # other correct (2-8)	**5.00** *SD* = 1.83	**7.31** *SD* = 1.12	**7.32** *SD* = 1.15
SQ self psych level (1-3)	**2.00** *SD* = 1.15	**1.89** *SD* = .75	**2.00** *SD* = .74
SQ other psych level (1-3)	**1.25** *SD* = .50	**1.86** *SD* = .80	**2.00** *SD* = .77
DT composite (0-100)	**15.02** *SD* = 14.06	**35.69** *SD* = 20.87	**37.20** *SD* = 23.35
Immaturity composite (0-100)	**83.90** *SD* = 3.51	**48.31** *SD* = 15.00	**47.28** *SD* = 14.89
SO composite (0-100)	**25.00** *SD* = 6.46	**61.95** *SD* = 23.55	**61.46** *SD* = 23.04
GR composite (0-100)	**6.79** *SD* = 1.25	**44.31** *SD* = 13.89	**41.46** *SD* = 15.00
Narrative composite (0-100)	**10.52** *SD* = 4.28	**47.93** *SD* = 15.97	**48.88** *SD* = 16.91

Person use, language scores and age

Person groups were compared, using non-parametric tests (Mann-Whitney U for two groups and the Kruskal-Wallis for three or more), on TELD raw and age in months scores. The GR "# of persons used" groups are not significantly different from one another on age in months, but do differ significantly on their TELD-2 raw scores. It is group 1 (users of only one person in recounting the story, *M* rank = 20.69) that differs from the other two (group 2, users of two persons, *M* rank = 28.04 and group 3, users of three persons, *M* rank = 25.69), chi square (2) = 6.725, asymp. $p < .05$.

Specifically, it was the referrers to "self only" ($n = 4$; the other four children in the "one person" group referred only to "we" or "us") who differed significantly from others in the sample on both age in months and TELD-2 raw score. The group composed of children who referred to the self only was the only group in the sample that differed from others in age. They were significantly younger than the others (*M* rank = 7.25 compared to 27.09 for others; Mann-Whitney U = 19.00, asymp. $p = .009$) and had significantly lower language scores (*M* rank = 3.63 compared to 27.40 for others; Mann-Whitney U = 4.500, asymp. $p = .002$).

The other important difference in TELD raw score among the person groups is between those who refer to "our" experience and those who do not. Those who refer to "we/us" have a mean TELD rank of 28.23 while those who do not have an average rank of 13.06 (Mann-Whitney U = 72.50, asymp. $p = .005$). The mean TELD score for the group that refers to shared experience is 8.61 points higher than those who do not.

Person use and psychological understanding

Significant differences among person use groups were found on two of the psychological understanding variables: RB task performance and DT mental state use. Results obtained are in the tables below.

The first table presents results found when a Kruskal-Wallis test was done to compare the performance on the RB task of those who used 1, 2 or 3 persons in their game retelling. Data for the separate "self" and "other think" questions, as well as children's performance as a whole (i.e., whether they failed, getting neither question right, or passed by getting both right, or are considered to be "in transition" because they got just one right) are presented so that it can easily be determined where the difference is found.

Table 23. Kruskal-Wallis test results of comparison of "GR # of persons used" groups on RB box task

* $p < .05$

# Persons used in GR	RB "self think"		RB "other think"		RB performance	
	M rank	*Chi square*	*M rank*	*Chi square*	*M rank*	*Chi square*
1	15.63	6.53 (2), p = .038*	17.63	4.75 (2), p=.093, *ns*	15.25	6.530 (2), p = .038*
2	29.81		29.88		30.50	
3	26.29		25.71		26.09	

Groups cannot be distinguished by their response to the question, "If your teacher came in right now, and saw the box closed just like this, what would she think is in

the box?" Their differences on the "self think" account for their different performances.

Once again, it was the "self only" (n = 4) children's performance and not the "shared only" (n = 4) that account for the fact that one-person users do significantly less well on the RB task. A Mann-Whitney U test comparing the "shared only" group (n = 4) with other children (n = 46) shows that there is no significant difference between the two groups in performance on the RB "self think," "other think," or RB task performance. When the "self only" group (n = 4) was compared with the rest of the sample (n = 46), however, the mean rank for both RB "self think" (12.50 compared to 26.63 for others) and for RB performance (11.50 vs. 26.72) were significantly lower than for the rest of the sample. The asymptotic significance levels were .032 and .030 respectively. Those who referred only to themselves in their game recountings were significantly less likely than others, to be able to correctly answer the question, "Before you opened the box... what did you think was in it?"

One-person referrers were also less likely to use mental state terms in their retelling of the Dragon's Tears story (M rank = 21.94) than those who used two persons when describing their game activity (M rank = 32.65), chi square (2) = 7.30, asymp. $p < .05$. On this variable however, there was no evidence that it was the "self only" group alone that accounted for the difference. Neither "self only" nor "shared only" group comparisons yield significant differences.

There is, in addition, one comparison that did not quite reach significance, but is considered worthy of report. The SQ psychological content level variable describes children as having done one of these three things: 1 = correctly answered only the

action questions ("Which animal did you play with?" "Where did you find the food?"), 2 = correctly answered at least one action question and one emotion question ("How did you feel when you saw what was in the box?"), or 3 = correctly answered at least one action question, one emotion question, and one epistemic state question ("Before you opened the box, what did you think was in the box?") So this variable parallels the GR variable in that separate scores are given for self and other, and it parallels the DT psychological content variable because it includes a level for "epistemic state." The person group that differed from others on this variable was the children who mentioned "us" in their recounting. Children who told the story of joint action demonstrated a higher level of psychological understanding in their answers to the questions about the other child's experience than did those who did not talk about what "we" did. The mean rank for children who used "we" was 27.27 compared to 17.44 for those who did not; the Mann-Whitney U is 112.00 and asymp. $p = .051$. There was no significant difference between the two groups on the level of psychological understanding demonstrated in their answers to the questions about the self.

Person use and narrative skill

Person use groups were first compared on their overall narrative skill. The most inclusive measure, because it includes variables from both DT and GR, is the NA (narrative skill) composite. On that score, those who refer to all three "persons" (i.e., 'I', 's/he' and 'we'; mean rank = 29.09, chi square =6.29 (2), asymp. $p = .044$) in the game recounting have significantly higher scores that those who refer to one person

only (*M* rank = 14.69). The difference is accounted for by the low scores of those children who refer only to the self (see table below); children who use "we" or "us" only do not differ significantly from the rest of the sample in narrative skill rank. Children who refer to two persons have a mean NA composite rank of 24.15.

A look at the DT composite and GR composite analyses reveals that the number of persons referred to in the game recounting did not relate to children's overall DT composite score, and that it was on the GR composite that person groups differed. That is not surprising given the fact that GR # self references and GR # other references are two of the variables that make up the GR composite score. Though reference to "we" or "us" is not part of that score, it is the children who used three persons in their recounting who have significantly higher GR composite scores (*M* rank = 32.72) than do those who use two (*M* rank = 19.62) or one (*M* rank = 8.88, chi square (2) = 19.65, asymp. $p = .000$). The differences between the three-person group and one-person group again reflect the lower scores of those who use "self" only rather than those who use "we/us" only. The mean difference in GR composite rank between "we only" users and those who do not refer to shared experience is 111.14 (Mann-Whitney U = 51.00, asymp. $p = .143$). The mean difference in rank between "self only" users and the rest of the sample is 25.00; Mann-Whitney U = .000, asymp. $p = .001$.

That pattern of difference between referrers to "self only" and the rest of the sample is seen, as well, in the comparisons of children's performance on more specific narrative variables. The table that follows shows the Mann-Whitney U comparison results on those narrative variables, as well as the other variables

mentioned above. It provides a summary of all the variables on which "self only" referrers receive lower scores than the rest of the children in the study.

Table 24. Significant results of Mann-Whitney U tests comparing those children who refer to self only in game recounting with the rest of the sample

* $p = < .05$
** $p = < .01$

SELF ONLY GROUP		*n*	Mean Rank	Mann-Whitney U	Asymp. Sig. (2-tailed)
GR % on-task utts	self only	4	4.50	8.00	.003**
	others	46	27.33		
GR # in-seq acts	self only	4	4.50	8.00	.002**
	others	46	27.33		
GR variety/connectives	self only	4	5.00	10.00	.002**
	others	46	27.28		
DT % subj-vb agree	self only	4	5.88	13.50	.005**
	others	46	27.28		
Age in months	self only	4	7.25	19.00	.009**
	others	46	27.09		
TELD-2 raw score	self only	4	3.63	4.500	.002**
	others	46	27.40		
GR composite score	self only	4	2.50	.000	.001**
	others	46	27.50		
Narrative skill	self only	4	3.38	3.500	.002**
	others	46	27.42		
RB "self think"	self only	4	12.50	40.00	.032*
	others	46	26.63		

Two person use groups proved to be more competent than those with whom they were compared. One was those who included the experience of their partner ("the other" in this study) in their game recounting. They differ significantly from the rest of the children in the study on their scores on the within-task narrative variables in Table 24. They also have lower scores on the immaturity (IM) composite (mean rank

of 23.00 compared with 31.93 for others, Mann-Whitney U = 162.00, p = .052) than others and higher scores on the number of GR in-sequence acts that they recount (*M* rank of 27.97 compared to 19.14 for others, Mann-Whitney U = 163.00, p = .052). On both variables, the significance levels do not quite reach the .05 level.

Table 25. Significant results of Mann-Whitney U tests comparing those children who refer to the 'other' in game recounting with those who do not

'OTHER' REFERENCE GROUP	*N*	Mean Rank	Mann-Whitney U	Asymp. Sig. (2-tailed)
GR composite 'other' ref	36	30.50	72.00	.000***
no ref to 'other'	14	12.64		
GR # acts recounted 'other' ref	36	28.35	149.50	.024*
no ref to 'other'	14	18.18		
GR variety/connectives 'other'	36	28.99	126.50	.004**
no 'other' ref	14	16.54		

The second group, those who referred to what "we" did in the game, differed even more from other children. They differed on several more variables and at higher levels of significance than did the "other" use group. The results of the Mann-Whitney U test comparing those who referred to the joint experience of both partners in the game playing with those who talked only about individuals are shown in Table 26.

Table 26. Mann-Whitney U test results of comparison between those children who refer to "us" in game recounting and those who do not

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

'WE/US' REFERENCE GROUPS		*n*	Mean Rank	Mann-Whitney U	Asymp. Sig. (2-tailed)
GR composite	'we' reference	41	27.90	86.00	.013*
	no reference	9	14.56		
NA composite	'we' reference	41	28.37	67.00	.003**
	no reference	9	12.44		
GR # acts recounted	'we' use	41	28.65	55.50	.001**
	no reference	9	11.17		
GR # in-seq acts	'we' reference	41	27.21	55.00	.001**
	no reference	9	17.72		
GR variety/connectives	'we' use	41	28.06	79.50	.005**
	no reference	9	13.83		
DT # acts recounted	'we' use	41	27.66	96.00	.025*
	no reference	9	15.67		
TELD raw score	'we' use	41	28.28	72.50	.005**
	no reference	9	13.06		
SQ other psych content	'we' use	41	27.27	112.00	.051*
	no reference	9	17.44		
Immaturity composite	'we' use	41	22.00	41.00	.000***
	no reference	9	41.44		

Psychological Understanding Groups

The following table provides information about the groups constructed to differentiate children from one another on dimensions of psychological understanding. For the purposes of this study, that understanding is defined as "children's ability to talk about people's actions, desires and goals, emotions, and mental states." The table below provides a description of the groups themselves, their levels, the number and percentage of the children who reached each given level.

Table 27. Composition of psychological understanding groups

DT = Dragon's Tears retelling; GR = Game Recounting
SQ = standardized questions; RB = raisin box task

GROUP VARIABLE NAME	LEVELS OF GROUP	*n*	% of *N*
GR highest psych content level (across persons)	1 = action only	28	56
	2 = action + non-epistemic psych ref	22	44
DT psych content Level	1 = action only	15	30
	2 = 1 + psych	22	44
	3 = 2 + mental state	13	26
SQ psych content Level	1 = action only	13	26
	2 = 1 + emotion	17	34
	3 = 2 + mental state	20	40
DT emotion state use	0 = none	22	44
	1 = one	20	40
	2 = two or more	8	16
DT mental state use	0 = none	37	74
	1 = one	10	20
	2 = two or more	3	6
SQ performance groups	0 = fail	29	58
	1 = transition	15	30
	2 = pass	6	12
RB performance Groups	0 = fail	22	44
	1 = transition	8	16
	2 = pass	20	40

Psychological understanding, language skill, and age

The findings here can in general be summarized in a single statement: Those who are able to respond to, and use mental state terms in discourse differ significantly in language skill from those who are not. That fact may explain the absence of GR psychological content groups from the list above. No child mentioned mental states

in her recounting of "yesterday's" game playing experience, and thus only two groups ("action only" and "action plus [non-epistemic] psychological states") could be formed. No significant differences in either language or age were found between them; in fact, they proved in general to be uninformative. Children who talked only about action did not differ in any measurable way from those who also mentioned non-epistemic psychological states like desires.

Three psychological understanding groups proved to have within them levels of performers that were significantly different from either other in age in months. Those who used no emotion state words in their retelling of the Dragon's Tears story performed at a significantly lower level (*M* range = 21.75) from those who used two or more (*M* rank = 36.31); children who used just one emotion word were closer in rank to those who used none than to those who used two (25.30, chi square (2) = 5.862, p = .053). The DT psychological content group that used action only (level 1) was younger (*M* rank 20.43 compared to 34.00, chi square (2) = 6.487, p = .039*) than the group that referred to actions, psychological and epistemic states in their story retelling. When distinguished by the number of mental state terms used, those who used 0 were younger by a mean of 9.6 months than those who used two or more. Performing a Mann-Whitney U test comparing the mean rank of age of the three groups with each other brought the following results: 22.51 for those who made no mental state references; 29.90 for those making one, and 47.67 for those making two or more references; chi square (2) = 9.402, asymp. *p*. is .009.

Language scores, however, are what distinguish children from each other not only in their use of mental states in the Dragon's Tears retelling, but also in their responses

to direct questions about the mental states of self and other. The following table provides the results of Kruskal-Wallis tests comparing the TELD-2 scores of psychological understanding groups.

Table 28. Kruskal-Wallis psychological group comparisons of TELD raw scores

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

	Levels	Mean ranks	Chi square df = 2	Asymp. p
DT psychological content level	action	20.43	11.908	.003**
	1 + psych states	23.93		
	2 + mental state	34.00		
DT mental state use	0 = none	22.08	8.05	.018*
	1 = 1 used	34.30		
	2 = 1 or more	38.33		
RB performance	0 = fail	19.36	7.02	.030*
	1 = transition	30.94		
	2 = pass	30.08		
SQ performance	0 = fail	21.91	7.39	.025*
	1 = transition	26.87		
	2 = pass	39.42		

Another interesting difference was found, as well, in the relationship between language and RB performance groups, and language and SQ performance groups. Language rank differentiates those who fail the RB task from those who are in transition, as well as those who pass. The language ranks of those who are in transition and those who pass are virtually the same. For the RB task then, those who are in transition are not at a "midpoint" between failing and passing; rather, they are, it seems, "on the verge" of passing.

Children's performance on the SQ task presents a contrast that is a different, but perhaps the more expected, one. The language skill increases as children take each step, from failing to transition to passing, but it is children who fail and those who pass that differ the most from each other.

Psychological understanding and Person Use

The puzzling difference just noted, above, relates to children's person use. Because the other 'think' on the two tasks are correlated (see table below), it is children's answer to the self question that makes the difference between the two (SQ and RB) task results. The relationship between "RB self" and "SQ self" is not significant.

Table 29. Spearman's rho correlations of "think question" responses

$N = 50$

	SQ other think	RB self think	RB other think
SQ self think	.240, *ns*	.243, *ns*	.165, *ns*
SQ other think		.387**	.485**
RB self think			.690**

The variable that proved to be most highly related to person use was the RB (raisin box) performance variable. Those who fail differed from those in transition on % of GR references that were to the self with those who fail referring to the self more exclusively (*M* rank = 32.18) than do those in transition (*M* rank = 15.31). Those who passed had a mean rank of 22.23, chi square = 9.58, p = .008. Those who fail the

RB task also devoted a lower percentage of their person use to "we" references (*M* rank = 19.41) than did those in transition did (34.81). Again, those who passed were in between the two with a rank of 28.48, chi square = 8.00, asymp. p = .018).

Psychological understanding and narrative skill

Mann-Whitney U and Kruskal-Wallis tests were performed in order to compare the psychological understanding groups on variables such as connective use and subject-verb agreement that reflect children's narrative skill. When GR psychological content groups (action vs. action and non-epistemic psychological state users) were compared on these variables, only one significant difference was found. The mean rank of GR in-sequence actions differed significantly depending upon whether children had used action only (22.05) or action and psychological states (29.89) in the recounting (Mann-Whitney U = 211.50, p = .057). Again, this two-level psychological content category in general does little to distinguish more from less-competent children.

When DT psychological level, emotion words, and mental state term use were compared on the narrative variables, the distinction of the tasks in this study was once again manifest. None of the levels of these three psychological groups, derived from the Dragon's Tears retelling task performance, were distinguishable for their scores on any Game Recounting narrative variable. The only differences among them, in narrative skill, were found on DT narrative variables or the narrative skill (NA) composite score, 50% of which is composed of DT variable scores. The following

three tables list the narrative variables on which the psychological groups were tested and reports the statistical significance levels of the differences found within the levels of the groups. The highest ranks are printed in bold type.

Table 30. Kruskal-Wallis comparisons of DT psychological content groups on narrative variables

	Mean Ranks by level			Chi square	Sig. level
	action	1 + psych (non-epistemic)	2 + mental states		
DT total # utterances	16.27	26.05	**35.23**	11.90	.003**
DT # acts retold	13.53	23.86	**42.08**	27.39	.000***
DT % subject-vb agree	15.57	23.14	**40.96**	22.27	.000***
DT variety/connectives	15.77	24.00	**39.27**	19.37	.000***
DT composite score	10.73	24.77	**43.77**	35.87	.000***
NA composite score	16.90	22.98	**39.69**	18.20	.000***

Table 31. Kruskal-Wallis comparisons of DT mental state use groups on narrative variables

	Mean Ranks by level			Chi square	Sig. level
	0 = no use	1 = one reference	2 = two or more refs		
DT total # utterances	22.08	33.45	**41.17**	8.51	.014*
DT # acts retold	19.68	**43.30**	38.00	23.18	.000***
DT % subject-verb agreement	20.07	**41.25**	40.00	19.87	.000***
DT variety/ connectives	20.66	37.10	**46.50**	17.39	.000***
DT composite score	19.08	43.30	**45.33**	27.64	.000***
NA composite score	20.51	**39.80**	39.33	16.66	.000***

Table 32. Kruskal-Wallis comparisons of DT emotion state use groups on narrative variables

	Mean Ranks by level			Chi square	Sig. level
	0 = no use	1 = one reference	2 = two or more refs		
DT total # utterances	20.32	29.30	**30.25**	5.00	.082, ns
DT # acts retold	16.77	**33.53**	29.44	14.63	.001**
DT % subject-vb agree	18.00	**31.92**	30.06	10.54	.005**
DT variety/connectives	18.91	**32.03**	27.31	9.02	.011*
DT composite score	13.59	34.15	**36.63**	26.39	.000***
NA composite score	19.05	**32.78**	25.06	9.30	.010**

Since both psychological content level and DT mental state use describe children's ability to refer to epistemic states in their DT retelling, the similarity of those two results is not surprising. Emotion word use tells a slightly different story; it is less significantly related to fewer variables than are the measures of mental state skill. The numbers of words used, whether one or two (or more), does not clearly differentiate child's level of narrative skill in general. The differences were small and so in each case the significant difference was between those who used no such words and those who used either one or two.

Narrative Skill Groups

Two narrative skills groups were composed for the purposes of group comparisons. The first was a "GR % autobiography group." This variable was excluded from the factor analysis because it did not load on any factor at a .3 level or higher, but its exploration is considered important for theoretical reasons. The ability

to tell a story about a specific event in one's own past is a form of discourse that is just developing among the youngest children in the study, and it is of interest to know whether any of the variables in this study are related to that development. To count as an autobiographical reference, a verb in the game recounting had to both be in the past tense and refer to a specific person or object. "Squirrels eat acorns" did not count; "I found my squirrel's acorn" did count. The groups were divided as evenly as possible (see table below). The second group that was constructed for comparison on person use and psychological understanding was derived from children's narrative skill composite scores.

Table 33. Composition of narrative skills groups
GR = Game Recounting

GROUP VARIABLE NAME	LEVELS OF GROUP	n	% of N
GR % autobiography groups	1 = low (0-69%)	16	32
	2 = medium (70-89)	16	32
	3 = high (90-100%)	18	36
Narrative Skill Composite groups	1 = low (7 – 40)	17	34
	2 = medium (41-57)	16	32
	3 = high (58 – 87)	17	34

Narrative skill and person use

The autobiography groups and narrative skill composite groups were tested on every person-related variable from the Game Recounting (e.g., # of references to self, etc), Standardized Questions (e.g., # of self questions), and the RB task (e.g., self

think). The Self-Other Answer composite was also explored for its relationship to these narrative skill groups.

Significant differences between levels on three variables were found. All were, like the autobiography variable itself, derived from the Game Recounting task: GR # references to the other child, GR # references to "we" or "us," and the # of persons mentioned (see table below).

Table 34. Kruskal-Wallis comparisons of autobiography groups on person-related variables

$N = 50$

	Mean Ranks by level			Chi	Significance
	low	medium	high	square	level
GR # 'other' references	19.09	**33.00**	24.85	7.871	.020
GR # 'we' references	20.71	**32.50**	23.71	5.90	.052
GR # persons mentioned	21.18	**32.72**	23.03	7.539	.023

Similar results were found when the narrative composite groups were explored for differences in person use. The three significant differences were, this time, on the number of 'self' references, number of 'we' references, and the percentage of references to "our" experience of the game.

Table 35. Kruskal-Wallis comparisons of narrative skills composite groups on person-related variables

$N = 50$

	Mean Ranks by level			Chi square	Significance level
	low	medium	high		
GR # self references	19.47	25.26	**33.00**	7.24	.027
GR # "we" references	13.89	27.97	**36.63**	21.07	.000
GR % "we" references	18.19	28.24	**31.17**	7.45	.024

Narrative skill and psychological understanding

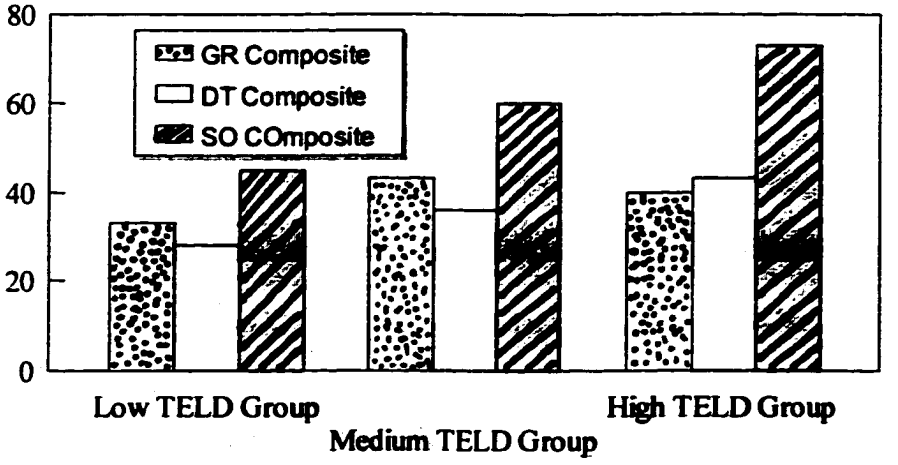
Both the GR autobiography groups and the narrative skill composite groups had levels within them that differed in their use of mental state terms in the retelling of the Dragon's Tears story. Surprisingly, the low autobiography group had a DT mental state use rank that was higher (26.69) than that of the medium group (19.00). The high group's mean rank was 30.82; chi square = 9.371 (2), $p = .009$.

It is important to note that the narrative skill composite score does not include within it any variables relating to the content of either GR or DT, so mental state use is not part of the variable. The DT mental state use mean ranks are as follows: low narrative skill group, 20.31; medium narrative skill group, 24.91, and high narrative skill group, 32.40; chi square = 9.67 (2), p = .008. No other significant differences were found.

Summary of Results

In this study children's verbally expressed understanding of persons and their psychological lives is highly task dependent. The factor analysis solution leads to the hypothesis that the task itself is the underlying cause of the correlations among variables in this study. Regression analysis further informs us that not all task-related composite scores (GR composite, DT composite, and SO composite) are predicted by the same measured aspect of a child's development. While language skill was the measure that most often correlated with other variables in this study, it does not predict performance across all tasks in this study (see figure 2 below).

Figure 2. Mean composite scores of GR, DT, and SO composites by language group



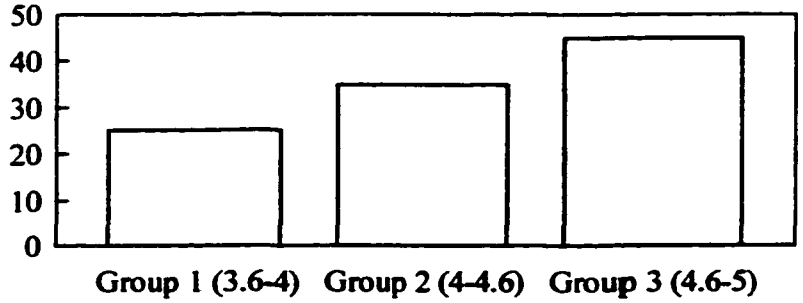
Language scores do, as the graph above indicates, best predict children's ability to answer the direct questions about the experiences of self and other, including the

questions that test children's understanding of false belief. A number of analyses point in that direction: correlations of TELD-2 scores and false belief task performance, multiple regression analyses, and group comparison. Clearly the ability to answer direct questions correctly, particularly questions about false belief, is the mark of a relatively skilled language user. Language scores account for 19% of the variance in SO composite scores.

The findings about the game recounting composite scores are mixed. The loadings of age and language on the GR factor are so low that they are not reported, but multiple regression analysis results indicate that language score is the best predictor of GR performance. The variance explained by them, however, is only 11%. The absence of significant differences between language groups is apparent in the graph above.

Though DT composite scores increase across age group, the finding is that children's age in months best predicts their scores. Twenty-eight percent of the variance in children's ability to retell the story called "The Dragon's Tears" is accounted for by their chronological age (see figure 3 below).

Figure 3. Mean "Dragon's Tears" retelling composite scores



Though the statistical relationship between language and age is not denied here, it is important to note that these tasks rely upon several aspects of maturity, as measured in this study, and that those aspects have different degrees of influence depending upon the task at hand. Maturation or experience with the practice of storytelling plays a stronger role in story retelling skill; general language ability has the greater influence on game recounting and question answering ability.

There is evidence in this study for the development of person use across the preschool period. "Self" and "other" reference, both in numbers and complexity, are common among children in this age group and highly correlated. Those who refer to self as one of two or three persons mentioned, and those who refer to the other child cannot be distinguished from other children in the study on measures of chronological age, linguistic or narrative skill. It is the reference to "we," though even more common in this sample than is reference to the "other," that correlates with higher language and narrative skill as well as with performance on tests of false belief.

DISCUSSION

Overview of chapter

The discussion of results will be organized according to the questions of the study. The discussion of the issue of task-neutrality will be first. The preliminary look at the issue will be followed by sections on the relationship between task neutrality and each of the three elements under study: person use and task neutrality, psychological understanding and task neutrality, false belief and task neutrality, and narrative skill and task neutrality.

The next section will provide an overview of the findings on the issue of person-neutrality. It will describe two important transitions in person use in early childhood, from "self only" reference to use of "us" and "we" in retelling. It will be followed by a discussion of the "self-other" findings. The final section will deal with the relationships between person use and psychological understanding, specifically "theory of mind" and narrative skill.

The issue of task-neutrality

A preliminary look

All of the data collected in this study were gathered to help answer this single, broad question: "What do children, aged 3 ½ to 5, say about the behavior and inner states of self, other, and 'us'?" Each task provided children with a distinct context in

which to verbally demonstrate their level of psychological understanding. Cross-task performance was of interest as an indicator of the task-dependence of children's understanding. Did those who passed the false belief test, for example, also attend to and comment on the mental states of self and other in the game recounting, or of mother, boy or dragon in the story retelling? Answers to such questions might suggest whether the understanding was "in the mind" of the child, as theory theorists would have it, or whether, for example, it was constructed as needed, in response to particular task and social demands.

The study, however, was not designed to focus specifically on the issue of the task dependency of children's psychological understanding. As a result, form and content of tasks were not controlled for; decisions were not made a priori about what elements of difference needed to be present for tasks to be considered distinct. The one exception is that care was taken to assure that the two deceptive box tasks (the one embedded in the game and the RB task) were alike on as many dimensions as possible.

There will be no attempt here to make grand generalizations about the task-dependency of children's psychological understanding. The question in this study is simply whether or not children's articulation of their understanding shows any general features across these diverse tasks, and what the presence or absence of those features might tell us about children's psychological understanding in the preschool period.

Initial correlations (see Appendix F) indicated that children's performance across tasks had some aspects in common. The number of acts that children mentioned in

their retelling of the Dragon's Tears (DT) story correlates with the total number of acts they recount from their game playing experience. The number of actions children mention in DT story also correlate with the number of unprompted in-sequence actions included in their game recounting (GR). Those who gave the more complete retellings of the DT story also provided more detailed and well-sequenced recountings of the game. Because the DT and GR tasks were so dissimilar in both form and content, it is striking that such cross-task correlations were found (for discussion, see narrative skill and task-neutrality section, below).

These correlations, however, provide no answer to the question posed by "theory theorists": Is there evidence that children either "have" or lack particular concepts (of the mind or emotions, for example) that are evident in their performance across tasks? On this question, the answer in this data is, in general, "no." Expected correlations of variables meant to measure "the same" competency (e.g., an understanding of mental or emotional states) were very often not found. For example, children's ability to use mental states in the DT retelling does not correlate significantly with their ability to answer standardized questions about what self and other "thought" was in the box before they opened it. Neither does it correlate with their ability to respond correctly to the self and other "think" questions in the RB task (for further discussion, see psychological content and task specificity section below).

The absence of expected cross-task correlations relating to the topics of interest is no more surprising, perhaps, than the presence of correlations that have no obvious conceptual relationship. In the data there are numbers of correlations that cross both task and topic. For example, children's use of the mental state words provided in the

telling of the Dragon's Tears story, correlates significantly with the number of standardized (SQ) 'self' and 'other' questions about the game playing experience that children answered correctly. Since the questions about mental states are not included in these SQ numbers, but constitute a separate variable, there is no conceptual or topical relationship between mental state use in the story retelling and the ability to answer direct questions about the experience of self and other. Ironically, children's DT mental state use does correlate significantly with their ability to answer direct questions about self and other's emotional states when seeing the box contents (see the further discussion of psychological understanding and narrative skill below). Puzzles like these are presented by the data and cannot be solved by reference to either the presence or the absence of cross-task, topic-related theories.

What then, if anything, does tell a more complete story? Why do we fail to find the expected content-related correlations? The factor analysis informs us that in general it is task not topic that determines children's performance. Variables entered into the analysis clustered together around the tasks themselves.

Why such clustering exists might seem at first to be quite obvious. The numbers are clearly related within tasks. Children who mention more of the actions in DT will of course need to use more connectives to link them. Children who make more references to the self and others in GR report a greater number of actions. Given the obvious and necessary relationship of variables within tasks, and the fact that no attempt was made to control for form or content across tasks, the clustering of same-task variables and the independence of tasks from one another seems exactly what one would predict.

It must be said, though, that most variables are not just indications that children have, within the particular task context, consistently talked more. Higher numbers do represent increases in the quality of children's discourse. The connective use of high scorers is more varied, not just more frequent. Children who tell longer stories tend to use a greater percent of subjects and verbs that agree, and to include mental state reference in their stories. These correlations tell us, then, that children whose DT stories are longer, in terms of the number of actions reported, are also better crafted and more psychologically complex.

The same is true of the game recounting (GR) task. The number of references to the self is not just an indication that more "noise" is being made by children. Frequent referrers to the self are using a greater variety of connectives and putting more actions in sequence. Measures of the quantity of children's talk correlate with those describing the quality of their task performance, giving us the ability to use their composite scores to distinguish levels of competence. Higher scores within tasks do represent higher levels of skill in describing people's actions and inner states.

If expressed psychological understanding is a characteristic of the cognitive competence of the child, rather than a product of task requirements or supports, we would expect that children who score highly on one task would be likely to do better on the others. That, though, is not the case here. Factor analysis indicates a lack of significant correlation among the three task-related factors, a result that is repeated when the relationships among task-specific composite scores are tested. No two composite scores share more than 4% of their variance. The question then is what

accounts for this striking lack of relationship that we find among three tasks that all draw upon children's psychological understanding.

The first question to ask is: Are some tasks harder than others for children in this age and language range? The factor analysis structure matrix (see Results section) suggests that the tasks represented by the DT, GR, and SO (self-other question answers) factors and composites do each present the children with different levels of challenge. The factor called "immaturity" was so named because both age and language load negatively on it. The IM factor indicates which tasks are "easiest" for children to perform. Aspects of tasks that load highly on immaturity indicate that most children performed them frequently or well; therefore those who don't are notably immature compared to the rest of the children who participated in the study. Listed here from most to least indicative of immaturity, the variables on the IM factor are: GR in-sequence acts, GR total # acts recounted, GR # of references to "we" or "us", GR % of on-task variables, GR variety of connectives, DT # of acts retold, SQ # of correct self questions, DT variety of connectives, and DT mental state use.

By this criterion, the game recounting is the easiest task for children to perform. In fact, the IM composite score, derived from those variables that load on the IM factor at .5 or higher, is made up entirely of GR variables. They are the first four listed above. Most children, for example, provided a sequenced account of their game playing experience; thus, the failure to do so is a mark of immaturity within the context of this study. One could also say that, in general, mentioning a number of the important events in the DT story was within the capability of most children to perform. The variables measuring use of connectives and mental state terms are

loaded just high enough on the factor to receive mention. Low connective scores or the absence of mental states in children's narrative are somewhat indicative of children in the lower age and language skill range.

We also learn from this that answering direct questions is most difficult for children. Only one of the SO (self-other answers) composite variables loads on the IM factor at the required level of .3 or higher and it is the one that is a measure of children's ability to answer action and non-epistemic state questions about the self. Even this variable, though, shares only 9% of its variance with the underlying hypothesized cause of the IM factor. Children's inability to answer the direct questions correctly does not indicate particularly low levels of age and language score compared to others in this sample. Thus it can be said that in general answering questions about the self and other is relatively challenging for the children in this study. The negative relationship between the SO and IM factor scores tells the same story and is the only significant correlation among the factors. The ability to answer direct questions increases as children's immaturity level decreases.

Categorizing tasks as simply either easy or hard would, though, be misleading. Each of the three tasks relates differently to age and language, a fact that may help to explain their independence from each other. While age and language are correlated with each other, factor analysis, multiple regression analysis, and group comparisons all confirm the following: DT performance is best predicted by children's age and SO performance is best predicted by their language scores. Neither age nor language group load at .3 level or higher on the GR factor, but multiple regression analysis indicates that language skill better predicts children's performance than age in months

does. So the relative difficulty of the tasks for a given child would depend in part upon that child's age and level of language skill. Assessments of task difficulty cannot be made simplistically, nor can they be applied to all children.

There is no doubt that failure to recount the game experience according to certain minimal criteria (staying on task, putting events in order, mentioning what "we" did) is a strong indication of immaturity in this sample. Performing the task at that minimal level of competence was possible for most children. The mean GR composite score is higher for the medium language group, but not significantly so, and it remains virtually the same for those with high language scores. Both parametric and nonparametric analyses show that none of the group differences is statistically significant. The ability of the least linguistically sophisticated children to perform this task, coupled with the absence of epistemic state terms (a consistent mark of linguistic sophistication in this study) in the recountings of those with the highest language scores results in relatively even GR performance across this age range. This "flat" pattern of scores explains the lack of significant correlation between the GR composite scores and the other two task composite scores. As SO scores increase with language, and DT scores increase with age, GR scores remain relatively even across age and language groups.

The different relationships of the three tasks to age and language alert us to the likelihood that competencies on the three may have some unshared causes and do not develop in precisely the same ways or at the same rates. Those attentive to task requirements and to the role of discourse context in children's performance would expect such differences. They would also note that even in the case of the self-other

answer composite, the one most highly correlated with language scores, seventy percent of the variance in children's scores remains unexplained.

Characteristics and discourse contexts differ among the three tasks. The table below provides a general summary of some of those differences.

Table 35. Summary of task differences in three main tasks of study

	Game Recounting task	**Self-Other question answering task**	**Dragon's Tears retelling task**
Form of target experience (game for GR and SO, story for DT)	action accompanied by investigator's language and (optional) children's language	action accompanied by investigator's language and (optional) children's language	investigator's language accompanied by pictures and (optional) children's language
Participants in target experience	two children and investigator	two children and investigator	child and investigator
Task discourse context	recounting of personal experience	question and answer	fictional story retelling
Child's involvement in target experience	active – action and (optional) speech	active – action and (optional) speech	receptive – listening and (optional) speech
Time between target experience and task	24 (to 48) hours	24 (to 48) hours	several minutes maximum between end of story and retelling
Role child assumes in task	Self, no role	Self, no role	Narrator of story
Comprehension requirements	Minimal: "Tell your mother what you did yesterday."	Words and grammatical forms of twelve questions	Words and grammatical forms of 27 sentences
Production requirements	No specific requirements; child's choice	At least one-word, correct responses	No specific requirements; child's choice
Provision of wording for task response	Child selected the words used	Child selected the words used	Investigator provided the words
Provision of task's narrative structure (tenses, connectives, etc)	Provided by child (supported by investigator prompts)	none required	provided by investigator
Constraints on "correctness" of child's performance	wide variation accepted	Strictly constrained right and wrong	wide variation accepted

Game retelling requires that children describe an experience from their own recent past. The words they are to use are not provided for them, so they need to put their memory of that event into words. They are, however, free to use simple constructions and to limit their reports to comments upon, for example, objects or actions only. To respond correctly to the SO questions, children must remember particular aspects of "yesterday's" game, but must also understand the meaning of the words and the grammatical constructions of the questions. The Dragon's Tears task is a listening and repetition task. Children who can understand, remember and repeat the words and even the grammatical form of those words have a distinct advantage in retelling the story. The task independence that we see may be telling us something about children's familiarity with, and mastery of, the different forms of discourse represented here.

Considering the forms of discourse makes some sense of the findings. Children as young as two years can talk about specific past events in which they have participated (Nelson, 1989) and parents regularly engage their very young children in conversations about the past (Fivush, 1991; Hudson, 1990; Nelson, 1993). So it may well be that the youngest children in this sample brought to the GR task considerable practice in engaging in the kind of discourse requested by the investigator: "Let's pretend you're telling your mother all about the game you and [partner's name] played." Whether children, having begun the task, kept their mothers in mind as the imagined audience is information that is not available to us from this data. It is also not possible to know how children's recountings were affected by the fact that they were describing their experience to someone who had shared it. Comparing these

children with a group who told the story of the game to their own mothers would be required to answer those questions. Young children, though, especially those who attend preschool and daycare may well have had many opportunities to respond to the request of their parents to describe daily experiences in which the parents had not themselves participated.

Direct questioning, however, especially by a teacher (or other adult) who poses a question to which she already knows the answer in order to evaluate the child's knowledge (Mehan, 1979) may be a form of discourse with which young children are far less familiar. If that is the case, it would not be surprising to find that the youngest children had more difficulty answering the questions about the game playing than they did in telling their own stories of what had happened in the same event.

Children's ability to retell a story like the Dragon's Tears, the study tells us, clearly does relate to children's age and thus, very likely, with children's experience with storybooks. The data, however, do not provide any means to evaluate the relative experience children have with such discourse forms. None of the personal information gathered about the children (daycare history in months, hours per week in daycare, family composition, etc.) correlated significantly with any of the variables in the study. Kruskal-Wallis comparisons of groups derived from personal experience data did not reveal any significant relationships between such groups and performance on the tasks in this study.

It can safely be said that in general, the tasks did make specific requirements of children and that children responded to those requirements by performing at different

levels of skill depending upon which task they performed. Specific aspects of task performance were, however, clearly related. A closer look at those specific aspects, both their differences and similarities, follows.

Person use and task dependence

An important question in this study is whether person use constitutes a cognitive task and challenge in its own right, with thinking about "self," "other," and "us" posing distinct challenges, or whether they are all reflections of a common underlying knowledge base or cognitive process. Is it harder for children to talk about the other child than it is about the self? Do children need more support from adults or task contexts when talking about "self" or "other" or "us"? That question will be dealt with more fully in the section on the person-specificity of children's performance.

For now, though, it is important to note that the factor analysis did not cluster variables according to person. The ability to make "self reference" or "other reference" were not distinguished from each other, as deriving from different causes, but rather were highly correlated with each other and loaded together on the tasks of which they were a part. The number of GR references to "self" and the number to "other" are highly correlated and both load only on the GR factor. They do not relate, for example, to children's ability to answer the "RB self" "RB other" questions correctly. The percentage of GR self and other references used are highly correlated, but tell us nothing about the likelihood that children will correctly answer SQ questions about the self and other.

The GR percentage of "other" reference is not significantly correlated with the number of standardized questions about the "other" that children answer correctly. There is no evidence that children have something that we might call "knowledge of the other" that is applied across contexts. In fact, the data on the self are an even more striking indicator that no such person-specific knowledge can be the hypothesized source of performance across tasks. "Percentage of self reference in GR" is negatively correlated, at the .001 level of significance ($r = -.44$), with children's ability to correctly answer direct questions about the actions and non-epistemic states of the self. Those who talk most exclusively about themselves prove least able to tell the investigator what they did or felt during the game.

A closer look at the four children who used 'self only' in their GR further demonstrates the difficulty in trying to find either a neat pattern of "person preference" or evidence of a knowledge base that is person-specific, across contexts. Those children who referred only to the self in the game recounting do not all answer more "self" than "other" questions correctly. One child does; another child does equally well on both. Two of the children answer more "other" than "self" questions correctly. No pattern of better understanding of self than other can be found. The linguistic requirements of specific tasks and discourse formats are a key component of the relationships between tasks on person use.

Psychological content level and task dependence

The observation by Bruner (1986) that narratives include both a landscape of action and a landscape of consciousness prompted a look at the question of whether or not children could be categorized across tasks as being attentive to "action only" or "action and inner states"; reference to epistemic states was expected to be made, as the literature points out, only by the most linguistically sophisticated of the children in the study. Three levels then, of psychological content were expected: action only, action and non-epistemic psychological states (desires, goals, etc), and action and psychological states including epistemic states. It was hoped, when the study began that there would be sufficient cross-task correlation that children could readily be categorized as "action only" or "action and psychological references" or "action, psychological references including epistemic states."

Prior to the study it was not clear how much of their psychological understanding children would reveal in these task contexts. That was especially true in the game recounting because it is a spontaneous narration of an event. Studies show that children demonstrate some understanding of inner states (both mental and emotional) in early childhood, but it is not until some time in middle childhood that children begin to include such states in their spontaneous narratives (Hudson & Shapiro, 1991; Smith, 1978; Wood, 1978). Results of this study support this finding. Not a single child mentioned a mental state in his narrative of "yesterday's" game. That absence is especially striking given that the investigator had encouraged children to pay attention to mental states by asking, "What do you think is in the box?" before each

child opened it. The question was intended to direct children's attention to mental states; hearing their own and their partner's answer was thought to make the false belief apparent to both, and thus, easier to remember and report. Despite these questions and answers, inner states were absent from the game recountings. The absence would come as no surprise to those who had observed the time lag described above.

An interesting and open question is what that time lag indicates. Is children's understanding of inner states still 'under construction' and thus dependent upon the verbal support of others (as is provided when questions are asked, and storybooks are read)? If so, what exactly is being constructed? Is it a network of word meanings that match the adult uses? Might it be that in order to tell an accurately sequenced story, children in this age range need to more fully understand the causes of the more complex emotions, or states of knowledge ("She must know because her friend told her" or "He's embarrassed because he guessed wrong about what would be in the box")? Another possibility is that, despite having some psychological understanding, at these ages, it is still action itself that remains of primary interest to the child who recounts a personal experience.

Questions of children's interest in mental states were not addressed in this study. The lack of mental state terms in the GR context, coupled with their use by some children in the DT and SO task contexts, provides the opportunity to think about the difference task requirements might make on children's demonstrated psychological understanding. One possibility is that children, during the 3½ to 5-year period, require cognitive support for their understanding of mental states. They can respond

to questions or repeat sentences containing mental state language, but are still constructing the concepts necessary to include mental states in their own accounts of people and their actions.

Children use words long before they have fully grasped their network of meanings in adult vocabularies is well-documented (Johnson & Maratsos, 1997; Levy & Nelson, 1994; Moore, Pure, and Furrow, 1990). The use of the word "surprised," for example, is first limited to states of unexpected happiness as one would feel when surprised by a gift or a birthday party (MacLaren & Olson, 1993; Wellman & Banerjee, 1991). When children in this study were asked whether Partner B who found her box empty, was "surprised" by that, the distribution of "yes" and "no" answers did not differ significantly from chance. This is consistent with the finding that not until they are at least five do children develop an understanding of surprise as a reaction to violated beliefs (Hadwin & Perner, 1991). Three-year-olds do not distinguish the meanings of "think" and "know," while four year olds demonstrate an understanding that "I think that..." indicates less certainty than "I know that..." does (Moore, Bryant, and Furrow, 1989). Given that, and the fact that only five children in this study consistently gave correct answers to direct questions about mental states, it seems likely that many children are still in the process of constructing the meanings of such words. They do not have sufficient mastery of them to use them in their own spontaneous narratives.

That possibility does not account for the absence of mental (or any inner state terms) from the recountings of those admittedly few children demonstrating relatively sophisticated understanding of mental state terms in other task contexts. Three

children, for example, did not just repeat the investigator's mental state term use verbatim, but expressed the same meaning in different words (for example, saying "Are you kidding?" instead of "Are you joking?"). Such substitutions seem to indicate that the child understood the word and remembered not just its sound, but its meaning. Because these children, like the others, did not include mental states in their recountings, other reasons for exclusion need to be considered.

Perhaps children did not report yesterday's mental states because they simply did not remember them. Because the investigator's questioning of the children followed their spontaneous game recounting, it may be that children required the investigator's prompting to evoke that memory. Fourteen children, though, correctly answered the question about what they had thought yesterday; thirteen accurately reported the other child's false belief. Six children got both right. Lack of memory for mental states does not account for the absence of mental states from their recountings.

The best explanation would be one that accounts for the absence of inner states from every story, regardless of its teller's linguistic or narrative sophistication. One such explanation is that the mental states of the partners were introduced into the game playing situation only through language, as the investigator asked children, before they opened the box, what they thought might be in it. With the exception of one child who said in recounting, "You was asking us what was in there [pointing to the box]," no child made a single reference to any verbal exchange. Children did not mention, much less report, the investigator's instructions, their own questions or comments, or even the conversation they had with the other child as they set the table so the animals could eat. That raises the question of whether it is speech itself or

mental states in particular that are being omitted. An experimental study would be required to settle that question. It could be that children of this age, engaged in action, do not attend to speech as a separate, but essential part of the game action. As such, the interchange about mental states would not be considered central to the game playing activity.

That many children were able to report speech, even word for word, was demonstrated in their retellings of the Dragon's Tears story. In that context, children demonstrated an awareness of the central role of speech in moving the story along by including in their retellings both the words and inflections provided by the investigator. One major difference between the contexts is that in the DT story many emotional and mental state words are provided for the children. The words not only pointed out the importance of the landscape of consciousness, but also provided children with a ready model to follow in their own retelling.

A second important difference between the two contexts is the role the child is asked to play. In the GR recounting, the child is herself, telling what happened from her own point of view. In the DT retelling situation, children were asked to take on the investigator's role as narrator. Children, imitating the reader's words, placed themselves in the position, not of participant in the story, but of omniscient observer who knew, not only what characters did, but what they said in private conversation (between mother and son, for example), what they thought and felt. Hearing and taking the role of the narrator in stories may well play a crucial role in children's ability to understand inner, specifically mental states. Through it, they "observe"

characters' inner life. They hear and practice the vocabulary of intentions, attitudes, thoughts and feelings that makes inner life apparent.

Because children did not report mental states in their game recountings, the overall psychological content level reached by each child (whether in self, other, or shared experience) in that task had only two levels, "action [and/or being and possession] only" or "action plus psychological states" (mention of goals, desires, etc). Comparisons revealed no significant differences between the groups with respect to composite scores. Further, mean language and age scores were virtually the same, within tenths of a point, across the groups. When compared on other measures of psychological understanding, the performances of the two groups are not significantly different from one another. It may be that finer distinctions between categories (perhaps making references to desire, for example, a separate category) would be necessary in order to differentiate levels of psychological understanding in this age range. It also may well be the case that the children in this study are capable of talking both about action and about the psychological states referred to in the game recounting. Their choice to do so may be a matter of salience or style or other factors unrelated to psychological sophistication.

If mental state use is the new "frontier" for children of this age, one would expect to find the most dramatic differences between those who do, and those who don't, refer to such states as 'thinking' or 'knowing.' In fact, group comparisons of those in each psychological content group in DT (action only, n=15; action and psychological reference, n = 22; and action and psychological reference including epistemic states, n=13) show precisely that. Kruskal-Wallis group comparisons show that significant

differences exist among the three groups on age in months, TELD raw scores, DT emotion word use, SQ number of correct 'self' questions and SQ number of 'other' questions that are answered correctly. In each case, it was the children who used epistemic states in the retelling who performed at a significantly higher level than one of the other two. In this study, it is only the use of epistemic states that distinguishes the levels of psychological understanding. The lack of correlation between GR psychological content level and DT psychological content level can be explained by the fact that children did not refer to epistemic states in GR; the two tasks do not both display the full range of children's knowledge.

Consistency in children's expressions of psychological understanding across the three tasks was simply not found. Emotion use was as task-specific as mental state use was. The use of emotion words in DT retelling does not predict use of such words in GR, since no child mentioned inner life, nor does it predict children's ability to correctly answer the SQ questions about how they and their partners felt about the contents they found in their box. Given this inconsistency, it was not possible to categorize some children as "action only" or "action plus landscape of consciousness," using Bruner's paradigm. The eight children who used action only in both GR and DT shared no other common characteristics. Their language (34 to 57) and composite (15% to 92% on the SO composite and 38% to 86% on immaturity) scores spanned the whole range. Given that, it cannot be said that "action only" indicates that the child is somehow at the beginning of psychological understanding. In fact, in some cases the use of action only served to focus the game recounting and provided a clear, sequential outline of events.

Each task must be viewed as providing both distinct constraints and supports for the expression of children's psychological understanding. In fact, these data might well make us wary of references to "psychological understanding" as if it were a consistent body of knowledge that children either "have" or "lack." If such a coherent body of knowledge "exists" it could only be tested in specific contexts, each of which would both inhibit and encourage particular aspects of that understanding to be expressed. The use of a wide variety of tasks as well as experimental designs, both cross-sectional and longitudinal, is required if we are to get a full picture of the dynamic and context-sensitive psychological knowledge that children use to make sense of the people in their world.

False belief and task dependence

The understanding that one's belief can be false has been determined by philosophers to be the mark of an understanding of the nature of mental life. Interest in the task-neutrality of false belief understanding required that its questions be posed to children in at least two different contexts. Those questions are: "Before you opened the box, what did you think was in it?" ("self think" question) and "Before s/he opened the box what did s/he think was in it?" (SQ "other think" question) or "If your teacher saw the box, closed like this, what would she think is in the box?" (RB "other think" question). One context was the set of standardized questions (SQ) asked about the game (during which children had opened a box and found surprising contents) and the other was the standard experimental version of the deceptive box

task (RB, for "raisin box"). The goal was to answer the question: Do children demonstrate consistency in false belief understanding (or failure to understand) across tasks?

The answer in general is "yes." Children were categorized for each task as having failed (their answers to both self and other "think" questions were wrong), passed (both answers were right), or being "in transition" (with one answer right and the other wrong). Using this person-neutral assessment of task performance, children's responses on the two tasks is significantly correlated ($N = 50$; $r = .53$, $p < .001$). Children's ability to answer the question about the other person's false belief was also correlated across tasks (RB 'other' think and SQ 'other' think $r = .485$, $p < .001$), and SQ 'other' question response correlated, as well, with the RB 'self' think.

The surprising exception to this pattern of correlation was the 'SQ self think' question. Children's responses to that question did not correlate significantly with their responses to any of the other three false belief responses. A careful look at the raw data, though, revealed the source of the anomaly. The decision to code each question independently and as objectively as possible, not making judgments about why the child might have given alternative answers, resulted in two children who failed all other questions getting the SQ 'think' question right. Their mistaken reporting of the real contents of the other child's box when asked about their own, happened to give them the correct answer to the question, "Before you opened the box, what did you think was in it?" That problem did not occur for any other subjects, so is simply reported here to explain the pattern. To test the effect of these two scores, they were temporarily changed to "incorrect"; the resulting correlations

among the four scores all reached significance. Children's responses to the false belief task questions are clearly correlated across tasks.

Before commenting on the "task neutrality" of children's false belief understanding, it is important to note that children's performance across the tasks was not perfectly correlated. While twenty nine children (58%) did demonstrate parallel performance on both tasks, 19 (38%) did better on the RB than the SQ task and only the two children (4%) described as anomalous above, scored higher on the SQ task. Chi square analysis showed that the distribution described above is significantly different from chance (chi square, df, 4 = 12.66, $p < .05$).

It is also interesting to note that the range of language scores for those who passed the RB task is 41 to 57 (M = 50.95); the minimum language score for a child who passed the SQ task is 50 (50.95). The SQ questions about "yesterday's" deceptive box task contents were harder for children to answer correctly than the RB task questions were.

The tasks were designed to be as similar as possible. The wording of the questions was precisely the same. Children in each case had in front of them a visual reminder of their prior false belief (the picture of raisins on the RB task box and the picture of food on the game boxes). The connection, however, between the box picture and the false belief, as well as the discrepancy between thought and reality, had to be kept in mind by children for at least twenty-four hours in order to correctly answer the SQ questions. In the RB task, the demonstration of the discrepancy was nearly simultaneous with the question. Children may be able to function successfully in the RB task using working memory alone.

The SQ questions about the game boxes also included a source of possible error that was not present in the RB task; player B could as reasonably expect to find money as food in the "food box" because the former is what player A actually found. Even a child who knew that player B's belief had been in conflict with reality could, if she did not remember what player B had said she thought, answer incorrectly. Memory for yesterday's speech was necessary in order to give an answer that was correct.

Only three, though, of the eighty mistakes that children made in the SQ task fall into that category. Three children gave evidence that they, though misreporting the object thought to be in the box (saying "leaves," for example, instead of "acorns" for player A or "money" instead of "food" for player B) nevertheless understood that reality and belief were different. Most mistakes children made were precisely the same kind they tend to make when answering the RB task questions. Children answer as if they had known all along what was really in the box. Ninety percent of the incorrect answers are reports that player A, looking at the closed box with the food pictured on it, thought there would be money in the box, and player B, even before opening it, thought her box held nothing. The children then do demonstrate that the basic conceptual demands of the tasks are the same.

The data, then, seem to support the claims of Gopnik, Wellman, and others who believe that children's understanding of mental life is a process of theory formation. At first glance it does seem that children either "have a theory" that allows them to understand that belief and reality can conflict, or they don't. Those who pass both the SQ and RB task certainly do, according to Gopnik and Wellman, have such a theory.

Failure on the SQ task only may indicate for them nothing more than, for example, a failure of memory (of yesterday's words, for example). Children do not have to use their newly formed theory under every circumstance in order to demonstrate that they have one.

The theory that children are forming theories, however, is just one among many explanations of children's developing understanding of their own and others' mental lives. Other theories explain the parallels between the task performances in different ways. Since the wording of the questions remains the same across contexts, language skill, both semantic and syntactic, remains a candidate for the cognitive capacity that allows children to pass the test of false belief. It is also possible, as dynamic systems theorists would claim, that success on the false belief test across task contexts is evidence that a stable state of organization has been reached. The many variables required for its performance (which might include the understanding of particular words and grammatical forms, sufficient experience with the answering of direct questions, the understanding of the causal link between seeing and knowing, etc) have come into new relationship to each other, and entered a period of stable organization. That stability results in a decrease in the variability of responses, and makes responses more consistent and predictable. There is yet, though, no general agreement in the field about which variable finally makes this "new" understanding possible. There is no one theory that sufficiently explains the false belief transition that is so well documented in this period.

The best explanation, however, would be one that takes both the continuity in performance across tasks, and the distinct performance in each task, into account.

The finding of a clear relationship between performance on the two tasks is indisputable. The fact that RB and SQ mistakes are of the same kind points to a shared conceptual base of performance in both contexts. The relative difficulty the children had on each task should be noted. In order to understand children's psychological knowledge, we need to know how and under what circumstances they are able to draw upon that understanding, and use it to make sense of the people in their worlds.

Narrative skill and task neutrality

The within-task coherence of narrative skill in early childhood is apparent from the intercorrelations among variables that load on the same task factor. The variables that have high loadings on the DT factor (DT # of acts, DT % subject-verb agreement, DT mental state use, DT connectives, and DT emotion states) illustrate that coherence. More striking, though, is the fact that the cross-task correlations that are found also relate to children's narrative skill. The construction of a GR and DT cross-task narrative composite score was possible because there were significant correlations between GR and DT # of acts included (Spearman's rho = .409, $p < .01$), and GR and DT variety of connectives used (Spearman's rho = .308, $p < .05$). Children (N = 50) used virtually the same set of connectives for both tasks and those who used a greater variety of them in GR had a tendency to do so in DT.

When constructing a story either about their own experience or from pictures and words provided in a storybook, children drew on their own narrative skills to give the

story coherence and structure. Across tasks, children showed themselves to be either brief or elaborated storytellers. So it can be said that of all the dimensions of children's cognitive skill studied here, narrative skill shows the most cross-task relevance to children's performance.

One could perhaps say that at the root of these correlations is "just" language ability. Of course, narrative skill cannot be entirely independent of general language skill. When TELD raw score is controlled for, the correlation between the GR longest in-sequence string of actions and DT number of acts is no longer significant. The other two significance levels, however, remain. Connective use and the number of total actions recounted continue to be significantly correlated across the two tasks. Narrative skill scores tell us something beyond what language alone, does.

Given the fact that the DT composite score is best predicted, not by language skill, but by age in months, it is important to find out whether age in months might account for the remaining relationships. When partial correlation analyses of DT and GR connectives, and then of DT and GR number of acts are performed, controlling for age in months, their correlations continue to be significant at the $< .05$ levels. So it does seem that some aspects of narrative performance cannot be predicted either by their relationship to language skill or age.

Which of the aspects of DT performance do relate to age in months? Regression analyses of the individual variables included in the DT composite score (see Results) reveal that, except for "DT variety of connectives", all of the DT variables are best predicted by age: DT # acts, DT mental state, DT % subject-verb agreement, DT

connectives, and DT emotion use. That is in striking contrast to both the GR and SO variables, all of which are best predicted by language.

At least two requirements are made of children who are asked to retell the Dragon's Tears story; meeting these requirements may itself require that children have repeated exposure to, and participation in, the specific discourse of storytelling. One requirement is that they pay close attention to the words themselves. A concentration on the pictures results in a story in which objects and simple, visible actions are central ("There's a dragon." "He's crying.") and internal states are rarely, if ever, mentioned.

Secondly, the story asks the children to pay attention to Bruner's "landscape of consciousness." Children who do not have experience with hearing books read to them may not understand those two basic requirements. They may not have developed the attentional strategies required (perhaps inhibiting their attention to the pictures a bit in order to listen more closely to the words) to attend to words. They may not yet understand that stories are "about" inner states as much as they are about objects and actions.

The fact that children use inner state language so much more often in the storybook context may point to a genuine benefit of story telling for psychological understanding. When stories are understood to be "about" what people think and feel as well as what they do, children may be assisted in turning their own attention to those invisible states. When the narrator's point of view is provided, children may be assisted in "seeing into" the minds of the characters and thus, start to understand that there is more to human behavior than meets the eye. Practice with narrative forms

may well enhance that ability. Further exploration of this possibility could prove quite fruitful.

This study, however, indicates that such exploration may not address the issues of "false belief" and "theory of mind" as philosophers and "theory theorists" construct them. The most often-mentioned relationship in the literature, one postulated to exist between theory of mind false belief tasks (here, in RB and SQ think questions) and narrative skill, was not found in this data at all. Narrative skill, as defined here, does not predict performance on either of the two sets of false belief questions. The use of causal connections postulated to be one possible component of narrative (Lewis, 1994) that might correlate with theory of mind understanding does not do so in this study. Children with TELD scores of 41 begin to pass the RB task; no child in this study whose TELD raw score was less than 52, used a causal connective in recounting the game or retelling the story of the Dragon's Tears. While an understanding of the sources of knowledge ("I know there are raisins in the box because I looked inside") may well be at the heart of competence on false belief tasks, causal connective use per se may tell us little about children's understanding of "the mind" as defined by philosophers. If experience with hearing and telling stories were in general, a benefit in RB task, one might expect to find that better DT composite scores, as well as narrative skill scores, would be significantly correlated with RB performance. No such correlations are found.

The issue of person neutrality

Person use: two important transitions

The central question of this study is: Does the recounting of "my," her/his" or "our" experience present the young child with distinct challenges, or does it make little cognitive difference to the child whose story he is telling? Focusing on children 3½ to 5-years old proved to be particularly fruitful for the study of this question. Children's recounting of their game playing experience indicates that children in this age range are, in general, able to talk about the experience of all three grammatical persons. No child told a story without actors; all mentioned at least one person in their retelling. Forty-four children (88%) in this study referred to the experienee of the self. 36 (72%) made explicit mention of the other child's experience in the game and 41 (82%) talked about what "we" did "yesterday." The number of persons referred to by children varied (see table 13 in Results section) with 29 children (58%) referring to all three persons (self, other, and "us"). The game playing experience prompted a range of person use in children's narratives and provided the opportunity to explore what the differences in person use might mean.

The "self" as starting point

As reported in the results section, the comparison of person use groups on their age and language scores told the story of two important transitions in person use

across this period. The first is from the use of "self only" (in this study, n = 4) in the game recounting, to more diverse person references. The limits of a four-person sample are obvious; it is clear that a much larger sample would be required in order to test this finding to see whether it can be generalized. For the purposes of this exploratory study, however, it is important to note both the presence of this group and to observe its patterns of skills and response.

One pattern is striking. In their retellings, which were, on average, 6.50 on-task utterances in length, each child in this group made just one person reference. So it might be said that their exclusive use of the self tells us very little about the development of person use. It may well be that if they had told longer narratives, their recountings would look like others', including more person references and a greater variety of persons. It might be tempting to write them off as outliers and dismiss them as person users of interest here.

Such a dismissal would, though, be unwise. First, the brevity of their recountings is of interest in this study, since it might point to a correlation between narrative skill and person use (see section below). Evidence suggests that the choice of "self" as the only person whose experience these children report is not simply a matter of chance, of individuals' interest, style or other personal variables. Comparisons of means show a pattern. "Self-only" children, when compared using the non-parametric Mann Whitney U test have a mean rank for age that is significantly lower ($p < .01$) than that of the other children in the study. Their mean age is 43.85 months or 3.6 (sd = 1.43) compared to 51.90 or 4.3 for the rest of the sample (sd = 5.40). Those referring only to the self

are also at the low end of the range of language scorers in this sample. Their difference from others in mean language rank is significant at the .002 level. In fact, three of them were the lowest three language scorers in the sample (range = 30 to 57) with scores of 30, 33, and 34; the fourth was number five with a score of 41.

Preliminary correlations (N=50) had pointed to exclusive reference to one's own experience in retelling as related to difficulties on other verbal tasks. The percentage of person references that are about "I" or "me" in the game retelling negatively correlates (see Appendix C, table 1) with: the total number of SQ questions answered correctly, children's ability to pass the RB, and the SQ false belief tests. Mann-Whitney U tests of rank showed that those who referred only to the self in game recounted answered correctly significantly fewer SQ questions than did the rest of the sample. They answered an average of 3.75 questions correctly compared to 6.83 for those who did not limit their person use to self-reference.

On two measures of maturity, age and language, those who referred to "self only" are among the very lowest children who participated in this study. That pattern indicates that this admittedly small sample may be telling us something important about the development of person use in early childhood. A younger sample, as well as a larger one, would enhance the ability to make generalizations about "self only" use, if such use by less mature children was confirmed.

Before discussing the use of "self only" further, it is important to emphasize that it is exclusive reference to the self, and not self reference per se, that differentiates this group of children from the others. The great majority of the children who participated in the study referred to the self, but with different degrees of

psychological and narrative complexity. Comparison of those children who referred to the self with those who did not revealed no significant differences in age or language ability.

Perhaps, then, the distinguishing characteristic of this group is not that they talk about the self, but rather that they narrated the experience of a single person. Results do show a significant difference in the mean language score of those who referred to just one person and those who talked about two. One could, it seems, reasonably suggest that differences in information processing capacity (Case, 1992), particularly the number of units of information that can be processed in working memory at any given time, might lead to these results. Children who can think only about one person when retelling the story of their game experience would be less effective narrators. The number of people in a child's story is certainly an important indicator of cognitive development.

Knowing the number of grammatical persons in a child's recounting, however, does not let us predict the child's maturity or narrative skill. To do that, we need to know whose story they told. While no child recounted only the experience of the other child, some participants did talk just about "us." The two groups of children who both use just one grammatical person in their recountings (self or "we") show strikingly different patterns of skill and performance.

Those who refer only to the self have language scores that are, on average, 14.5 points lower, and ages that are ten months younger, than those of children who tell the story of what "we" did. Correlational analysis of all 50 children shows a similar pattern. More exclusive reference to the self correlates significantly with lower

language scores (see Appendix F, table 2) while more exclusive reference to what "we" did "yesterday" correlates significantly with higher language scores. Clearly it is only the exclusive use of the self that points to lower levels of maturity. The reporting of "my" experience seems to be the starting point for children in this study, and the move away from "self only" to the use of other grammatical persons marks an important transition.

A number of possible reasons come to mind. The inability to give a more elaborate account of the game playing experience and the people involved in it might be the result of the failure to encode, store, or retrieve from long-term memory many of the details of "yesterday's" game. That possibility cannot be dismissed, given the content of the self reference that these children made: "I had that one [flipping to a page of the book that showed each animal puppet with its appropriate food]." "I see [unintelligible] last time [pointing to a page in the book]." "I only had one [animal puppet]" and "That's me right there [pointing to the parrot]." Their references were primarily to objects, prompted, it seems in at least three cases, by the availability of visual reminders of the game activity. The reminders in each case resulted in an accurate report of the animal each had played with, and provided evidence in at least three cases, that children were in fact not just pointing to objects of current interest, but were talking about the game they had played the day before. Perhaps then, these children needed the visual reminders of their own activity, and simply did not remember the objects or actions of their game partner.

One reason for such a failure might be the child's need to concentrate exclusively on what he himself was doing during the game playing session. Less-developed

language skills may mean that these children were focusing their attention on following the instructions given, or on taking the specific actions they were asked to take. They may remember little about the game, and especially about the other child's experience, because they did not attend to, and thus, encode that experience at all. This problem, however, may have been lessened by the fact that children in this game took turns playing. While one played the other could simply stand and observe his partner's action.

The child's own activity level, though, may still have limited her attention to the other player in the game. Though a systematic look at this dimension of children's behavior was not taken, the investigator's impression was that it was, in general, the youngest children who had the most difficulty inhibiting their impulse to "get ahead" of the game by looking in the boxes or walking around the table. As a result, while their partners were taking their turns, younger children were busy, moving toward objects of interest or responding to the investigator's requests that they "please stand there" or "wait just a minute." The possibility, suggested by Carlson, Moses, and Hix (1998) that the failure of inhibitory processes may affect false belief task performance, may apply here as well. It is hard to imagine that children's inability to inhibit their own "off-task" behavior does not have a powerful effect on their ability to attend to and remember the activity of the child with whom they played the game. "Self only" reports would be expected from those who were moving toward their own goals and exploring objects of their own interest, throughout the game. A more careful look, however, at children's activity level across age and language groups would be required to assess whether this impression has any empirical support.

These hypothesized causes can be tested, at least in part, by looking at children's answers to the standardized questions the investigator asked them about the game. These four children's answers revealed that each of the children who told "self only" stories of the game were attentive to the objects used by their partner in the game playing. All four correctly identified the animal their partner used. Two of the four also correctly answered the question about what box their partners went to, to look for their animal's food. Their answers were inconsistent across persons; for example, one child did not correctly report the contents of his own box, but was accurate in his description of his partners'. That inconsistency indicates some confusion about the meaning of the questions or about the game events themselves. Each child, though, clearly had some memory of the other child's object use.

In fact, object use is the only dimension of the game for which we have evidence that these children have memories they can articulate. Whether or not other aspects of the game were remembered, but were too difficult to report, cannot be said. There is the possibility that having pictures and objects available may have biased children toward mention of those objects. Without the game objects in front of them, however, even those limited memories may not have been retrieved and reported. The combination of game recounting and standardized question answers, however, tells us that memory for objects was not person-dependent for these four children.

The fact that the memory for the other child's animal was not reported may indicate some limits in the working memory required to reconstruct the situation verbally, or in the language and narrative skills required to do that. While working memory was not tested here, there is evidence from the assessment of their narratives

that children who referred only to the self were unskilled narrators compared with other children. In fact, their recountings do not constitute "narrative" even in its broadest definition. Narrative, most would agree, requires the reporting of events in time. None of these children correctly order and place two events in the past. One child talks only in the present tense: "that's me right there [pointing to the parrot]... the money's in there" and the object orientation gives the listener the sense that people in action are not central to the children.

Secondly, and especially relevant to this study, is the sense one gets that these four children do not see, and tell stories of, people as agents. Those actions that were mentioned "put money in there," for example, have no subject. None of these children uses action verbs to describe specific human actors. No particular person in these stories "picked an animal" or "looked in the box" or "walked over to the store". The verbs used, to refer to the experience of the self, are being, having, and seeing. Very short narratives, impossible for the naïve listener to understand, are the result. From these recountings it is impossible to tell what the game entailed, what its goal was, who participated, and what each participant did.

Piaget's notion of "egocentrism" comes readily to mind when one reads these recountings. Only a person who had shared the experience and thus entered into the experiential world of the child would be able to make any sense of these recountings. The children themselves did not put themselves in the position of the listener and report accurately what the listener would need to know in order to make sense of the child's experience. So these recountings are "person-neutral" in the sense that the

child's own experience is all there is. It not only provides the only human content of the story; it is the only perspective from which the story makes any sense.

Those who, with Mead and Vygotsky, claim that language itself make perspectives possible would not be surprised that those least skilled in language demonstrate a lack of ability to differentiate self from world and self from other selves. They do not yet have the linguistic tools to make themselves and others objects of reflection, nor can they structure their experiences using such linguistically provided tools as sequence and tense.

In this sense, Piaget's "egocentrism" is a useful concept because it claims that children's own experience is primary. If there is only one reality, it must be the one the child has access to through her own senses, her movements, and her own level of understanding. In fact, the self who experiences just "one reality" may not yet be a "me," in James' term, but may be narrating the experience of the "I" that is, awareness itself.

Perhaps, as Harris (1991) suggests in his discussion of mental states, the experience of "the self" in general is more "direct." In the case of the game recounting, that might mean that the child's own experience is more affectively charged or reinforced in a greater variety of ways. The child watched his partner child open a box; he walked toward his own box, touched it, guessed aloud what might be in it, and felt surprised or disappointed when he saw its contents. It may be easier for children to remember and report experiences that are felt in so many different ways.

More specifically, it might be that understanding and remembering the experience of the self might rely less upon linguistic skill and mediation. Though children need to know how to put their experience into words in order to report it, they may require more linguistic support in order to attend to, and remember the experience of the other child. Children who are more attentive to language, and more skilled in using it as the means to understand a reality that is not the same as their own, are likely to be more adept at reporting the experience of another child.

A far simpler social explanation can also be offered. Young children, especially those who spend considerable time away from their mothers in daycare, are likely to be used to hearing their mothers ask, "What did you do in school today?" Since it is unlikely that their mothers often ask about the activities of another child unless that has an impact upon their own child's experience, children are no doubt used to being asked to tell their own story. Though the investigator in this case, said, "Let's pretend you're telling your mother all about how you and [partner's name] played the game," children who were relying upon their understanding of how this kind of storytelling "usually goes," would tell their own story.

Both correlational evidence (N=50) and the anecdotal evidence from a sample of four, point in the same direction. "Self only" reference is the preference of those whose linguistic and narrative skill is lowest in this study. When reporting skills are minimal, the person whose story children tell is "mine."

to "we/us" reference

There is a second transition in these data that is of particular interest because, despite the fact that it is not mentioned in the literature, it is the one for which the argument is the strongest in these data. Reference not just to the experiences of individuals, but to the experiences that "we" shared distinguishes more-skilled language users in this sample from those who are less skilled. As reported above, when the entire sample is included (N=50) in correlational analysis, we find that more exclusive use of "we" or "us" correlates with higher language scores.

When users of "we" or "us" (n = 41) are compared, using the Mann-Whitney U test of ranks, with children who refer only to "I" and "she" (n = 9), those who refer to "we" and "us" have significantly higher scores on many measures: language scores, GR number of in-sequence acts, GR variety of connectives, DT number of acts reported, SQ number of self questions answered correctly, GR composite score, and Narrative composite score. Those who report shared experiences had significantly lower immaturity scores than those who do not.

There is much evidence in the data that though "I" and "s/he" and "we" are each a single grammatical person, they tell different stories about the cognitive development of the children who use them. One hypothesized story was disconfirmed by the data. Children do not differentiate "I" and "s/he" out of an undifferentiated "we." If that had been the case, we would have found in the data: lower TELD and narrative skill scores for users of 'we' than for those who referred to "I" and "s/he," and mistakes in attributing actions to persons (i.e., confusing "my" actions with "his" or using "we" to

describe the actions of individuals as in "we picked the parrot"). "We" users might well have received lower scores on measures of narrative skill than those who referred to "self" and "other" did. None of these things were found.

In fact, a review of children's "we" references with these findings in mind confirmed that children who used "we" were not confused about its referent. They clearly used it to describe activities in which both self and other participated, whether in turn or at once. Just one subtle mistake with potential to confuse the listener, was found. Nine of the forty-one children who referred to "we" used that pronoun at least once with a singular noun as the object of shared action, giving the false impression that the two children had acted simultaneously on a single object. That is, when adults might say, "We each took a card and then we each picked an animal," or "We took cards and picked animals," children said, for example, "We taked a card and picked a animal." Most children used that construction in sentences which were otherwise grammatically correct ("we picked one of them [referring to the cards]" "we put the dollar in the store" and "we got to look in a box"). While it is possible that children believe that they did act together on one object, that is unlikely. Virtually all children distinguish the objects used by the self from those used by the other, both in answer to direct questions and in even the simplest game recountings. It is possible that agreement between subjects and objects is a challenge for children this age. It is likely as well that they have not yet mastered the expression of the fine-tuned difference between serial and simultaneous action (i.e., a problem that the insertion of "each" would solve).

In any case, the children's choice of verbs in the sentences containing this construction tells us that even these children do understand the meaning of "we." The actions reported are those performed by both self and other during the course of the game. In the sample as a whole, "we" is used appropriately in the overwhelming number of utterances: "We started the snack," "We drove the cars," "We looked in there [pointing to the book] to see what kind of food they eated, " "We went back downstairs."

A few children made relatively sophisticated references to "we." One child clearly conveyed the fact that the two players had indeed reached for a single object: "We both picked this one [card] and then again and then again, and then I choose this one and she choose this one." Two children made references not just to activities that "we" each did in the game, but to something that occurred between them: "...and then we had a fight [about who would choose her card first]" and "we played nicely and we didn't fight." These last two statements, among the rare evaluative comments found in children's narrative of the game, tell an interactive story of self and other.

What, then, does children's ability to use "we" and "us" accurately in the game recounting tell us about the development of person use in early childhood? What cognitive skills might be required of those who refer to "us" that is not required by children who talk about what they themselves, or another individual child, did? One possibility is that rather being the "undifferentiated" state out of which "I" and other emerge, "we" use may be the result of children's ability to integrate and coordinate their understandings of "I" and "other." As Werner and Kaplan propose, "we" as used here, indicates that children have been able to differentiate self from other and

the activities of the self from those of the other, and have chosen to report the common aspects of that experience. The evidence is that, unlike the "self only" users, those who refer to "we" have moved out of the state of undifferentiation to one of increasing differentiation of self and other, and hierarchic integration of the two to talk about what "we" did. The correct use of "we" to include only those activities that were, in fact, shared by two or more people, may be an indicator of just such integration.

As mentioned above, information processing theorists and Robbie Case in particular, from a neo-Piagetian perspective, would point us to the number of elements children need to hold in working memory in order to successfully complete a task. Griffin (1992) and McKeogh (1992) follow Case's lead in claiming that limits in working memory account for children's inability to coordinate different elements of psychological understanding. Griffin explores the relationship between events and intentions; McKeogh discusses the difficulties children have in coordinating the "landscape of action" and the "landscape of consciousness." No one has explored, however, the relationship of working memory and person use. Specifically, are persons "elements" in Case's sense, or does that term apply only to the specific event that is being described? If people are "elements" in working memory, is "we" one element or two? If it is considered just one, then the data provide evidence against this theory, given the fact that both "me" and "us" are one element, and seem to require different levels of cognitive skill to narrate. If two persons share an intention or a goal, do they become one "element?" These questions would need to be

answered before the usefulness of this approach to the question of person use could be determined.

Another approach that one might take to "we" reference is to note its possible relationship to the topic of categorization in human thought. Categorization is a much studied issue in cognitive development. Whether or not children learn Rosch's (1975) "basic level" category of "dog" before learning the superordinate "animal" or subordinate "poodle" has long been debated and explored. The categorization of persons, though, a central topic in social psychology as it relates to issues of stereotyping, for example, has been less studied by developmentalists. Infants' ability to differentiate human intentional action from the movement of inanimate objects has been studied, as has the development of gender categories in the preschool period. The question of interest here, however, has not been studied: Is the individual the "basic level" category in children's developing taxonomy of persons?

A functional understanding of categorization (Nelson, 1996) would lead us to expect that because it is individuals who act, children may first attend to and tell the stories of individuals. "We" don't grasp particular objects, speak particular words, sit or stand or run; individuals do. The use of "we," then, might be thought of as being a superordinate category, as "animal" is to "dog." It is an abstraction, a general class of distinct, functioning individuals. Though the issue of whether Rosch's "basic level" categories are always learned first has been debated, the claim would be here that "we" is, for narrative purposes at least, a later development and thus, indicative of higher levels of linguistic and narrative skill.

One could also claim, though, that for purposes of narration, the added difficulty for those who use "we" is not that "we" itself is a superordinate category, but rather that it requires a greater use of such categories in the sentences and story itself. When telling the story of what "we" did, children often need to summarize activities and to refer to superordinate categories of objects. They need to say, "We picked animals" instead of "We picked a parrot and a goat." That children have difficulty constructing sentences that include "we" is noted above. Two observations, however, point away from this as the primary difficulty children encounter. For one thing, the mistakes children made tended not to be mistakes in categorization as much as in number. They failed to use the plural form when needed, saying "We picked a animal" instead of "We picked animals." Secondly, it was most often true that a superordinate category was not required to make a correct sentence. Children accurately said, for example, "We drove cars" or "We put forks on the table." They did not need, in this case at least, to say, "We drove vehicles" and "We put silverware on the table." Nevertheless, the use of "we" was indicative of higher levels of language, and narrative skill. Further research focused on the issue of categorization of person would be required to see whether looking at the use of "we" from that perspective is fruitful.

It would also be important to test the use of "we" in situations in which the other was not a peer. Could it be, for example, that the development of "we" language differs depending upon who the "other" is? One could imagine that children might much more readily refer to "we" when talking about activities shared with family members. That might be especially true if the "other" were the child's mother or

another primary caregiver with whom the child has long engaged in coordinated activities as basic as feeding and bedtime rituals. It might even be that some uses of "we" in those contexts would be of the undifferentiated kind in which the distinction between self and other is blurred. Nelson's (1989) analysis of two-year-old Emily's first uses of "we" in her bedtime monologues gives such ideas plausibility. Emily, Nelson observes, begins to refer to "we" when describing normative ("my have to get checkup, so *we* take my jamas off.") and group activities ("We bought a baby [doll]") (Nelson, p. 289). Both take place in the context of events shared with family members, and the first might be construed as indication of a blurring of the boundary between self and other. Many questions remain to be asked and answered about children's developing reference to shared experience.

Whatever else is said about it, the findings here are strong and convincing, especially because they are derived from the entire 50-person sample. The use of "we" in the recounting of this particular game is the singularly most telling person-related indicator of linguistic maturity in this study. The issue of its developing use, thus far little studied, is an area of investigation that could well provide us with rich insight into both cognitive and social development.

The self-other issue

The self-other story, much more often debated in the literature, is the subtler story to tell here. At first glance, it seems that there is little difference to report. "Self" and "other" reference correlate with each other (see Appendix F) in: 1) number of

references made in the children's game recountings, 2) GR level of psychological content, 3) the number of direct questions about the experience of each that were answered correctly, 4) the psychological content of the self and other questions that were answered correctly and 5) the answers to the "think" questions in the raisin box task and the standardized questions. When "self" reference, for example, does not correlate significantly with age or TELD scores, neither does "other" reference. The number of "self" and "other" references load, at similar levels, on each of the four factors. Neither tells us anything about the immaturity level of the child, nor do they say anything about the likelihood that children will correctly answer the direct questions posed to them. Articulation of the experiences of self and other seem to go together in these data, making it difficult to easily rank them according to the relative challenges they pose. Generally, children who can refer to, or answer questions about the self can also do the same when the topic is the other child. That evenness would seem to point to person-neutrality in psychological understanding and narrative skill, at least as far as self and peer are concerned.

Given the exploratory nature of this study, however, it seems important to point out that the story may not be as straightforward as it first appears. There are hints in the data that talking about the other child may be more difficult for children than talking about the self is. There are subtle differences between "self" and "other" performance that are worth noting.

First, there is the evidence that thinking and talking about the self is the "floor" in these data. Though the sample size is small (n=4), the above discussion of the "self only" group indicates that self-reference is where children begin in game recounting.

The shortest and simplest recountings, given by children whose immaturity scores were higher, and whose composite scores on Dragon's Tears, game recounting, and standardized questions were all significantly lower than the rest of the sample, were stories about "me" and "my" experience in the game.

No one in this sample tells a brief story that is only about what "he" or "she" did while playing the game. The absence of exclusive reference to the other child in the data set raises fewer questions than its presence would have. The expectation that children will narrate their own life story, including other people as relevant to them and their own activity, is well-established. To hear a child tell the story of a game in which they themselves participated, and mention only the experiences of their partner, would sound "strange" and might even prove a bit worrisome to an adult listener, unless it was clear to the adult what the partner's experience meant to the child, and why it was so important to the teller. While adults often offer narratives that are not "about" the "self," it is generally expected that the narrative will reveal something about the speaker. We expect narratives to inform us of the narrator's experiences, attitudes, interests, group identifications, and worldview. Children who talk directly about their own experiences may be expressing a beginning understanding of the self-revelatory purpose of narrative in this culture.

Secondly, while self and other performance in general, correlate positively with each other, each tells a slightly different story about the children who use them. When stepwise regression analysis is done, entering both percentage of self reference and percentage of other reference, the self-other composite score proves to be predicted by percentage of self used, but not by the percentage of other. The beta

weights are negative indicating that higher percentages of self reference predict lower self-other composite scores (see Results). Missing "self" questions is more predictive of immaturity than missing "other" questions is; in fact, the percentage of self references used in the game recounting predicts the immaturity composite score while percentage of other references does not.

So once again it would appear that the use of self is a kind of default category or baseline. A larger sample size, and perhaps children younger than 3.6 would be needed to strengthen the case. In this study, though, the story seems to be that recounting of the experience of the self, both in narrative form and in answer to direct questions, is less difficult for children than is talking about the experience of the other.

Why this difference exists, as subtle as it is, is the very question that prompted this study. In general, children's concepts of action or inner states at this period of their development, are not person specific; they apply to self and other and "us." It is nevertheless the case that the self seems to be the first person whose experiences children articulate. In this study, the self is at first simply the possessor or user of objects, and does not seem to be taken "as an object to the self" who is telling the story. Only later do children begin to refer to both "self" and "other," and then to us. And though a correlational study does not permit claims of causation, it is clear here that language skill is an essential dimension of the change we observe here. Low language scores mark exclusive use of the self, and those who talk about what "we" did are among the higher scorers in this study.

Psychological content level and person-neutrality

The results section is full of evidence that children's verbally-expressed understanding of the psychological lives of self and other run parallel during this period in early childhood. In the game recounting, standardized question answers, and the deceptive box tasks, self and other references showed, in general, the same level of psychological understanding. Children did not, for example, talk about what "I thought" was in the box or felt about the game, and then report only the other child's action. The stories of self and other were told, in general, at the same psychological level.

The answers children gave to the standardized questions about the game had that same characteristic. Children not only answered approximately the same number of self and other questions correctly, but when coded for psychological level (i.e., whether questions about "object and action only" or "action and emotion" or "action, emotion, and epistemic state" were answered correctly), it was apparent here as well that "self" and "other" answers were generally the same on this dimension. The Spearman's rho correlation of SQ self psychological content and SQ other is significant.

It is important to say here that children who answered both correctly could not simply repeat the same answer twice. The game had been designed in order to assure that self and other would have different experiences, so that the investigator could tell whether children were accurately reporting the experience of each, or whether they had a tendency to make egocentric or allocentric mistakes. As a result of those

different experiences, the self-other questions could not be answered without understanding the way that the other child's experience differed from one's own. The ability to correctly respond to the question about the emotional state of the self and the other after seeing what was in the box, required, for example, that the child report two opposite states. The child who found the money felt "good" or "happy," and the child who found nothing felt "bad" or "sad." So, when the correlation of self and other emotion answers is found to be significant at the .000 level, as it is in this study, that evidence is convincing.

The verbal data in this study lead to the conclusion that psychological understanding in preschool period does not differ depending upon the particular person whose inner life is being elaborated. Children in this setting do not seem to know more about "me" than they do about "him" or "her." At least they don't express differential knowledge verbally in this context.

A second way to approach the issue of person use and psychological understanding is to ask about the differences among the person use groups mentioned above. As previously noted, the story recountings of the "self only" children did not offer a single clue as to the inner state of the self; all four of the children were categorized as having made GR "action only" reference. It is important to remember, though, that the range of psychological content response was narrow, and no significant differences were found across person use groups. Analysis of the psychological content level of the SQ questions that "self only" users answered correctly, confirms their orientation to action. The group missed inner state questions, and could answer only those questions which were about action: "Which

animal did you/ s/he pick?" and "Which box did you/ s/he go to?" Only one child in the "self only" group answered an inner state question correctly. When asked, "How did you feel when you saw what was in your box?" The child, who had found money, said, "good." All four failed both tests of false belief understanding.

"We" users, in contrast to the "self only" users, answered significantly more of the SQ (non-epistemic state) questions correctly than the rest of the sample did. They also answered both self and other questions, including epistemic states, at a higher psychological level than those who did not refer to "us." For both sets of questions, the difference in rank was significant. Spearman's rho correlations, performed to explore the relationships between percentage of "we" use and psychological understanding revealed that as the percentage of "us" use increased both SQ self and other psychological level increased.

One explanation of this increase is simply that children who talked about what "we" did had more complete and accurate memories of the game playing experience than the rest of the children did. That memory would enhance their ability to answer the direct questions about the game playing accurately. There is, in fact, evidence that children who referred to "we" referred to significantly more actions of the game than the nine children who did not refer to us. But we also have evidence that the ability to answer the epistemic state questions that are included here do not depend just on memory. The fact that children who use "we" can answer more mental state questions correctly than other children can will be discussed further in the section below.

Theory of mind and person-neutrality

Person use in GR recounting has a clear relationship to the ability to pass the false belief task in this study. The ability to answer the "RB self think" question ("Before I opened the box, when it was closed just like this [showing a picture of raisins on its cover], what did you think was in the box?") correlates with the percentage of self reference and the percentage of shared reference, but in opposite directions. Children who talked more exclusively about the self in the game retelling were less likely to answer the RB "self think" question correctly; those for whom the percentage of shared reference was high were more likely to answer correctly. The tendency to talk about "us" instead of "me" correlates with increased ability to accurately describe one's own prior mental state. The percentage of references to "us" correlates significantly not only with the answer to the "self" questions posed to the children, but to their general performance on both the SQ and RB false belief tasks.

The question of person-neutrality has at least two aspects in this study. The first reflects upon the narrator herself. What is the difference between a "self only" user and a "we" user that might make those two people differ in their performance on the false belief tasks? These data inform us that at least one of the important distinguishing characteristics of the two narrators is language skill. In this case, we can be quite certain that it is not, in particular, narrative skill that makes the difference since we have consistently found that there is no significant correlation between false belief task scores and narrative skill. General language skill, however, is highly correlated with "we" use and false belief task performance. Domain specific

theories would not predict these relations between children's ability to tell a story with particular characters and their theory of mind. The correlation, of course, does not help us to determine whether or not a specific conceptual or linguistic "achievement" is common to both narrative of shared experience and false belief task success.

"Theory theorists" assume that the understanding that self and other are thinkers should be constructed simultaneously by the child. This aspect of person-neutrality, whether a given "narrator" would understand the false belief of self and other similarity is given considerable support in this study. Half of those who get just one "SQ think" question right answer the "self" question correctly; the other half get the "SQ other think" question right. The finding that children seem to either "have it" or not is hard to deny.

What it is that they "have" remains an open question. Harris' information processing theory (1991) would just as easily explain the data on the RB task as the "theory theory" does. Harris' simulation theory predicts that both questions would be equally hard for children because each requires an overriding of a current state of affairs. The "self think" question requires that the child override her knowledge that there are really crayons in the box, in order to correctly say that she had thought there would be "raisins" in it. When asked what the teacher would think, the child must do that again. He must put aside his own knowledge of what is really in the box to imagine that the teacher would think there were "raisins" in it. Harris' theory would also make sense of the increased difficulty that children have when there are two boxes as in the SQ task. Children must override not only their current understanding

of what is in this box, but must also distinguish it from the reality (and thoughts) about the contents of the other box.

Likewise, those who believe that success on the 'theory of mind' task is dependent upon language skill could make a strong case from this data. The ability to answer these questions does differentiate those with high language skills from those who are less linguistically mature. Children who are able to use "think" correctly are also able to tell stories of both self and other. The understanding that "think" is non-factive is not person-related, and the knowledge, for example, of the source of beliefs ("She doesn't know what's in the box because she hasn't seen it") may well derive from the experience of hearing people describe such causal connections in words. It would not, then, be unreasonable to think that both of those language-related characteristics, linguistic skill and ability to reflect upon both self and other, would be manifest in children's response to these tasks.

The question remains, however, of whether the false belief tasks reflect the same, if subtle, priority of self that is apparent in their children's ability to answer non-epistemic questions about the self and the other. Remembering that the number of self questions wrong is indicative of immaturity (and the number of other questions wrong is not) and that the ability to answer "other" questions is a sign of maturity, we might wonder whether the same applies for the false belief tasks. The false belief tasks, it seems, are much more constrained and thus, the range of responses is narrow. Especially with the RB task, there is a clear tendency to get either both right or both wrong, though those who are in transition are more likely to answer just the "self" question correctly.

So the false belief task does not show as clearly that knowledge of one's own mental life has priority in development. Noting that false belief task success is unrelated to the use of DT mental states or narrative skill, it is possible that success depends upon mastering a concept that is unrelated to psychological understanding. Two good candidates are children's difficulty in inhibiting their reality-based responses, and their limited knowledge of the sources of belief.

Narrative skill and person-neutrality

Clearly the number and identity (see Results section) of those persons included in children's game recounting tells us something about a child's narrative skill. "Self only" reference is an indication that a child is having trouble telling a story at all. Those who referred to self only (n=4) had a mean narrative composite score that was 37 points lower, as well as a significantly lower mean rank, than that of the rest of the children. Those who mentioned "we only" and those who referred to the "other" did not differ significantly in narrative skill from the rest of the children who participated in the study.

Those who talked about what "we" did proved to be the most skilled narrators. Their narrative composite score, which includes variables from both DT and GR, was 21 points higher than those who referred only to individuals and their mean rank was significantly higher as well. They were not, it should be noted, simply the "best performers" in the study. For the most part the group did not differ from others on measures of psychological content or on their ability to answer direct questions. "We

users" do outscore non-users on the test of language development and on the specific narrative variables of GR in-sequence acts, and GR variety of connectives. Interestingly, the only DT variable on which they differ from others is on the number of actions they report.

The specific question of the use of "self," "other," and "we" cannot be asked about the DT task since it was not a story of children's own experience. To test what difference the provision of words has on person use, a story of children's own experience would have to be told by the experimenter. Perhaps such provision would aid children whose language skill was lower to make themselves and others an object, a character in their own story. However, the ability to attend to language as an alternative representation of reality may be something that more experienced story listeners and tellers are better able to do. If that level of language use is necessary for a successful retelling, then children in the lowest language groups may not find the provision of words terribly useful to them. Children of low language scores in this study were less able than higher groups to retell the story of the Dragon's Tears.

Narrative skill variables in this study tend to correlate with each other. Part of the reason is, as stated in the task-neutrality section, the interrelationships of the numbers themselves. However, it is striking here that narrative skill does not predict theory of mind performance at all, much less the particular abilities of children to answer either the "self" or "other" questions correctly. Person use, then, seems to be related to narrative skill, but not to other aspects of performance tested in this study. The exception of course, is their common relationship to language skill in general.

The relationship between task and person

This study has dealt with two main issues: the task-neutrality and person-neutrality of children's psychological understanding. In the table below the relationship between person use and task performance is summarized. A minus sign indicates that when the group named (e.g., "self only, n = 4") was compared by mean rank with the rest of the participants in the study (n = 46), its score on the variable listed (e.g., "age in months") was significantly ($p < .05$) lower than that of the others. A plus sign indicates that the group mean rank was significantly higher than the rest of the participants.

Table 36. Results of Mann-Whitney U test comparisons of person use groups with the rest of the sample on age, language, and task-related variables

IM = immaturity; GR = game recounting
DT = Dragon's Tears; SO = self and other question answers
NA = narrative composite
* = asymp. p = .052

	SELF ONLY (n=4)	OTHER REF (n=36)	'WE' REF (n=41)
Age in months	-		
Language score	-		+
IM composite	+	-*	-
GR composite	-	+	+
DT composite	-		
NA composite	-		+
SO composite	-		
RB & SQ 'think'	-		

The most striking finding is that inability crosses task boundaries more readily than ability does. Such cross-task consistency indicates that use of "self only" reflects a more fundamental characteristic of children's cognition that affects performance even on those tasks (DT retelling, for example) that do not require "self" or "other" or "us" reference. The current data, though correlational, point to general immaturity, especially of language skill, as the one that accounts for children's relative difficulty with these tasks. "Self only" users' performance is, in that sense, truly task-neutral.

The person reference groups whose members are more linguistically competent demonstrate greater task-sensitivity. Perhaps in part because they can make use of the linguistic support provided by the investigator, they may well be better able to

differentiate task requirements from each other and respond more appropriately to each task as describe.

The data seem to indicate that the ability to refer to the experience of the "self" when it is not used exclusively, or to the "other" child does not signal a broader change in children's cognition. The groups of children who use them do not differ significantly from others on variables not related to the game recounting.

The data do seem to indicate, however, that "we" reference is related to more general capacities. The percentage of "we" reference correlates with the ability to answer more self-related questions correctly, for example. Most striking, though, is its positive relationship to language scores.

The tasks in the study have vastly different characteristics and requirements; no attempt was made to eliminate those differences. So it is striking to find that reference to "self only" and to "us" relates in any way at all to children's performance on tasks that are so different. A case can be made that it is the language skill of their users that gives them that relationship to other verbal tasks.

The development of person use

One of the main findings of this study is that children at the lowest level of language skill recounted a shared game playing experience as if "I" was the only player. If further research supports the anecdotal finding here, it may be that very young children, with narrative skills that are not yet well-developed are likely, when asked to recount a shared experience, to talk about what "I did." The descriptions

they give of their own experience are likely, as well, to have some strict limits. Children in this study focused their attention on the objects within their reach, and described themselves in relationship to those objects ("I had this one"). They were apparently unaware of, or unable to articulate, psychological states of any kind.

The slight "edge" that self understanding has in this data may well indicate to us that there is something about experiences that are perceived through, and in one's own body and that are relevant to one's own physical, emotional, and social well-being that have a particular salience for children (and arguably, for adults as well). That salience would make the self an apt starting point for children's developing understanding of people.

The ability to articulate the experiences of both self and other, as well as "we" seems to signal a new awareness of "person." Children who mention either self and other, or "us" show evidence of being able to remember, and to articulate the different experiences of self and other, and to answer questions about them in parallel. Children who answered questions about emotion correctly in reference to the self also did so when asked about the other, for example. In addition, children who use higher percentages of "we" are more likely to be able to articulate the mental states of both self and other in the RB and SQ false belief task contexts.

The findings then depict move from self only to self and other and "us." Whether this move can be described as a sequence or a clear series of stages is doubtful. Like all aspects of language development, it is likely to have different courses and timing for different children. Whether general patterns of development can be discerned in

the presence of such variation is yet to be determined. The findings here indicate that person use in the preschool period is worthy of much further exploration.

CONCLUSIONS AND IMPLICATIONS

The joy of exploration is greatest when one finds not only what one was looking for, but also discovers the unexpected. This study of the psychological understanding of "self" and "other" in early childhood led to both. It fulfilled its original purpose of allowing us to compare young children's understanding of "self" as actor and thinker with their understanding of the "other" as one who acts and thinks. The general finding of significant correlations between self- and other- understanding on all dimensions measured has important implications for our understanding of the process by which children's concept of person develops. Those implications will be discussed below.

The study, though, did more than answer the question posed. Children's recountings of their playing of the game were not limited to descriptions of what "I" or "s/he" did. Children told the story of "our" experience and in so doing made possible the description of a relationship that had not previously been noted in the literature: More exclusive reference to what "we" did together is associated with greater linguistic skill among the children in this study. Higher percentage of person reference devoted to what "we" did is significantly correlated with better language skills, higher narrative composite scores, and greater likelihood of correctly answering the mental state questions posed about the deceptive boxes. In this study, it is the consistent use of the first person plural form in recounting the game experience that relates to cognitive maturity.

That surprising finding prompts countless questions that call for further study. It will be important to find out, for example, whether the relationship between language

skill and the use of "we" is general across contexts in this age range, or whether it is specific to interactions with peers, or related in some way to preschool experience. Much could be learned by tracing individual children's developing use of "we" and "us" longitudinally, as Nelson (1989) did, to see the way its linguistic and social functions change with development. Children in this age range are not using these pronoun forms for the first time, but are perhaps beginning to use them more consistently in extended forms of discourse. It makes sense to ask, then, whether familiarity with the forms themselves (in this case, narratives of personal experience) is a key component in children's enhanced ability to talk about "us."

We will want to find out, as well, what the relationship is between first person plural pronoun use and children's developing understanding of themselves as partners in shared activity or as those whose experience is both like and unlike others in certain ways. Underlying all these questions is that of the relationship between thought and language. Raised whenever linguistic data are used as a means of exploring children's developing thought, assumptions about that relationship will need to be made explicit and revised, if necessary, as new data become available.

The findings and interpretation of the current study will serve as a foundation for further research into the relationships among person use, psychological understanding, and language and narrative skill. While making no claims to have found "the cause" of children's developing concepts of persons as actors or thinkers, this study did find a striking pattern of relationships that points toward an answer to the question posed in this study. The question is: What does children's relative understanding of "self" and "other" tell us about the process by which they construct

an understanding of persons as actors and thinkers? In addition to the finding that "we" use is correlated with language skill, the most important findings are three: 1) In recounting the game experience, reference to the "self only" was the choice of those with limited language skill; 2) "Self" reference and "other" reference were correlated, across the sample, both in number and in content, and 3) A pattern of relationship between person ("self only" and "we") use and linguistic skill was found.

The finding of self-reference as a starting point is consistent with Piaget's notion of "egocentrism." Piaget declares that the "primacy of his own point of view is... perfectly natural [for the young child]" (Piaget, 1923, page 161), and so Piaget would not be surprised to find that children in this sample first tell the story of "my" experience. Gopnik and others who believe that theory construction is the means by which children come to understand that people are thinkers would find their theory supported by the correlation of "self" and "other" reference in this study. Gopnik's assertion has been confirmed here: On deceptive box tasks, children tend to answer either correctly or incorrectly on both "self" and "other" questions. Children in transition, who answer just one of the two correctly, seem to do so randomly and no general pattern of "self first" understanding is found. Each of these theories, then, interprets one aspect of the data presented here.

Neither, though, can explain all aspects of the findings. Egocentrism does not account for the parallel references to "self" and "other." Positing the understanding of mental states as the result of theory construction does not explain why children whose language skill is limited might talk only about the self. Neither theory addresses the relationship of language to the pattern of person use found here.

When one takes all three aspects of the findings into account, it is the theory of George Herbert Mead that comes to mind. Mead claims that it is only through language that the child becomes "an object to himself by taking the attitude of other individuals toward himself..." (Mead, 1934, p. 225). Nelson (1996) describes that process in depth, noting the child's initial immersion in events in which person, object, and action form an inseparable whole. The child so immersed is, according to Nelson, an "experiencing self," absorbed in the events in which she is engaged. It is only through language, Nelson says, the child gains the ability to step back from those events and to reflect upon them and upon "me" as participant in them.

Nelson describes the functions performed by language that make that reflection possible. Language allows children to name things, and thus to separate person from object and object from action. To meet the demands of language, especially in its narrative forms, the child must cognitively slow events down and put even simultaneously occurring events in linear order. She begins to be able to tell the story of a self who is not the one experiencing a current event, but the one who "played a game yesterday." The child becomes an "object to herself," a character in her own life story, so to speak.

Considering the role of language may also make sense of the finding of a relationship between language skill and "we" use. To talk about what "we" did requires the ability to recount the experience at the appropriate level of generality. The child must leave out some of the concrete details that are so salient to young children, referring to "the purple octopus" as one of the "animals we picked." Language functions to enable such generalization and categorization by drawing our

attention to shared aspects of objects, persons, and experiences, and giving labels to things, like "animals" that do not exist except in language.

The child also gains access through language to domains of cultural knowledge that would remain hidden without it. Among them are those of subjective experience, and particularly of mental states. Through language, then, the child constructs an understanding of persons, including the self, which would not exist apart from language. It would be nonsense, then, to talk as if conceptual life is simply expressed "as is" through language, or that the child is expressing in language precisely the same concept he would have constructed without it. Nonlinguistic forms of thought and language transform each other continually. They are parts of a complex cognitive system in which elements interact and bring about change in each other, and in the whole. When we study language, we look at a part of that whole, an essential and powerfully influential part that is publicly shareable, measurable, and dynamic.

A closer look at that dynamism is an important next step to take in the study of the relationship among language, person use, and psychological understanding. Admittedly, the cross-sectional design stopped the clock at a particular moment in each child's development. To do so served the purposes, and gave us the data, described above, making it possible to say some things about young children's development that could not otherwise be said. These are not, though, by any means, the last words on the issues raised here. Valsiner (1987), Thelen and Smith (1994) and others, Nelson among them, encourage us not to be content with our research and theories until we have described the very process of change we are committed to

studying. A number of reasons and means to do that became clear as this study progressed.

The limitation of a cross-sectional study in which individuals are placed in groups and their behavior is averaged, was felt most strongly as the investigator sought to assess the relative understanding children had of "self" and "other." The global outcome, measured across children, indicated that psychological understanding of "self" and of the "other" were not distinguishable from each other. In the moment in which they were studied, children tended to respond to questions, or describe experiences, similarly, whether the subject was "I" or "s/he." That global finding, however, covered a multitude of subtleties in the data. The inability to correctly answer a question about the "self," for example, was more indicative of immaturity than was the inability to respond correctly to a question about the "other." The percentage of "self" references used in the game recounting predicts the immaturity composite score while percentage of "other" references does not. We saw in the game recountings that "self" reference is a kind of floor, or starting point, in this study. And while "self" and "other" use are correlated, they are of course, not perfectly correlated.

In other words, there is evidence in this study that "self" and "other" understanding do not march lock step through development. These data prompt the investigator to hypothesize that, given longitudinal data we would find that children's construction of a psychological concept begins with its relationship to the "self," and that the child's ability to use that same concept in reference to the "other" is an indication, to use dynamic systems language, that the concept has stabilized.

Thelen and Smith (1992) explain that, according to dynamic systems theory, one knows that a system is in a stable state when the behavioral degrees of freedom are compressed and the behavior of the organism is consistent and predictable. The predictability of answers across "self" and "other" questions in the deceptive box task tells us, systems theorists would say, that the child's concept of false belief is in a stable state. That concept has many hypothesized elements: understanding that "think" is nonfactive, mastery of the syntactic forms of the question as asked, understanding that seeing leads to knowing what's in the box, understanding that mental states can misrepresent reality, and so on. What we see when we observe a stable state is that the elements or this concept are "cooperating," giving "rise to behavior with a unitary character" (Thelen & Smith, 1987, page xix). The dynamic elements continue to interact, but harmoniously, and in a predictable manner.

Stability, say Thelen and Smith, then is not terribly informative to those who want to study the process of change. Its predictability gives us the "illusion of structure" (Thelen & Smith, 1987, page xix), and lulls us into believing that we are observing a static phenomenon. We might be too ready to declare, for example, that self and other understanding are equivalent for young children.

For one interested in the relationship between "self" and "other" understanding (of false belief, for example), one needs to know about process, and particularly here, whether self and other understanding are linked in development. Do they change together across time? Does change in one prompt a change in the other? It would be especially telling if such a link were demonstrated in moments of instability, when all of the conceptual pieces of a particular puzzle are not yet in place. If a question were

asked that confused the child's notion of her own belief, or some aspect of it, would his notion of that aspect of the other's belief change in response? Answers to such questions would help us to determine whether "self" and "other" understanding develop by the same process and at the same rate. Longitudinal studies and studies of children who are in a moment of dynamic change have much to offer us as we continue this exploration.

Finally, it should be said that there is much left to learn about the relationship of narrative to children's psychological understanding of "self" and "other" and "us." Providing children with mental state words in the context of a story prompted some children to use them in their own retellings. While that might be precisely what one would expect, it stands in striking contrast to the absence of mental, and inner states in general, from children's recounting of the game. In the game, too, the investigator provided mental state words, but the words, in that context, did not prompt children to use them. Many possible reasons for that are discussed above.

In summary, though, it should be said that though this study does not provide the data needed to make the case, the possibility remains open that narrative serves a special function in children's developing psychological understanding. Narrative may reunite aspects of experience that are isolated in language. Objects, people, and events are related to one another; the relationships between people's inner and outer worlds are made explicit for children. Children may be learning to integrate Bruner's "landscape of action" and the "landscape of consciousness." The whole, which is narrative, may prove to be, for children's developing psychological understanding,

much more than the sum of its parts. It may bring a sense of wholeness and linguistic coherence to children's understanding of their own and others' experience.

There is no end to the questions that can be posed about the relationships among person use, psychological understanding, language and narrative. Fortunately this study is just the beginning.

Appendix A

"The Dragon's Tears"

Far away in a strange country there lived a dragon. And the dragon's home was a deep mountain cave; from the cave his eyes shone out like headlights.

Very often, when the people living nearby were gathered in the evening by the fire, one would say: "What a terrible dragon is living near us. And another would agree, saying, "Someone should kill him."

Whenever children were told about the dragon, they were frightened. But there was one little boy who was never frightened. All the neighbors said: "Isn't he a funny little boy?"

When it was almost time for this funny little boy's birthday, his mother asked him: "Whom would you like to invite for your birthday?" Then that little boy said: "Mother, I would like to ask the dragon." His mother was very much surprised and asked: "Are you joking?" "No," said the little boy very seriously, "I mean what I say; I want to invite the dragon."

And sure enough, on the day before his birthday, the little boy stole quietly out of his house. He walked and walked and he walked till he reached the mountain where the dragon lived.

"Hello. Hello. Mr. Dragon," the little boy called down the valley in his loudest voice. "What's the matter? Who's calling me?" rumbled the dragon, coming out of his cave. Then the little boy said: "Tomorrow is my birthday and there will be lots of good things to eat, so please come to my party. I came all the way to invite you."

At first the dragon couldn't believe his ears and kept roaring at the boy. But the boy wasn't frightened at all and kept saying: "Please, Mr. Dragon, please come to my party." Finally the dragon understood that the boy meant what he said and was actually asking him, a dragon, to his birthday party. Then the dragon stopped roaring and began to cry, "What a happy thing to happen to me." The dragon sobbed. "I never had a kind invitation from anyone before." The dragon's tears flowed and flowed until at last they became a river.

Then the dragon said: "Come, climb on my back and I'll give you a ride home." The boy climbed bravely onto the back of the ferocious dragon and away the dragon went, swimming down the river of his own tears.

AT FIRST THE DRAGON COULDN'T BELIEVE HIS EARS AND KEPT ROARING AT THE BOY. BUT THE BOY WASN'T FRIGHTENED AT ALL AND KEPT SAYING: "PLEASE, MR. DRAGON PLEASE COME TO MY PARTY."

APPENDIX B: Sample Game Recountings
(on-task utterances only; references to investigator, animals, and indefinite 'you' were not coded)

GP = investigator used a general prompt
SP = investigator pointed to a particular item and prompted child to talk about it
S = support ("uh huh" or repetition)
U = experimenter expresses confusion or asks for repetition

SELF ONLY REFERRERS

1)	GP	what happen that?
	SP/book	that one.
	S	I had that one [pointing to animal puppet]
	GP (9)	[child is touching and looking at the store]
	SP/store	the monies?
	SP/small table	table!
2)	GP	got food.
	GP	a animal.
	GP	food.
	GP	I see <have a> last time [pointing to goat]
		This one.
	SP/boxes	two dollars in there.
	GP	store.
	SP/store	put money in there.
3)	GP	a octopus one
	GP	I only had one [animal].
	SP/store	that's the store.
	SP	food there.
4)	GP (several)	
	SP/book	that's me right there [pointing to parrot].
	SP/boxes	he [reference to parrot, not 'other] eats this one.
	S	seeds
		he eats chicken
	GP	this one there's [?] nothing.
	SP	the money's in there.
	GP	the seeds.

'WE' ONLY REFERRERS

1)	GP	a chipmunk.
	SP	food.
	GP	set the table.
	GP	buy money.
	GP	put them in a chair.
	SP/book	these [pointing to acorns]
	SP/boxes	we picked food.
	SP/store	we opened it up.
	S	and put the money in there so we could buy the food.
	GP	we set the table.
2)	GP	you took them out.
		and then you pick.
		and then you let us take one
	GP	we taked it
		and then we had a animal
		and we let it feed
		and we feed it
		and then when they're ready to eat the food
		they put the plates on the co...
		got the tea
		and then they got the coffee
	GP	plates
		the knives
		the cups and stuff
		knives
		all that stuff
	SP/boxes	take one and [unintelligible]
		that one there was nothing in them
		there was something in this one
		and then we opened this one [touching the box]
		and then something was in there
	SP/store	we took the food
		and then we ate it all
		nuts
	GP	and then we took the cars and we ride them.
3)	GP	we looked in one of those boxes.
	GP	we went over to the store
		and then we drove back
		we went to buy some food
		and then we came back and ate some snack

we looked in there to see what kind of food they eated.

4)	GP	there was none in there.
	GP	there was some in there.
	GP	that's where you get the money.
	SP/table	then we sit at the table and eat.
	GP	truck to drive home.

TWO-PERSON REFERRERS

1)	GP	a goat and a parrot and a squirrel and an octopus
	GP	they eat
	SP/boxes	we didn't find my food in one of those
		and the leaves.
	SP/store	I paid for my food
	SP/table	eat
		they ate
2)	GP	[Partner's name] got to go first.
	SP/book	first is that one [pointing to each of the animals on the cover in turn]
		second is this one.
		third is that one and fifth is this one
	GP	it was fun
		then [Partner's name] took the seeds
		then I took the plants.
	SP/store	it's a store and it has different dollars in it.
	SP/table	there was lunch
	S	they eat it all.
3)	GP	I pick one
	S	and [Partner's name] picked one.
	SP/book	I was that one.
		she was that one.
		and [Partner's name] was that one.
	SP	[Partner's name] ws that one and that one.
		this one
		these two
		I was this
	GP	[Partner's name] had a squirrel and the...
	GP	I had [false start] got the shell
	SP/store	I got the shell from here
	SP/ table	there was food at the table.

4)	GP	eat
	S/GP	went to the food store
	GP	eat
	GP	drive
	SP/cards	we took some puppets
	SP/book	we took some puppets on that
	SP/boxes	there was some money
	SP/store	eat
		then we looked in the food store
		[Partner's name] [false start] [Partner's name] opened up the box
		and there was nothing
	S	so he took one dollar right?

REFERRERS TO ALL THREE PERSONS

1)	GP	I beed the one
	GP	and [partner's name] beed the two.
	GP	we start to play the game
	GP	I had this one [pointing to the parrot]
	GP	[Partner's name] beed the chipmunk.
	SP/boxes	there were seeds in the boxes
		and there was money to feed the animals
		and then I took some food like uh
		not seashell
		not nuts
		I took nuts!
		and I put them in the bowl
		and then they started to eat
	GP	well I put my money in here yesterday
		and then I got out some seeds
		and then there was this and not this
		just those [unintelligible]
		and then I tooked [?] seeds

2)	GP	and then we had a fight
	GP	and we had to close our eyes
		And you put a card on our hands
		And I had a two and [partner's name] had a one.
		And [partner's name] got to go first.
		She got the octopus
		I got the goat.
		And I [false start] she got the money
		And I [false start] and she gave me money when I wanted to buy my food
		and I looked in the box

		and there was nothing there
		and um then we sat down at the table
		and then we ate
		and we put the stuff on the table
		and we did [false start] and we had some ice cream.
3)	GP	[Partner's name] got this one
		and I got this one
		and I winned.
	SP	we play a house
	GP	we could play some
		Then we go downstairs in the Munchkin [class' name] room
	SP/book	octopus eat [false start] uh octopus eat shell
	SP/boxes	[Partner's name] didn't get any.
	SP/store	take the dollar and put in there and help [Partner's name]
	SP/table	you put you put [false start] I put the octopus right there.
4)	GP	[Partner's name] picked the first card
		the one
		this is the one that whichever one picked
		we picked
	S	we picked these
	S	and I was that one [referring to the chipmunk]
		And he want to pick that one [referring to the parrot]
	SP/book	then we had to find [flipping pages] food.
	S	this one.
	GP	[Partner's name] [false start] [Partner's name] looked over there
		And he saw like seeds.
		And he saw one that [unintelligible' on it that showed what his animal ate
		and then he went to the store
	GP	he buy his food
	GP	set up the table
	SP/boxes	I remember that [partner's name] he looked in those boxes and [unintelligible] all money.

APPENDIX C

Outline of Game Actions

I. Picking the cards
   - A. Partners decide who will pick first
   - B. They make selections
     1. Player A ("A") chooses number 1
     2. Player B ("B") chooses number 2

II. Choosing the puppet animals
   - A. A chooses animal X
     - i. points to animal X on card 1
     - ii. investigator puts it on A's finger
   - B. B chooses animal Y
     - i. points to animal Y on card 2
     - ii. investigator puts it on B's finger

III. Naming the animals
   - A. A names animal X
   - B. B names animal Y

IV. Looking in the book for appropriate food for animals

A (or B) A sees her/his animal's food

B (or A) B sees her/his animal's food

V. Finding box contents
   - A. Player A opens box
     - i. A goes to the box with food X on top
     - ii. A answers questions about what s/he thinks is in the box
     - iii. A opens box and finds money in it
   - B. Player B opens box
     - i. B goes to the box with food X on top
     - ii. B answers questions about what s/he thinks is inside

iii. B opens box and finds nothing inside

VI. Looking in the store

A. A gets food X

i. A goes to store and opens it

ii. A sees appropriate food

iii. A receives instruction from investigator

iv. A puts money in pocket

v. A takes food out

B. B gets food Y

i. B goes to store and opens it

ii. B sees appropriate food

iii. B receives instruction from investigator

iv. B receives money from A

v. puts money in pocket

vi. B takes food out

VII. Taking animals home from store

A. A drives car X

B. B drives car Y

VIII. Setting table for snack

A. They put plates, utensils, and food on table

B. They put the animals in their chairs

IX. Animals eat their snack

APPENDIX D

VARIABLE NAMES AND DEFINITIONS

DT = Dragon's Tears retelling task
GR = Game Recounting task
RB = Raisin box task
SQ = Standardized Questions

Age group	1 = 3.6 to 4 years 2 = 4 years to 4.6 3 = 4.6 to 5 years
Age in months	age in months
DT composite score	Dragon's Tears composite score derived from the variables that loaded at .5 or higher on DT factor (converted to percentages)
DT # on-task utterances	number of utterances that were part of the child's Dragon's Tears story retelling
DT emotion word use	child's use of mental state terms in retelling of Dragon's Tears 0 = no emotion words used 1 = one emotion word used 2 = two or more emotion state words used
DT # of acts (of 27)	number of actions (of 27) the child mentioned in the retelling of the Dragon's Tears
DT % subject-verb	percentage of verbs in the Dragon's Tears retelling that were in the appropriate tense, and agree with subject
DT variety of connectives	variety of connectives used in the Dragon's Tears on a Guttman scale 0 = no connectives used 1 = "and" only 2 = ("and" +) "then" 3 = either of above + "so" and/or "because" and/or "but" 4 = any of above + "when" 5 = any of above + "where"

DT mental state use	child's use of mental state terms in retelling of Dragon's Tears 0 = no mental states used 1 = one mental state word used 2 = two or more mental state words used
DT psychological content level	psychological content the child mentioned in retelling the story of the Dragon's Tears 1 = action only 2 = action + [non-epistemic] psychological states 3 = 2 + epistemic states
Expressive	TELD subscore for expressive language
GR composite score	Game recounting composite score derived from the variables that loaded at .5 or higher on DT factor (converted to percentages)
GR # acts recounted (of 9)	number of (the nine main) game actions child recounted
GR # in-sequence acts	longest string of in-sequence actions recounted, uninterrupted by investigator prompts
GR # on-task utterances	number of on-task utterances
GR # 'other' references	number of references to the experience of the child's partner in the game
GR # 'self' references	number of references to the game playing experience of the child her/himself
GR # 'we/us' references	number of references to the pair's experience in the game
GR % autobiography	% of game subject-verb units that are in the past tense and have a specific subject
GR % on-task utterances	percentage of utterances in child's game recounting that was a reference to "yesterday's" game
GR % person ref = other	percentage of utterances in child's game recounting that was reference to the child's game partner

GR % person ref = self	percentage of utterances in child's game recounting that was reference to the self
GR 'other' psychological content level	psychological content of game recounting references to the other child 1 = action only 2 = action + [non-epistemic] psychological states
GR 'self' psychological content level	psychological content of game recounting references to the self 1 = action only 2 = action + [non-epistemic] psychological states
GR 'we/us' psychological content level	psychological content of 'us/we' references in game recounting 1 = action only 2 = action + [non-epistemic] psychological states
GR variety of connectives	variety of connectives used, on Guttman scale 1 = no connectives used 2 = "and" only 3 = ("and" +) "then" 4 = either of above + "but" and/or "so" and/or "because" 5 = any of above + "when"
Immaturity (IM) composite	This is derived from the three most highly-loading variables on factor 4, called the Immaturity factor. All three variables are from the game recounting.
Language group	1 = low language scorers 2 = medium language scorers 3 = high language scorers
Narrative (NA) composite	This is derived from the six variables that relate to children's ability to structure their storytelling in Dragon's Tears factor and the game recounting factor (i.e., they use three of the four variables that went into DT and GR composites)
NA composite groups	children's scores on the narrative composite were divided into three groups: 1 = low narrative scores 2 = medium narrative scores 3 = high narrative scores

Other use group	children who in their game recounting made at least one reference to the other child
Raisin Box deceptive box	child's performance on the raisin box ('self' and 'other' think) questions 0 = fail (got both wrong) 1 = in transition (got one right) 2 = pass (got both right)
RB 'self' think	child's response to the Raisin Box task question, "Before I opened the box... what did you think was in it?" 0 = incorrect 1 = correct
RB 'other' think	child's response to the Raisin Box task question, "If your teacher came in right now and saw the box, closed just like this, what would she think is in the box?" 0 = incorrect 1 = correct
Receptive	TELD subscore for receptive language
Self Only group	game recounting - children who referred exclusively to the experience of the self
Semantic	TELD subscore for semantic language
SO composite score	composite score derived from the variables that loaded on the "self-other answer" factor at the .5 level or higher (converted into percentages)
Standardized Questions deceptive box	child's performance on the SQ (embedded in game and asked about in SQ) deceptive box ('self' and 'other' think) questions 0 = fail (got both wrong) 1 = in transition (got one right) 2 = pass (got both right)
SQ & RB 'other' think	0 = got neither SQ or RB other think question correct 1 = got one or the other correct, but not both 2 = got both questions right
SQ + RB 'self' think	0 = got neither SQ or RB 'self' think question correct 1 = got one or the other correct, but not both

	2 = got both questions right
SQ # correct 'self' questions	number of standardized questions about the experience of the self in the game that were answered correctly
SQ # correct 'other' questions	number of standardized questions about the experience of the other child in the game that were answered correctly
SQ 'self' psychological level	level of psychological content of correct 'self' answers 1 = action only 2 = action + [non-epistemic] psychological state 3 = action + epistemic state
SQ 'self' emotion	child's answer to the standardized question, "How did you feel when you saw what was in the box?" 0 = incorrect 1 = correct
SQ 'other' emotion	child's answer to the standardized question "How did s/he feel when s/he saw what was in the box?" 0 = incorrect 1 = correct
SQ 'other' think	child's answer to the standardized question about the game, "Before [partner's name] opened the box, what did s/he think was in it?" 0 = incorrect 1 = correct
SQ 'self' think	child's answer to the standardized question about the game, "Before you opened the box... what did you think was in it?" 0 = incorrect 1 = correct
Syntactic	TELD subscore for syntactic language
TELD raw score	raw (language) score for TELD-2 test
'We' only group	children who, in the game recounting, referred exclusively to the experience of the pair (making no reference to self or other as individuals)
'We' use group	children who, in the game recounting, made at least one reference to the experience of the pair

APPENDIX E

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
age in mos	50	42.00	60.00	51.2556	5.6097
teld raw score	50	30	57	48.62	6.83
GR # on-task utt	50	4	29	13.10	5.85
GR acts (of 9)	50	1	8	4.68	1.79
GR # in-seq acts	50	1	10	3.92	2.29
GR % on-task utt	50	27	100	74.76	19.47
GR % autobio	50	0	100	73.30	27.71
GR conn variety	50	0	4	1.86	1.11
GR # 'self' refs	50	0	12	3.04	2.59
GR # 'other' refs	50	0	10	2.24	2.35
GR # 'we' refs	50	0	17	3.58	3.63
GR % pers=self	50	0	100	38.94	28.18
GR % pers=other	50	0	71	22.56	20.62
GR % pers= 'we'	50	0	100	38.51	30.19
GR 'self' psych level	50	0	2	1.22	.65
GR 'other' psych level	50	0	2	.86	.64
GR 'we' psych level	50	0	2	.98	.62
DT # on-task utterances	50	4	46	16.68	8.50
DT act (of 27)	50	0	20	7.86	4.77
DT % subj-vb agreement	50	0	98	47.34	31.35
DT conn variety	50	0	5	1.98	1.35
DT PSYCH LEVEL	50	1	3	1.96	.75
DT mental state use	50	0	2	.32	.59
DT emotion word use	50	0	2	.72	.73
SQ 'self ' # correct	50	2	9	6.58	1.46
SQ 'other' # correct	50	2	8	6.12	1.29
SQ self psych level	50	1	3	1.90	.76
SQ other psych level	50	1	3	1.90	.79
SQ 'self' feel	50	0	1	.52	.50
SQ 'other' feel	50	0	1	.48	.50
SQ 'self' think	50	0	1	.28	.45
SQ 'other' think	50	0	1	.26	.44
RB 'self' think	50	0	1	.52	.50
RB 'other' think	50	0	1	.44	.50
SQ + RB 'self' think	50	0	2	.76	.77
SQ + RB 'other' think	50	0	2	.70	.81
SQ DB	50	0	2	.50	.71
RB DB	50	0	2	.96	.92
DT composite	50	1.00	95.99	35.8374	23.0985
SO composite	50	14.93	99.98	59.1345	25.4039
GR composite	50	6.17	68.83	38.3865	16.9441
IM composite	50	18.00	87.71	51.6233	17.1986
NA composite	50	6.67	86.75	45.0733	18.8737
Valid N (listwise)	50				

	1	2	3	4	5	6	7	8	9	10	11
1. Age in months		533**	.141	.002	.377**	.263	.235	.177	.088	-.055	.179
2. TELD raw score			.416**	.037	.449**	.416**	.173	.143	.024	-.025	.387**
3. GR acts recounted (of 9)				.646**	.537**	.713**	-.050	.544**	.377**	.321*	.564**
4. GR # on-task utterances					.471**	.554**	-.065	.494**	.430**	.562**	.427**
5. GR % on-task utterances						.413**	-.051	.277	.227	.151	.312*
6. GR # in-sequence acts							.115	.550**	.329*	.322*	.664*
7. GR % autobiography								-.164	.227	.183	.136
8. GR connective variety									.230	.400**	.479**
9. GR # 'self' references										.505**	.029
10. GR # 'other' references											.142
11. GR # 'we' references											

Appendix F Table 1. Spearman's rho correlations of 38 variables

N = 50

PAGE 1 OF 10

* = $p < .05$
** = $p < .01$
*** = $p < .001$

	GR % person reference = 'self'	GR % person reference = 'other'	GR % person reference = 'we'	GR 'self' psychological content level	GR 'other' psychological content level	GR 'we' psychological content level	DT # on-task utterances	DT acts (of 27)	DT % subj-verb agree
1	-.130	.007	.197	-.093	-.161	.224	.126	.385**	.394**
2	-.244	-.035	.371**	.002	-.085	.497**	-.072	.306*	.267
3	-.202	.094	.277	.153	.326*	.389**	.000	.409**	.253
4	-.210	.336*	.030	.142	.447**	.113	.005	.166	.081
5	-.116	.045	.159	-.032	.119	.264	.027	.217	.072
6	-.286*	.061	.341*	.199	.254	.476**	.046	.359*	.175
7	.029	.120	-.046	.147	.067	.092	.368**	.209	.236
8	-.334*	.229	.275	.028	.319*	.389**	.087	.363**	.264
9	.506**	.216	-.394**	.625**	.404**	-.115	.275	.271	.222
10	-.152	.865**	-.280*	.406**	.824**	-.042	.097	.023	-.007
11	-.627**	-.118	.789**	-.100	.092	.725**	-.004	.438**	.284*

Appendix F Table 1. Spearman's rho correlations of 38 variables

N = 50

* = $p < .05$, ** = $p < .01$, *** = $p < .001$

1 = age in months
2 = TELD (language) score
3 = GR acts recounted (of 9)
4 = GR # of on-task utterances
5 = GR % on-task utterances
6 = GR # in-sequence actions
7 = GR % of autobiography
8 = GR variety of connectives
9 = GR # of 'self' references
10 = GR # 'other' references
11 = GR # 'we' references

	DT connective variety	DT psychological content level	DT mental state use	DT emotion word use	SQ # correct 'self' questions	SQ # correct 'other' questions	SQ 'self' psych. level	SQ 'other' psych. level
1	.216	.344*	.380**	.307*	.340*	.197	.127	.258
2	.189	.160	.405**	.089	.499**	.414**	.325*	.448*
3	.231	.126	.113	.004	.154	.096	.012	.209
4	.188	-.024	-.070	-.052	-.055	.031	-.238	-.056
5	.114	-.013	.044	-.150	.304*	.255	.032	.192
6	.252	.034	.154	-.166	.090	-.026	-.034	.051
7	.077	.207	.212	.120	-.087	.195	-.185	-.175
8	.308*	.104	.092	.101	.183	.013	.048	.196
9	.222	-.033	-.061	-.015	-.137	-.014	-.166	-.186
10	-.007	-.136	-.142	.037	.058	.032	-.144	-.123
11	.284*	.173	.207	.021	.242	.125	.147	.183

Appendix F Table 1. Spearman's rho correlations of 38 variables
N = 50
* $p < .05$, ** $p < .01$, *** $p < .001$

1= age in months
2 = TELD (language) score
3 = GR acts recounted (of 9)
4 = GR # of on-task utterances
5 =GR % on-task utterances
6 = GR # in-sequence actions
7 = GR % of autobiography
8 = GR variety of connectives
9 = GR # of 'self' references
10 = GR # 'other' references
11 = GR # 'we' references

	SQ 'self' emotion	SQ 'other' emotion	SQ 'self' think	SQ 'other' think	RB 'self' think	RB 'other' think	SQ + RB 'self' think	SQ + RB 'other' think	SQ deceptive box task	RB deceptive box task
1	.018	.176	.017	.232	.089	.096	.100	.144	.234	.104
2	.196	.142	.128	.462**	.339*	.293*	.374**	.421**	.446**	.346*
3	.172	.228	-.179	.117	.312*	.204	.210	.194	.077	.283*
4	.064	.028	-.353*	-.046	.093	.148	-.050	.084	-.154	.128
5	.003	.082	-.073	.179	.228	.272	.199	.283*	.149	.272
6	.171	.182	-.206	-.037	.149	.045	.068	.008	-.042	.107
7	-.091	-.139	-.165	-.056	.058	.114	-.049	.052	-.087	.095
8	.282*	.141	-.262	.154	.196	.240	.072	.221	.030	.238
9	.197	.083	-.328*	-.213	-.058	.079	-.167	-.047	-.298*	.009
10	-.033	-.061	-.233	-.111	.058	.091	-.018	.028	-.122	.081
11	.200	.196	-.048	.121	.289*	.247	.247	.223	.183	.293*

Appendix F Table 1. Spearman's rho correlations of 38 variables
N= 50
* $p < .05$, ** $p < .01$, *** $p < .001$

1 = age in months
2 = TELD (language) score
3 = GR acts recounted (of 9)
4 = GR # of on-task utterances
5 =GR % on-task utterances
6 = GR # in-sequence actions
7 = GR % of autobiography
8 = GR variety of connectives
9 = GR # of 'self' references
10 = GR # 'other' references
11 = GR # 'we' references

	12	13	14	15	16	17	18	19	20	21
12. GR % person reference = 'self'		-.189	-.754**	.486**	-.124	-.545**	.125	-.147	-.119	-.180
13. GR % person reference = 'other'			-.360*	.259	.799**	-.143	-.031	-.154	-.049	-.186
14. GR % person reference = we/us'				-.473**	-.231	.702**	-.110	.335*	.246	.260
15. GR 'self' psych. content level					.453**	-.111	-.029	.054	.060	-.029
16. GR 'other' psych. content level						-.062	-.037	.019	-.017	-.096
17. GR 'we/us' psych. content level							-.124	.367**	.261	.245
18. DT # on-task utterances								.526**	.416**	.559**
19. DT # acts (of 27)									.796**	.629**
20. DT % subj-vb agreement										.585**
21. DT variety of connectives										

Appendix F Table 1. Spearman's rho correlations of 38 variables

N= 50

* p < .05, ** p < .01, *** p < .001

	DT psych. content level	DT mental state use	DT emotion word use	SQ # correct 'self'	SQ # correct 'other'	SQ 'self' psych. level	SQ 'other' psych. level	SQ 'self' emotion	SQ 'other' emotion	SQ 'self' think	SQ 'other' think
12	-.118	-.163	-.059	-.345*	-.184	-.324*	-.249	-.070	-.128	-.128	-.195
13	-.134	-.126	.097	.213	.120	-.078	-.063	-.087	-.166	-.161	.006
14	.189	.238	.040	.301*	.139	.390**	.298*	.195	.245	.139	.194
15	-.148	-.142	-.040	-.058	-.131	-.158	-.123	.085	-.136	-.076	-.058
16	-.134	-.157	.000	.124	.020	-.068	-.108	-.017	-.174	-.154	-.080
17	.122	.224	.016	.302*	.173	.173	.325*	.229	.162	-.052	.314*
18	.493**	.412**	.306*	-.186	-.080	.024	-.212	.133	.015	-.101	-.193
19	.733**	.668**	.468**	.210	.126	.267	.195	.408**	.273	-.063	.081
20	.647**	.629**	.416**	.271	.192	.182	.241	.442**	.153	-.212	.219
21	.614**	.592**	.347*	.063	.128	.047	.069	.309*	.150	-.123	.052

Appendix F Table 1. Spearman's rho correlations of 38 variables
N= 50
* $p < .05$, ** $p < .01$, *** $p < .001$

12. GR % person reference = 'self'
13. GR % person reference = 'other'
14. GR % person reference = 'we/us'
15. GR 'self' psychological level
16. GR 'other' psychological level
17. GR 'we/us' psychological level
18. DT # on-task utterances
19. DT # acts (of 27)
20. DT % subject-verb agreement
21. DT variety of connectives

	RB 'self' think	RB 'other' think	SQ + RB 'self' think	SQ + RB 'other' think	SQ deceptive box task	RB deceptive box task
12	.289*	-.221	-.398**	-.245	-.356*	-.327*
13	-.371**	.089	.021	.073	-.018	.083
14	.065	.200	.383**	.226	.346*	.298*
15	.340*	-.051	-.033	-.051	-.118	-.021
16	.017	.085	.077	.034	-.050	.109
17	.118	.160	.215	.262	.312*	.216
18	.229	-.049	-.182	-.145	-.143	-.118
19	-.168	.112	.101	.093	.113	.133
20	.129	.186	.137	.196	.105	.251
21	.272	.091	.048	.062	-.031	.151

Appendix F Table 1. Spearman's rho correlations of 38 variables
N= 50
* $p < .05$, ** $p < .01$, *** $p < .001$

12. GR % person reference = 'self'
13. GR % person reference = 'other'
14. GR % person reference = 'we/us'
15. GR 'self' psychological level
16. GR 'other' psychological level
17. GR 'we/us' psychological level
18. DT # on-task utterances
19. DT # acts (of 27)
20. DT % subject-verb agreement
21. DT variety of connectives

	22	23	24	25	26	27	28	29	30	31
22. DT psychological content level		.808**	.610**	.153	.355*	.175	.160	.213	.207	.033
23. DT mental state use			.371**	.313*	.380**	.286*	.366**	.286*	.333*	.064
24. DT emotion state use				.159	.085	.061	.039	.148	.063	.070
25. SQ # 'self' correct					.467**	.585**	.574**	.364**	.221	.283*
26. SQ # 'other' correct						.273	.438**	.154	.259	.101
27. SQ 'self' psych. content level							.540**	.438**	.303*	.644**
28. SQ 'other' psych. content level								.304*	.461**	.248
29. SQ 'self' emotion									.522**	-.203
30. SQ 'other' emotion										-.064
31. SQ 'self' think										

Appendix F Table 1. Spearman's rho correlations of 38 variables

N= 50

* $p < .05$, ** $p < .01$, *** $p < .001$

	SQ 'other' think	RB 'self' think	RB 'other' think	SQ + RB 'self' think	SQ + RB 'other' think	SQ deceptive box task	RB deceptive box task
22	.086	.161	.208	.109	.158	.110	.202
23	.260	.225	.188	.212	.229	.268	.226
24	.048	-.021	.021	-.046	.011	.050	.002
25	.480**	.354*	.428**	.466**	.512**	.525**	.427**
26	.427**	.484**	.661**	.446**	.654**	.410**	.623**
27	.316*	.401**	.388**	.678**	.414**	.674**	.430**
28	.809**	.439**	.467**	.507**	.698**	.744**	.494**
29	.022	.119	.206	.036	.137	-.026	.178
30	-.022	.122	.116	.102	.055	.076	.133
31	.240	.243	.165	.612**	.228	.673**	.221

Appendix F Table 1. Spearman's rho correlations of 38 variables
N= 50
* $p < .05$, ** $p < .01$, *** $p < .001$

22. DT psychological content
23. DT mental state use
24. DT emotion state use
25. SQ # correct 'self' answers
26. SQ # correct 'other' answers
27. SQ 'self' psychological content
28. SQ 'other' psychological content
29. SQ 'self' emotion
30. SQ 'other' emotion
31. SQ 'self' think

	32	33	34	35	36	37	38
32. SQ 'other' think		.387**	.485**	.436**	.795**	.805**	.474**
33. RB 'self' think			.690**	.885**	.662**	.491**	.922**
34. RB 'other' think				.614**	.909**	.475**	.916**
35. SQ + RB 'self' think					.636**	.739**	.818**
36. SQ + RB 'other' think						.703**	.852**
37. SQ deceptive box							.526**
38. RB deceptive box							

Appendix F Table 1. Spearman's rho correlations of 38 variables

N= 50

* p < .05, ** p < .01, *** p < .001

	1	2	3	4	5	6	7	8	9	10	11
1. Age in months		.589**	.214	-.025	.331*	.272	.217	.256	.039	-.123	.131
2. TELD raw score			.377**	.106	.512**	.412**	.333*	.287*	.041	.052	.302*
3. GR acts recounted (of 9)				.633**	.565**	.670**	.158	.606**	.350*	.251	.532*
4. GR # on-task utterances					.456**	.501**	.119	.500**	.383**	.461**	.571*
5. GR % on-task utterances						.377**	.123	.307*	.221	.120	.322*
6. GR # in-sequence acts							.207	.566**	.306*	.253	.591**
7. GR % autobiography								.026	.235	.279*	.185
8. GR connective variety									.265	.319*	.412**
9. GR # 'self' references										.293*	-.076
10. GR # 'other' references											.043
11. GR # 'we' references											

Appendix F Table 2. Pearson product moment correlations of 38 variables

N = 50

PAGE 1 OF 10

* = $p < .05$
** = $p < .01$
*** = $p < .001$

	GR % person reference = 'self'	GR % person reference = 'other'	GR % person reference = 'we'	GR 'self' psychological content level	GR 'other' psychological content level	GR 'we' psychological content level	DT # on-task utterances	DT acts (of 27)	DT % subj-verb agree
1	-.241	.018	.212	-.085	-.169	.245	.149	.413**	.427**
2	-.429**	.083	.344*	-.018	.039	.512**	-.008	.375**	.309*
3	-.338*	.069	.269	.132	.335*	.416**	-.020	.394**	.308*
4	-.249	.261	.055	.134	.397**	.141	-.054	.145	.093
5	-.243	.108	.153	-.027	.181	.296	.014	.237	.149
6	-.300*	.000	.281*	.246	.215	.456**	-.037	.339*	.192
7	-.142	.215	-.014	.142	.199	.166	.270	.250	.261
8	-.419**	.174	.272	.044	.347*	.411**	.143	.348*	.327*
9	.415**	.068	-.433**	.553**	.348*	-.176	.120	.246	.215
10	-.234	.784**	-.317*	.380**	.688**	.017	.009	-.081	-.068
11	-.544*	-.225	.671**	-.125	.036	.592**	.009	.414*	.210

Appendix

Appendix F Table 2. Pearson product moment correlations of 38 variables

N = 50

* = $p < .05$, ** = $p < .01$, *** = $p < .001$

PAGE 2 OF 10

1 = age in months
2 = TELD (language) score
3 = GR acts recounted (of 9)
4 = GR # of on-task utterances
5 = GR % on-task utterances
6 = GR # in-sequence actions
7 = GR % of autobiography
8 = GR variety of connectives
9 = GR # of 'self' references
10 = GR # 'other' references
11 = GR # 'we' references

	DT connective variety	DT psychological content level	DT mental state use	DT emotion word use	SQ # correct 'self' questions	SQ # correct 'other' questions	SQ 'self' psych. level	SQ 'other' psych. level
1	.226	.359*	.436**	.323*	.386**	.254	.143	.282*
2	.179	.155	.362**	.117	.586**	.534**	.314*	.447**
3	.251	.126	.099	-.023	.292*	.194	.051	.223
4	.202	-.004	-.099	-.027	.086	.066	-.194	-.029
5	.122	.015	.093	-.186	.425**	.376**	.107	.255
6	.257	.033	.126	-.172	.161	-.004	-.075	.029
7	.086	.126	.190	.094	.115	.267	-.131	-.098
8	.340*	.115	.102	.077	.279*	.069	.031	.194
9	.099	.001	-.102	-.048	-.093	.017	-.153	-.217
10	-.037	-.190	-.160	.004	.089	.058	-.168	-.053
11	.353*	.217	.189	-.038	.263	.124	.125	.156

Appendix F Table 2. Pearson product moment correlations of 38 variables

N = 50

* $p < .05$, ** $p < .01$, *** $p < .001$

PAGE 3 OF 10

1 = age in months
2 = TELD (language) score
3 = GR acts recounted (of 9)
4 = GR # of on-task utterances
5 = GR % on-task utterances
6 = GR # in-sequence actions
7 = GR % of autobiography
8 = GR variety of connectives
9 = GR # of 'self' references
10 = GR # 'other' references
11 = GR # 'we' references

	SQ 'self' emotion	SQ 'other' emotion	SQ 'self' think	SQ 'other' think	RB 'self' think	RB 'other' think	SQ + RB 'self' think	SQ + RB 'other' think	SQ deceptive box task	RB deceptive box task
1	.035	.186	.013	.248	.105	.104	.136	.199	.229	.114
2	.207	.143	.042	.438**	.366**	.324*	.366*	.438**	.412**	.376**
3	.188	.219	-.189	.133	.346*	.228	.224	.213	.081	.313*
4	.072	-.010	-.288*	-.002	.120	.193	-.022	.118	-.111	.171
5	.099	.140	-.122	.199	.255	.323*	.212	.307*	.174	.314*
6	.142	.157	-.213	-.019	.107	.013	.012	-.002	-.076	.066
7	.014	-.166	-.216	.035	.121	.196	-.004	.140	-.068	.172
8	.279*	.123	-.286*	.159	.206	.224	.055	.224	.013	.234
9	.202	.047	-.322*	-.222	-.079	.065	-.199	-.081	-.300*	-.008
10	-.039	-.030	-.217	-.061	.082	.064	-.024	.006	-.123	.080
11	.110	.101	.048	.120	.244	.249	.240	.219	.163	.268

Appendix F Table 2. Pearson product moment correlations of 38 variables
N= 50
* $p < .05$, ** $p < .01$, *** $p < .001$

1 = age in months
2 = TELD (language) score
3 = GR acts recounted (of 9)
4 = GR # of on-task utterances
5 =GR % on-task utterances
6 = GR # in-sequence actions
7 = GR % of autobiography
8 = GR variety of connectives
9 = GR # of 'self' references
10 = GR # 'other' references
11 = GR # 'we' references

	12	13	14	15	16	17	18	19	20	21
12. GR % person reference = 'self'		-.265	-.752**	.427**	-.212	-.608**	-.002	-.198	-.178	-.147
13. GR % person reference = 'other'			-.436**	.232	.703**	-.189	.004	-.239	-.051	-.185
14. GR % person reference = we/us'				-.557**	-.282*	.657**	-.001	.349*	.201	.263
15. GR 'self' psych. content level					.470**	-.141	-.094	.010	.067	-.018
16. GR 'other' psych. content level						-.058	-.065	-.033	-.022	-.098
17. GR 'we/us' psych. content level							-.013	.412**	.242	.267
18. DT # on-task utterances								.496**	.442**	.572**
19. DT # acts (of 27)									.767**	.635**
20. DT % subj-vb agreement										.583**
21. DT variety of connectives										

Appendix F Table 2. Pearson product moment correlations of 38 variables

N= 50

* p < .05, ** p < .01, *** p < .001

	DT psych. content level	DT mental state use	DT emotion word use	SQ # correct 'self'	SQ # correct 'other'	SQ 'self' psych. level	SQ 'other' psych. level	SQ 'self' emotion	SQ 'other' emotion	SQ 'self' think	SQ 'other' think
12	-.098	-.177	-.042	-.444**	-.294*	-.305*	-.258	-.077	-.103	-.039	-.213
13	-.150	-.064	.062	.238	.158	-.105	-.038	-.142	-.193	-.148	.036
14	.195	.209	-.002	.251	.166	.357*	.267	.170	.228	.137	.174
15	-.149	-.135	-.040	.013	-.155	-.161	-.116	.080	-.142	-.075	-.061
16	-.139	-.150	.002	.220	.070	-.071	-.109	-.023	-.167	-.144	-.085
17	.129	.185	-.013	.350*	.232	.168	.328*	.229	.161	-.052	.315*
18	.488**	.459**	.249	-.042	.052	.108	-.020	.163	-.002	-.045	.044
19	.741**	.658**	.434**	.267	.209	.288*	.213	.413**	.308*	-.019	.085
20	.635**	.585**	.365**	.314*	.245	.170	.248	.443**	.164	-.208	.211
21	.621**	.602**	.306*	.131	.131	.038	.094	.316*	.164	-.124	.077

Appendix F Table 2. Pearson product moment correlations of 38 variables

N= 50

* $p < .05$, ** $p < .01$, *** $p < .001$

12. GR % person reference = 'self'
13. GR % person reference = 'other'
14. GR % person reference = 'we/us'
15. GR 'self' psychological level
16. GR 'other' psychological level
17. GR 'we/us' psychological level
18. DT # on-task utterances
19. DT # acts (of 27)
20. DT % subject-verb agreement
21. DT variety of connectives

	RB 'self' think	RB 'other' Think	SQ + RB 'self' think	SQ + RB 'other' think	SQ deceptive box task	RB deceptive box task
12	-.350*	-.220	-.367**	-.252	-.284*	-.310*
13	.061	.051	.011	.051	-.009	.061
14	.284*	.170	.334*	.199	.270	.247
15	.017	-.053	-.056	-.066	-.111	-.019
16	.104	.069	.055	-.004	-.068	.094
17	.229	.160	.202	.270	.255	.211
18	-.127	.058	-.096	.060	.014	-.038
19	.133	.086	.135	.099	.106	.119
20	.285*	.192	.132	.233	.072	.260
21	.196	.104	.054	.106	-.032	.163

Appendix F Table 2. Pearson product moment correlations of 38 variables
N= 50
*** $p < .05$, ** $p < .01$, *** $p < .001$**

12. GR % person reference = 'self'
13. GR % person reference = 'other'
14. GR % person reference = 'we/us'
15. GR 'self' psychological level
16. GR 'other' psychological level
17. GR 'we/us' psychological level
18. DT # on-task utterances
19. DT # acts (of 27)
20. DT % subject-verb agreement
21. DT variety of connectives

	22	23	24	25	26	27	28	29	30	31
22. DT psychological content level		.766**	.572**	.151	.341*	.170	.164	.216	.212	.033
23. DT mental state use			.309*	.303*	.353*	.301*	.335*	.254	.298*	.117
24. DT emotion state use				.175	.080	.022	.021	.126	.040	.057
25. SQ # 'self' correct					.484**	.512**	.513**	.583*	.169	.243
26. SQ # 'other' correct						.303*	.454**	.185	.286*	.046
27. SQ 'self' psych. content level							.526**	.403**	.286*	.673**
28. SQ 'other' psych. content level								.287*	.431**	.251
29. SQ 'self' emotion									.522**	-.203
30. SQ 'other' emotion										-.064
31. SQ 'self' think										

Appendix F Table 2. Pearson product moment correlations of 38 variables

N= 50

* $p < .05$, ** $p < .01$, *** $p < .001$

	SQ 'other' think	RB 'self' think	RB 'other' think	SQ + RB 'self' think	SQ + RB 'other' think	SQ deceptive box task	RB deceptive box task
22	.093	.163	.209	.123	.179	.076	.202
23	.223	.254	.136	.263	.205	.246	.212
24	.040	-.040	.009	-.049	.027	.000	-.017
25	.457**	.331	.397**	.417**	.493**	.505**	.396**
26	.409**	.467**	.612**	.420**	.599**	.381**	.587**
27	.320*	.403**	.384**	.722**	.411**	.700**	.428**
28	.835**	.441**	.475**	.497**	.746**	.750**	.498**
29	.022	.119	.206	.013	.139	-.057	.177
30	-.022	.122	.116	.092	.060	.000	.129
31	.240	.234	.165	.663**	.232	.700**	.222

Appendix F Table 2. Pearson product moment correlations of 38 variables

N= 50

* $p < .05$, ** $p < .01$, *** $p < .001$

22. DT psychological content
23. DT mental state use
24. DT emotion state use
25. SQ # correct 'self' answers
26. SQ # correct 'other' answers
27. SQ 'self' psychological content
28. SQ 'other' psychological content
29. SQ 'self' emotion
30. SQ 'other' emotion
31. SQ 'self' think

	32	33	34	35	36	37	38
32. SQ 'other' think		.387**	.485**	.425**	.843**	.814**	.474**
33. RB 'self' think			.690**	.852**	.636**	.458**	.920**
34. RB 'other' think				.596**	.880**	.460**	.919**
35. SQ + RB 'self' think					.598**	.749**	.788**
36. SQ + RB 'other' think						.726**	.824**
37. SQ deceptive box							.499**
38. RB deceptive box							

Appendix F Table 2. Pearson product moment correlations of 38 variables

N= 50

* $p < .05$, ** $p < .01$, *** $p < .001$

APPENDIX G: Factor Analysis Results

Goodness-of-fit Test

Chi-Square	df	Sig.
100.271	101	.502

Structure Matrix

	Factor			
	1	2	3	4
DT act (of 27)	.889	-.434	.197	.182
DT % subj-vb agreement	.818	-.254	.307	.205
DT mental state use	.754	-.323	.302	-.224
DT conn variety	.678	-.325	.166	.131
DT emotion word use	.515			-.112
GR # in-seq acts	.111	-.840		.521
LANGUAGE GROUP	.231	-.676	.531	-.274
GR acts (of 9)	.187	-.649	.360	.632
GR # 'we' refs	.237	-.644	.276	.268
GR % on-task utt		-.524	.460	.242
AGE GROUP	.343	-.431	.269	-.113
SQ + RB 'other' think	.122	-.235	.856	
SQ 'other' # correct	.237	-.249	.707	-.152
RB 'self' think	.148	-.250	.703	
SQ 'self ' # correct	.261	-.393	.624	
GR conn variety	.213	-.491	.269	.604
GR # 'self' refs	.114	-.124		.546
SQ 'self' think			.208	-.452
GR # 'other' refs	-.168	-.113		.426

Extraction Method: Maximum Likelihood.
Rotation Method: Oblimin with Kaiser Normalization.

Factor Correlation Matrix

Factor	1	2	3	4
1	1.000	-.241	.196	-1.437E-02
2	-.241	1.000	-.375	-.215
3	.196	-.375	1.000	-5.903E-02
4	-1.437E-02	-.215	-5.903E-02	1.000

Extraction Method: Maximum Likelihood.
Rotation Method: Oblimin with Kaiser Normalization.

REFERENCES

Abbeduto, L., & Rosenberg, S. (1985). Children's knowledge of the presuppositions of know and other cognitive verbs. Journal of Child Language, 12, 621-641.

Astington, J. W. (1990). Narrative and the child's theory of mind. In B. K. Britton & A. D. Pellegrini (Eds.), Narrative thought and narrative language (pp. 151-171). Hillsdale, NJ: Lawrence Erlbaum Associates.

Baldwin, J. M. (1906). Social and ethical interpretations in mental development. (4th ed.). London: Macmillan.

Bartsch, K. & Wellman, H. M. (1995). Children talk about the mind. New York: Oxford University Press.

Bowlby, J. (1980). Attachment and loss. (Vol. 3). New York: Basic.

Bretherton, I. (1985). Attachment theory: Retrospect and prospect. In I. Bretherton & E. Waters (Eds.), Growing points of attachment theory and research (Vol. 50, nos. 1-2, serial no. 209, pp. 3-38).

Bretherton, I., & Beeghly, M. (1982). Talking about internal states: The acquisiton of an explicit theory of mind. Developmental Psychology, 18, 906-921.

Brown, A. L., & Smiley, S. S. (1977). Rating the importance of structural units of prose passages: A problem of metacognitive development. Child Development, 48, 1-8.

Bruner, J. (1986). Actual minds, possible worlds. Cambridge, MA: Harvard University Press.

Bruner, J. (1990). Acts of meaning. Cambridge, MA: Harvard University Press.

Carlson, S. M., Moses, L. J., & Hix, H. R. (1998). The role of inhibitory processes in young children's difficulties with deception and false belief. Child Development, 69, 672-691.

Carey, S. (1985). Conceptual change in childhood. Cambridge, MA: MIT Press.

Carlson, S. M., Moses, L. J., & Hix, H. R. (1998). The role of inhibitory processes in young children's difficulties with deception and false belief. Child Development, 69, 672-691.

Case, R. (1985). Intellectual Development. Orlando, FL: Academic Press.

Case, R. (1991). The mind's staircase: Exploring the conceptual underpinnings of children's thought and knowledge. Hillsdale, NJ: Lawrence Erlbaum Associates.

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

Cicchetti, D., & Beegly, M. (1990a). Perspectives on the study of self in transition. In D. B. Cicchetti, M. (Ed.), The self in transition: Infancy to childhood (pp. 1-15). Chicago: University of Chicago Press.

Cicchetti, D., & Beegly, M. (1990b). The self in transition: Infancy to childhood. Chicago: University of Chicago Press.

Charman, T. & Shmueli-Goetz, Y. (1998). The relationship between theory of mind, language, and narrative discourse: An experimental study. Cahiers de Psychologie Cognitive, 17, 245-271.

Damon, W., & Hart, D. (1988). Self-understanding in childhood and adolescence. New York: Cambridge University Press.

Dennett, D. C. (1978). Beliefs about beliefs. Behavioral and Brain Sciences, 1, 568-570.

Donald, M. (1991). Origins of the modern mind: Three stages in the evolution of culture and cognition. Cambridge, MA: Harvard University Press.

Fivush, R., & Hamond, N. R. (1990). Autobiographical memory across the preschool years: Toward conceptualizing childhood amnesia. In R. Fivush & J. A. Hudson (Eds.), Knowing and remembering in young children (pp. 223-248). New York: Cambridge University Press.

Fivush, R. (1991). The social construction of personal memories. Merrill Palmer Quarterly, 37, 59-82.

Flavell, J. H. (1985). Cognitive development. (2nd ed.) Englewood Cliffs, NJ: Prentice-Hall.

Flavell, J. H. (1988). The development of children's knowledge about the mind: From cognitive connections to mental representations. In J. Astington, P. Harris, & D. Olson (Eds.), Developing theories of mind. New York: Cambridge University Press.

Flavell, J. H. (1992). Perspectives on perspective taking. In H. P. Beilin, P. B. Pufall (Eds.). Piaget's theory: Prospects and possibilities. Hillsdale, NJ: Lawrence Erlbaum Associates.

French, L. A. (1988). The development of children's understanding of "because" and "so." Journal of Experimental Psychology, 45, 262-279.

French, L. A. & Pak, M. K. (1995). Young children's play dialogues with mothers and peers. In K. E. Nelson & Z. Reger (Eds.) Children's language, vol. 8 (pp. 65-101). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. Cognitive Development, 10, 483-527.

Gopnik, A. (1993). How we know our minds: The illusion of first-person knowledge of intentionality. Behavioral and Brain Sciences, 16, 1-14.

Gopnik, A., & Wellman, H. (1992). Why the child's theory of mind really is a theory. Mind and Language, 7(1-2), 145-171.

Gopnik, A., & Wellman, H. (1994). The theory theory. In L. A. Hirschfeld & S. A. Gelman (Eds.), Mapping the mind (pp. 257-293). New York: Cambridge University Press.

Griffin, S. (1991). Young children's awareness of their inner world: A neo-structuralist analysis of the development of intrapersonal intelligence. In R. Case (Ed.), The Mind's Staircase (pp. 189-206). Hillsdale, NJ: Lawrence Erlbaum Associates.

Hadwin, J. & Perner, J. (1991). Pleased and surprised: Children's cognitive theory of emotion. British Journal of Developmental Psychology, 9, 215-234.

Harris, P. L. (1991). The work of the imagination. In A. Whiten (Ed.), Natural theories of mind: Evolution, development and simulation of everyday mindreading (pp. 283-304). Cambridge, MA: Basil Blackwell.

Harris, P. L. (1994). Thinking by children and scientists: False analogies and neglected similarities. In L. A. Hirschfeld & S. A. Gelman (Eds.) Mapping the mind.

New York: Cambridge University Press.

Harter, S. (1996). Developmental changes in self-understanding. In A. J. Sameroff & M. M. Haith (Eds.) The five to seven year shift: The age of reason and responsibility (pp. 207-236). Chicago: The University of Chicago.

Hirshfeld, L. A. & Gelman, S. A. (Eds.) (1994). Mapping the mind: Domain specificity in cognition and culture. New York: Cambridge University Press.

Hresko, W., Reid, D., Hammill, D. (1981). The Test of Early Language Development. Austin, TX: Pro-Ed.

Hudson, J. A. (1990). The emergence of autobiographical memory in mother-child conversations. In R. Fivush & J. A. Hudson (Eds.), Knowing and remembering in young children (pp. 166-196). New York: Cambridge University Press.

Hudson, J. A. & Shapiro, L. R. (1991). From knowing to telling: The development of children's scripts, stories, & personal narratives. In McCabe, A. & Peterson, C. (Eds.), Developing narrative structure (pp. 89 – 136). Hillsdale, NJ: Lawrence Erlbaum Associates.

Hudson, J. A. & Slackman, E. (1986). The acquisition and development of scripts. In K. Nelson (Ed.), Event knowledge: Structure and function in development (pp. 71-96). Hillsdale, NJ: Lawrence Erlbaum Associates.

James, W. (1985). Psychology: The briefer course. Notre Dame, Indiana: University of Notre Dame Press. (Original work published 1892).

Johnson, C. (1988). Theory of mind and the structure of conscious experience. In P. Harris & D. Olson (Eds.), Developing theories of mind (pp. 47-63). New York: Cambridge University Press.

Johnson, C. & Maratsos, M. P. (1977). Early comprehension of mental verbs: Think and know. Child Development, 48, 1743-1747.

Karmiloff-Smith, A. (1992). Beyond Modularity. Cambridge, MA: MIT Press.

Kohlberg, L. (1969). Stage and sequence: The cognitive-developmental approach to socialization. In D. A. Goslin (Ed.), Handbook of socialization theory and research. Chicago: Rand McNally.

Kozulin, A. (1986). Vygotsky in context. In A. Kozulin (Ed.), Thought and language (pp. Xi-lxi). Cambridge, MA: MIT Press.

Laks, D. R., Beckwith, L., & Cohen, S. E. (1989). Mothers' use of personal pronouns when talking with toddlers. The Journal of Genetic Psychology, 151, pp. 25-32.

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

Leslie, A. M. & Roth, D. (1993). What autism teaches us about metarepresentation. In S. Baron-Cohen, H. Tager-Flusberg, & D. Cohen (Eds.). Understanding other minds: Perspectives from autism (pp. 83-111). New York: Oxford University Press.

Levy, E. & Nelson, K. (1994). Words in discourse: A dialectical approach to the acquisition of meaning and use. Journal of Child Language, 21, 367-389.

Lewis, C. (1994). Episodes, event, and narratives in the child's understanding of mind. In C. Lewis & P. Mitchell (Eds.), Children's early understanding of mind: Origins and development (pp. 457-480). Hillsdale, NJ: Lawrence Erlbaum Associates.

McCabe, A. (1991). Structure as a way of understanding. In A. McCabe & C. Peterson (Eds.). Developing narrative structure (pp. ix-xvii). Hillsdale, NJ: Lawrence Erlbaum Associates.

McCabe, A. & Peterson, C. (Eds.) (1991). Developing narrative structure. Hillsdale, NJ: Lawrence Erlbaum Associates.

McLaren, R. & Olson, D. (1993). Trick or treat: Children's understanding of surprise. Cognitive Development, 8, 27-46.

MacWhinney, Brian (1995). The CHILDES project: Tools for analyzing talk (2nd ed). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Mead, G. H. (1934). Mind, self, and society from the standpoint of a social behaviorist. (Vol. 1) Chicago: University of Chicago.

Mehan, H. (1979). What time is it, Denise? Asking known information questions in classroom discourse. Theory into Practice, 18, 285-294.

Menig-Peterson, C. L. (1975). The modification of communicative behavior in preschool-aged children as a function of the listener's perspective. Child Development, 46, 1015-1018.

Moore, C., Bryant, D., & Furrow, D. (1989). Mental terms and the development of certainty. Child Development, 60, 167-171.

Moore, C., Pure, K., & Furrow, D. (1990). Children's understanding of the modal expression of speaker certainty and uncertainty and its relation to the development of a representational theory of mind. Child Development, 61, 722-730.

Murphy, G., & Medin, D. L. (1985). The role of theories in conceptual coherence. Psychological Review, 92, 289-316.

Nelson, K. (1974). Concept, word, and sentence: Interrelations in acquisition and development. Psychological Review, 81, 267-285.

Nelson, K. (Ed.) (1986). Event knowledge: Structure and function in development. Hillsdale, NJ: Lawrence Erlbaum Associates.

Nelson, K. (1989). Monologue as the linguistic construction of self in time. In K. Nelson (Ed.) Narratives from the crib. Cambridge, MA: Harvard University Press.

Nelson, K. (1993). The psychological and social origins of autobiographical memory. Psychological Science, 4, 7-14.

Nelson, K. (1996). Language in cognitive development: The emergence of the mediated mind. Cambridge: Cambridge University Press.

Nelson, K. (1997). Finding one's self in time. In J. G. Snodgrass, R.L. Thompson (Eds.) The self across psychology: Self recognition, self-awareness, and the self concept. Annals of the New York Academy of Sciences, 818. New York: New York Academy of Sciences.

Nelson, K., Plesa, D., & Henseler, S. (1998). Children's theory of mind: An experiential interpretation. Human Development, 41, 7-29.

Niesser, U. (1988). Five kinds of self knowledge. Philosophical Psychology, I, 35-59.

Neisser, U. (Ed.) (1993). The perceived self: Ecological and interpersonal sources of self-knowledge. (Vol. 5). Cambridge: Cambridge University Press.

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

Peterson, C. & McCabe, A. (1991). Linking children's connective use and narrative macrostructure. In McCabe, A. & Peterson, C. (Eds.). Developing narrative structure (pp. 29-53). Hillsdale, NJ: Lawrence Erlbaum Associates.

Piaget, J. (1955). The language and thought of the child. Cleveland: Meridian. (Original work published 1923).

Piaget, J. (1962). Play, dreams and imitation in childhood. New York: Norton & Co. (Original work published 1945).

Plesa, D., Goldman, S., & Edmondson, D. (1995, April). Negotiation of meaning in a false belief task. Poster presented at the biennial meeting of the Society for Research in Child Development, Indianapolis, IN.

Rosch, E. (1975). Cognitive representation of semantic categories. Journal of Experimental Psychology: General, 104, 192-233.

Russell, J. (1996). Agency: Its role in mental development. Hove, U.K.: Erlbaum.

Sakade, F. (1957). Japanese children's stories. Rutland, VT: Tuttle.

Selman, R. (1980). The growth of interpersonal understanding. New York: Academic Press.

Shatz, M., & Gelman, R. (1973). The development of communication skills: Modifications in the speech of young children as a function of the listener. Monographs of the Society for Research in Child Development, 38 (5, Serial No. 152).

Shatz, M., Wellman, H. M., & Silber, S. (1983). The acquisition of mental verbs: A systematic investigation of the first reference to mental state. Cognition, 14, 301-321.

Smith, M. C. (1978). Cognizing the behavior stream: The recognition of intentional action. Child Development, 49, 736-743.

Spelke, E. S. (1991). Physical knowledge in infancy: Reflections on Piaget's theory. In S. G. Carey, R. Gelman (Ed.), The epigenesis of mind: Essays on biology and cognition (pp. 133-170). Hillsdale, NJ: Lawrence Erlbaum Associates.

Thelen, E. & Smith, L. B. (1994). A dynamic systems approach to the development of cognition and action. Cambridge, Massachusetts: The MIT Press.

Tomasello, M. (1993). On the interpersonal origins of self-concept. In U. Neisser (Ed.), The perceived self: Ecological and interpersonal sources of self-knowledge (Vol. 5, pp. 174-184). Cambridge: Cambridge University Press.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. Cognitive psychology, 5, 207-232.

Valsiner, J. (1987). Culture and the development of children's actions: A cultural-historical theory of developmental psychology. Chichester, England: John Wiley & Sons.

Vygotsky, L. (1978). Mind in society. Cambridge, MA: Harvard University Press.

Vygotsky, L. (1979). Consciousness as a problem of psychology of behavior. Soviet Psychology, 17.

Vygotsky, L. (1981). The genesis of higher mental functions. In J. V. Wertsch (Ed.), The concept of activity in Soviet psychology. Armonk, NY: Sharpe.

Vygotsky, L. (1986). Thought and language. (Kozulin, A., Trans.). Cambridge, MA: MIT Press. (Original work published 1934).

Wellman, H. (1990). The child's theory of mind. Cambridge, MA: MIT press.

Wellman, H. & Banerjee, M. (1991). Mind and emotion: Children's understanding of the emotional consequences of beliefs and desires. British Journal of Developmental Psychology, 9, 191-214.

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

Wimmer, H. & Hartl, M. (1991). Against the Cartesian view of mind: Young children's difficulty with own false belief. British Journal of Developmental Psychology, 9, 125-138.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. Cognition, 13. 103-128.

Wood, M. E. (1978). Children's developing understanding of other people's motives for behavior. Developmental Psychology, 14, 561-562.