Children's understanding of relevancy violations and theory of mind acquisition

Kenneth Brian Smyth

Submitted for the Degree of

Doctor of Philosophy

Department of Psychology
University of Strathclyde

November 2009

Violations of conversational relevance (Grice, 1975) may take a variety of conventionalised forms, including references to topics involving out of sight objects, past events, and emotions and beliefs that are not immediately apparent. Past qualitative research (Ninio & Snow, 1996), supported by examination of the CHILDES Language Database (MacWhinney, 2001), indicates that young children's understanding of such references emerges sequentially, in this order, between the ages of 18 months and 4 years, as each is promoted in turn by parents introducing them and providing conversational, historical and psychological support for their interpretation. It is argued that this process might provide the basis for an account of the acquisition and development of children's theory of mind through the mechanism of conversation, since it would effectively provide training not just in the pragmatics of managing conversations, but in reading the intentions and perspectives of others. In order to examine this possibility further, four studies were conducted with children aged 1.5 to 4.0 years (Studies 1 and 2) and 2.5 to 4.0 years (Studies 3 and 4) with the objectives of 1) establishing whether the sequence of emergence of understanding each type of reference is consistent with Ninio & Snow's account; 2) examining how far these competences are related to wider language ability, working memory and the management of processing demands, and performance on different forms of false belief task. Each study employed a request task of the type devised by Babelot & Marcos (1999), in which children's ability to identify and retrieve toys from the immediate environment in response to experimenter request was compared to that when these requests referred to toys out of sight behind the child (Study 1), toys that had formed the focus of a shared past event (Study 2), toys that were the subject of a non-evident emotion (Study 3), and toys that had non-evident beliefs (Study 4). As a further check on competence, in each study requests of both kinds varied between direct ('give me the...') and indirect forms ('have you got the...?'). This made it possible to determine how far understanding of each reference type was secure enough to permit management of these additional pragmatic demands. Study 1 found that children aged 3.5 to 4.0 years showed good grasp of references to out of sight objects, as indexed by both performance and management of the different request forms, but even those aged 1.5 to 2.0 years showed signs of competence. In contrast, Study 2 found that older children showed restricted competence in responding appropriately to references to a shared past event, with younger children performing at no better than chance levels. Studies 3 and 4 found no difference between the extent of grasp of references to evident and non-evident emotions and beliefs, but with children aged 3.5 to 4.0 performing better than those aged 2.5 to 3.0 years, and at levels of competence that were higher than those apparent among children of the same age in Study 2, contrary to prediction. As hypothesised, performance on the request task in Studies 3 and 4 was related to performance on a corresponding false belief task (the 'look first' manipulation of Siegal & Beattie, 1991), but there were also ubiquitous relationships between grasp of non-immediate references and measures of expressive language and working memory, as well as relationships to age that were not explained by any other directly measured variable. Overall, in line with the conversational account, the research suggests children's development is a gradual, incremental progression rather than the result of sudden conceptual insight. However, the relationship to other factors and the departure from the sequence of emergence described by Ninio & Snow indicate that whilst conversational practice may play an important role in the emergence of pragmatic ability and theory of mind, this is only one influence in a complex network of internal and external processes of development. Future research needs to concentrate on investigating these processes in greater depth.

Acknowledgements

This PhD thesis could not have been written without the help and inspiration of many people, both within and outside of academia, and I am pleased to acknowledge those who have made it possible here.

~~~~~

Firstly, I could not have wished for better supervisors, and I am deeply indebted in this regard to Professor Andy Tolmie for his great insight, guidance, support, encouragement and faith in me.

A major debt of gratitude is also owed to Professor Christine Howe, my second supervisor, whose insight, advice and help was invaluable to me.

Many thanks go to Dr Alison Sanford, Director of Postgraduate Studies, Dr Stephen Butler, Dr Doreen Thompson, Dr Birgit Schroeter and Mrs. Patricia Gallagher of the University of Strathclyde for their help and contribution.

I would also like to acknowledge my supervisors in the MSc in Psychological Therapy in Primary Care, 2006-2007, University of Dundee and Stirling, and many thanks are extended to Dr Mary Howley, Clinical Psychologist, NHS Greater Glasgow & Clyde and to Dr Will Goodall, Course Coordinator, University of Stirling. Their support of me to persevere with work on both the PhD and MSc was of great encouragement in what was a particularly busy year for me.

~~~~~

A special thanks must go to my wife, Janet whose selflessness and unwavering faith in me over the years has made my dream of a second career in Psychology possible and I will always be deeply grateful and indebted for her support and inspiration.

My son Ryan, for his expert computer knowledge, and daughter Emma, youngest son, Jamie, and son-in-law Robert are all deserving of my gratitude for their constant support. As is my 5-year-old granddaughter, Caitlin who never fails to remind me that Developmental Psychology is the most fascinating and rewarding, yet unpredictable of subjects.

A thank you is also due to my mum, Isobel Smyth and younger brother Gordon, who despite living in Northern Ireland have continued to offer great support for my studies from across the sea and during their visits to Scotland.

~~~~~

I would also like to thank Emma Watson, who helped me in the early pilot studies, and I also find myself indebted to the children and staff from the various nurseries in Glasgow and North Ayrshire that I visited. Their enthusiasm, when participating or assisting in the studies, was always a great source of encouragement to me.

Finally, my debt of gratitude is also extended to both the Department of Psychology, University of Strathclyde, and also to the Economic and Social Research Council whose funding made the research and PhD possible.

~~~~~

Contents

Children's understanding of relevancy violations and theory of mind acquisition

Chapter 1

Literature review

1.1 Rationale and the research background

The rationale for this thesis emanates from a growing body of work that challenges traditional viewpoints on how children come to acquire a theory of mind. By and large, pre-school children's lack of success on theory of mind tests in experimental settings e.g. the standard unexpected transfer test (Wimmer & Perner, 1983) or the unexpected contents task (UCT) (Perner, Leekam & Wimmer, 1987) has long been attributed to the absence of a key conceptual insight. These children are claimed to lack knowledge of how the intentions, beliefs and desires of others arise, and can be distinct from their own. More recent research challenges such claims, reinterpreting children's performance in terms of their developing understanding of pragmatic aspects of language i.e. the management of topics of conversation (Siegal, 1993; Siegal & Peterson, 1994; Garfield, Peterson & Perry, 2001), and arguing that such understanding is integral to theory of mind abilities more generally.

With the purpose of identifying how a theory of mind may be acquired, Flavell (1999) poses two questions. Firstly, what behaviours do infants of different ages show that seem relevant to the development of knowledge about people? Second, how should these behaviours be interpreted? That is, exactly how much and what kind of knowledge about the mind (if any) should we attribute to infants and young children who exhibit such behaviours?

In response to the first question, this thesis will argue that young children show behaviours that imply their insight into other's mental states is acquired through, and deployed within a developing grasp of a variety of aspects of the pragmatics of conversational practices. In answer to the second question, this thesis will also contend that theory of mind acquisition should not be interpreted as a unitary conceptual achievement in isolation from language. That is, the pragmatic skills that children exhibit should be interpreted as evidence that a growing knowledge and understanding of other people's minds is in place early on in pre-school children.

This chapter commences by considering the nature of theory of mind, with discussion being focused by four theories that are pre-eminent in the debate on the acquisition of such understanding: functional viewpoints, modular approaches, theory theory accounts, and social constructivist theory. The literature review will then move towards an outline of the growth of pragmatic language understanding between adults and children, drawing on the CHILDES Language Database (MacWhinney, 2001). Adult support for pragmatic language understanding and its implications for performance on false belief tests will also be examined before an outline of the research and hypotheses are presented.

1.2 The nature of theory of mind

Theory of mind has engaged researchers since Premack & Woodruff's (1978) paper stimulated debate by defining a theory of mind as an ability to impute mental states to others and to the self that permits interpersonal understanding. Their research culminated with the controversial claim that some non-human primates can attribute the behaviour of a protagonist to an intended goal, and so may possess a 'theory of mind,' at least in a primitive form – although critics argued that non-linguistic creatures cannot possess mind reading abilities, suggesting that the results of primate studies may have reflected chance or non-mentalistic processes such as associative learning (Heyes, 1993; 1998).

Children's theory of mind appears much more sophisticated but no less contentious. Theoretical divisions over its emergence and development are endemic, with four main accounts being offered: (1) functional accounts focused on 'lean' or 'rich' interpretation of adult-child behaviour and language, especially the meaning of acts of reference; (2) modular accounts, focused on the idea that theory of mind derives from an innate structure that exhibits maturational development; (3) representational or theory theory accounts, that posit a radical conceptual shift in children's thinking that occurs at around 4 years of age; and (4) social constructivist theory, which argues that mind reading abilities derive from culturally influenced interactions with caregivers. Each of these is examined in turn below.

1.3 Lean and rich interpretation accounts

Functional accounts argue that children's knowledge concerning people derives from infants' attempts at reference, and at understanding others' references in the first two years of life. One strand of research has investigated how growth of understanding of mental states may be causally linked to perceptual inputs. Butterworth & Jarrett (1991) argue that infants show an awareness of embodied minds through direct perception i.e. they are pre-attuned to the informational potential of certain aspects of others' behaviour. The 18 month old child exhibits understanding that their own visual field is shared with that of caregivers, for instance, by turning to look in the same direction as their mother's gaze. Infants' gaze at this age was found to be highly calibrated with their mother's gaze, demonstrating that something more than imitation or reinforcement might be at work, and suggesting that they have developed geometrical co-ordination to identify the location of the mothers' line-of-regard. Butterworth & Jarrett adopt a lean interpretation of these findings, contending that a child may simply recognise that changes in the direction of their mother's gaze are good predictors of where an object might be located, and that to credit children with an abstract theory of mind per se on the basis of these abilities may be unwarranted, although they may form an important precursor.

In contrast, Baldwin & Moses (1994) propose a richer interpretation of children's perceptual abilities, within which infants aged 14 to 19 months are so highly attuned to a speaker's referential action that it guides new word-object mapping and facilitates the development of correct vocabulary. More recent research with older infants by Diesendruck, Markson, Akhtar & Reudor (2004) shows that 2-year-olds appear to monitor the intentional character of references carefully, treating words associated with deliberate actions differently to those associated with accidental actions.

Baldwin & Moses (1994) also extended theorising about joint attention into the socio-emotional domain, where emotions, akin to words, have a referential character, as one is happy, sad or frightened about an object, action or event. They note that infants as young as 12 months can grasp the relevance of adult's attentional cues and locate the intended referent of a caregiver's emotional signal. Social referencing of emotion is also supported by Campos & Stenberg (1980), who found that mothers only had to display an emotional reaction in the presence of a novel object, not actually to look at it, to achieve an effect on infants' reaction to that object i.e. the child attaches a perceived shift in emotional tone to the presence of the novel object.

Based on their rich interpretation stance, Baldwin & Moses propose that infants are developing a fledgling insight into theory of mind, by performing a mentalistic analysis of others' communicative and emotional behaviour. They go on to suggest that "Infants' budding linguistic knowledge might provide an especially strong impetus to construct a more powerful theory of mind" (1994, p. 152). This view is supported by Baron–Cohen (1993), who proposes that infants possess 'attention-goal psychology', where the ability to interpret an adult's attentional cue establishes an understanding in the child that the caregiver is interested in an external object.

1.4 Parental and child rich interpretation

Baldwin & Moses' (1994) theory of rich interpretation concentrates on children's understanding of attention and referential intent towards currently accessible objects. However, adults' everyday discourse with each other, and with children, is beset with intentional references to intangible objects that are either spatially distant or temporally absent (e.g. a parental reference to an older child or sibling who has gone to school, or to the family dog out of view in the back garden). Baldwin & Moses acknowledge that it remains largely uncharted how children come to comprehend that attention and thought can persist and allow reference to objects not perceptibly present in the immediate surroundings, and they recognise that the study of such development may pave the way towards a better understanding of their mental state attribution. The understanding of such references is therefore the focus of the present research.

The intimate connection between interpersonal communication and shared knowledge was first acknowledged by Baldwin (1884; 1906), and later by Rommetveit (1971, 1972, 1979a, 1979b), who note that messages which occur out of context are ambiguous statements that adult conversationalists can only make sense of by drawing upon implicit shared representations. Expanding upon this, Clark (1977) noted that most everyday conversation is elliptical, and that successful communication depends on the listener's prior knowledge and their ability to make a 'bridging inference' based on the minimal information given. Bretherton (1991) argued that these principles can also be applied to infant-adult communication: adults rely on such inferences – a key form of rich interpretation – to decode children's one word utterances and intentional gestures: "They [adults] use the timing and directiveness of infant signals in conjunction with shared presuppositions to make sense of what the infant is trying to convey" (Bretherton, 1991, p. 69.) The significance of this is noted by Lock (1978): by being taken to mean something, children learn *how* to mean in a controlled fashion.

Greenfield & Smith (1976) proposed that parental rich interpretation requires at least two basic assumptions about infants: (a) that infants perceive relations between agents and objects more or less like adults (making shared knowledge possible); and (b) that their gestures or words indicate important elements of the situation in which communication occurs (helping to establish an intersubjective framework). The research of Dunn (1994) indicates that these assumptions hold, and that infants make reciprocal assumptions about adults' understanding during communication. Dunn examined the quality and quantity of conversations between mothers and children in the home and found that children aged 18-24 months labelled a wide range of feeling states and made causal statements regarding their feelings to influence mothers. By 33 months of age children's talk about causality with adults concerned both overt behaviour and internal states.

The research of Bretherton (1993), Cassidy & Marvin (1992), Dunn (1994) and Marvin & Greenberg (1982) indicates that during the second year of life infant's rich interpretation of adult speech and behaviour allows them to build internal working models of rules for social behaviour and for interaction (cf. Cassidy & Shaver, 1999). This gradual development strongly suggests a functional grasp of the meanings of others that is taken to be a precursor to more generalised understanding of other minds.

1.5 Modular and theory theory viewpoints

The abilities noted by functional theorists are accorded a rather different status by modular theories. According to the latter, infants possess a 'theory of mind mechanism', an innate, predetermined structure that unfolds according to a maturational timetable, and whose operation has been likened to the Chomskian Language-Acquisition Mechanism (Fodor, 1992). Leslie (1987) proposed that this theory of mind mechanism (ToMM) comes online in children by the middle of their 2^{nd} year and is manifest in their understanding of pretence. Baron-Cohen & Ring (1994) endorsed Leslie's proposal by arguing for the existence of three other component modules of the mind reading system that function prior to the ToMM, during the infants' first 18 months.

Understanding of mental states within these 4 component modules occurs via the processing of what Leslie & Roth (1993) describe as M-Representations. From birth to 9 months, such processing is possible to a limited extent via the operation of an Intentionality Detector (ID), a primitive perceptual mechanism that interprets the actions of self–propelling stimuli, and builds representations indicating goal intentions, driven by desire. Within this same period, an Eye-Direction Detector (EDD) comes into operation. The function of this is to detect gaze, which is similarly interpreted in terms of desire. Between 9 and 18 months, a Shared Attention Mechanism (SAM) begins to function, the purpose of which is to check if the child themselves and another person are attending to the same object. The SAM is reliant upon the ID and EDD for input, using this to construct primitive m-representations of the form 'I see mother wants banana'. When the ToMM comes into play around 48 months, it fills out these representations by expanding 'wants' into a range of other possible mental connections to objects. With the fully developed TOMM comes the ability to impute mental states to the self and others (Baron-Cohen & Ring, 1994).

For 'theory theorists', this watershed for a representational theory of mind is marked by a child's ability to pass a series of false belief tests at around 4 years of age (Wimmer & Perner, 1983). To pass a false belief task requires an understanding of a sequence of events, which is revealed by answers to questions about the likely future actions of others. From the experimenter's viewpoint the purpose of asking these questions is to determine whether the child can represent the mistaken belief of others as a state of mind distinct from their own privileged knowledge. For example, in the classic ‘unexpected transfer task’ (Wimmer & Perner, 1983) a child listens to a story enacted using small dolls, in which a character places his/her chocolate in a blue cupboard and leaves the room. Unbeknown to this character, a second person then transfers the chocolate to a green cupboard. The participating child is asked to predict where the first character will look for the chocolate on his/her return.

If the listening child states that the first character would look in the blue (empty) cupboard, this is taken to indicate that they can represent that character’s false belief, and distinguish between this and what they knew had actually occurred. Wimmer & Perner (1983) detected a shift between the representational abilities of 3 and 4 year old children, with younger children failing to acknowledge the first character’s false belief, judging that they would look for the chocolate in its transfer location, the green cupboard.

To guard against the possibility that poor performance on the false belief test was due to 3 year old children’s difficulties in keeping track of a narrative chain of events, Perner, Leekam & Wimmer (1987) devised the unexpected contents test (UCT) or deceptive box test. In this, child participants were shown a ‘Smarties’ tube and asked what they thought it contained: ‘Smarties’ (chocolate sweets) was children’s customary answer. However, the child’s expectation was then confounded by the experimenter revealing that the tube contained crayons. The crayons were then returned to the ‘Smarties tube’ and the lid closed. Children were then asked what they thought the next child, who was outside the room (and therefore not aware of the unexpected contents), would think the tube contained. Three year old children typically gave a realist judgement by predicting that newcomers would think that it contained crayons.

A further test that children are assessed on is the appearance-reality task (Flavell, Flavell & Green, 1983). A child is shown an object that resembles a rock. However, when children handle it they soon discover that it is actually a sponge. When asked what the object is and what it looks like, 4 year old children are apt to respond correctly by stating that is a 'sponge' that looks like a 'rock'. In contrast, 3 year old children will answer 'sponge' to both questions indicating their difficulty with the concept of simultaneous contrasting representations, and a lack of grasp about how these can arise i.e. rock by appearance, sponge by touch.

According to theory theorists, this shift in performance is attributable to radical cognitive restructuring at around 4 years, possibly due to the ToMM switching on. This restructuring leads to a sudden, none to all shift in ability to understand such situations, and thus references to hidden mental states as well. Theory theorists agree that there are precursors to this (e.g. attention-goal concepts), but only in the sense that these provide building blocks, with limited functional value in themselves.

Theory theory has been challenged, however, with several studies suggesting that the age trend in passing/failing false belief tests may not be as sharp as predicted. Robinson & Mitchell (1995), for instance, conducted six investigations of the unexpected transfer task with various age ranges from 3 to 5 years. They failed to find an age trend in 4 of the 6 investigations. Significant age trends were only found in those experiments with the widest age range, e.g. 3.6 to 5.4 months. Similarly, Saltmarsh, Mitchell & Robinson (1995) in a study using deceptive box tests found that older children were no more likely to understand false belief than younger ones (age range 3.5 to 5.2 months). Furthermore, Stevenson (1995) found that a minority of children as old as 6 years failed to acknowledge false belief.

In a further experiment, Stevenson & Mitchell (1995b) produced evidence that current reality may retain a salience for children beyond their 4th year, grabbing their attention and hindering their ability to acknowledge false belief. Children aged 4 to 6 years who had passed a standard deceptive box task (or UCT) subsequently showed a current reality bias in a modified follow up, in which they were asked the question, "When you first saw this box you thought there was a pencil inside, didn't you?" According to Mitchell (1996), the vast majority of children agreed, incorrectly, with this statement matching current reality. Moreover, these responses were not attributable to children being motivated by the need to maintain their ability in the eyes of the experimenter, and accepting the suggestion that they knew all along what the contents of the tube were: the same pattern of responses was found when the design was extended to children considering the belief of a glove puppet, whose perceived ability should not matter to children. That children can pass a false belief test but later show a current reality bias presents considerable difficulties for theory theory. As Mitchell argues (1994, p. 41), "These findings provide strong evidence to suggest that how children answer false belief questions depends on what criterion (realist versus (vs.) representational) that they attend to".

On the strength of these investigations, Mitchell (1996) argues that developmentalists should not become fixated on the age at which children pass false belief tests, as they can only be either right or wrong, and there is no scope to detect shades of correctness between the extremes. This is a view bolstered by Newton, Reddy & Bull's (2000) research on everyday deceptions, which indicates a very poor correspondence between clear-cut cases of representational reasoning situated in conversations and performance on standard false belief tasks. Thus it would appear that when measures are used which are more sensitive to shades of correctness, standard tasks do not relate very well to these, and may in fact be biased towards identifying shift-like rather than gradual changes in cognition.

1.6 Social constructivist accounts

Thus far, we have seen that functional accounts of the child's acquisition of a theory of mind emphasise the role played by infants' early interactions with caregivers, which provide direct information on the thoughts and feelings of others, and a context within which the infant can reciprocate by providing similar information to those others. The child is seen as building up a broader mentalistic understanding by generalisation from these more specific experiences, though the precise processes involved are unclear. Whilst functional accounts point effectively to how young children’s abilities manifest, they say relatively little about how or why they change. Modular and theory theory approaches, in contrast, regard these earlier manifestations of mental understanding as the product of qualitatively different cognitive mechanisms, with full representational theory of mind not being possible until more sophisticated mechanisms have come into operation. These accounts are much stronger in terms of explaining change, but do relatively poorly in explaining evidence on early abilities in non-test situations, or apparent inconsistencies in later performance.

In fact, the evidence appears on balance to be more consistent with the continuity of understanding proposed by functional accounts, with children exhibiting both more and less advanced reasoning contemporaneously. Evidence of continuity is not the same as explaining how progress is achieved, however, and some account is needed of how children acquire the insights necessary to pass false belief tasks, even if the application of these is initially piecemeal. It is this which is offered by social constructivist approaches to the acquisition of a theory of mind.

Such approaches argue that in general the developing child acquires the ‘tools’ of thought and action through interaction with members of his or her social group, who provide an induction into the accepted usages of these tools within that group’s culture. According to Vygotsky (1978), tools mediate the relationship between actor and object, allowing controlled action. While all tool use has this effect, words are the pre-eminent tool since they have a pervasive influence that no other tool possesses. “Thought development is determined by language i.e. by the linguistic tools of thought and by the social cultural experience of the child" (Vygotsky, 1962, p. 51). Vygotsky claimed that a culture's conventions of signs and its words in natural language drive mental activity, and provide the basis of psychological growth in the developing child. The beginnings of semiotic mediation emerge in early social speech, where parents give meaning to children crying, and start to use language in response to infant's early vocalisations, to direct the child's behaviour (Vygotsky, 1962; Lock, 1978). Later, the child uses overt language to him- or herself, as a means of directing their own behaviour. Finally, language becomes internalised as 'inner speech' or discursive thought, acquiring an intrapersonal function in addition to interpersonal use via the creation of higher psychological processes (Vygotsky, 1978). Thus what is initially intermental (shared between people) becomes intramental (within the individual).

According to Garfield *et al*. (2001, p. 531), "Vygotsky is tantalisingly close to developing an account of theory of mind per se". The basis of this claim is Vygotsky's insistence that the transition from the early speech phase to the child as a language user marks the transition from animal cognition to fully human thought: this transition allows the child to conceptualise, to posit and to use in explanations the concepts implicated in theory of mind. As Vygotsky (1978, p. 26) states, "Thus, with the help of speech children, unlike apes, acquire the capacity to be both the subjects and objects of their own behaviour". Although Vygotsky (1978) does not deny the contribution of ontogenetic processes in the development of thought and language, his contention is that the social environment provides both the initial reason to acquire the skill of learning to think and the necessary support to enable its acquisition: maturation is therefore viewed as a secondary factor. Central to Vygotsky's theorising is the concept of the zone of proximal development (ZPD) (Vygotsky, 1978). This presupposes two levels, a child's 'actual developmental level', marked by the ability of the child to work unaided on problems and tasks, and their 'potential development level', defined as the level of capability that a child can exhibit with the help and guidance of another person. The ZPD is the difference between these two levels, within which the more expert other scaffolds (Bruner, 1975) the child's performance by using language to control and direct certain elements of their activity; as the child appropriates and internalises this language to direct themselves, their competence grows.

In order to determine if it is appropriate to apply Vygostsky's social constructivism in the present context, it is necessary to address two questions: whether there is any evidence to support the argument that key aspects of theory of mind are acquired via guided activity within the ZPD; and, if so, which aspects of understanding are likely to be acquired by means of such guidance. As far as the first question is concerned, it is also important to note that guidance within the ZPD is by definition focused on the *performance* of an activity, not on abstract understanding, and it is therefore in such contexts that evidence must be sought.

There is in fact indirect evidence of the positive effects of adult support in the context of performance on theory of mind tasks themselves. Peterson & Siegal (1995) tested 26 deaf children, proficient in sign language and of normal intelligence, aged 8 to 13 years, and found that 65% failed a simple false belief task which normal pre-schoolers and children with learning difficulties pass at around 4 years of age. This delay in the development of theory of mind was attributed to deaf children having no access or exposure to the conversational support of caregivers e.g. early discussion of mental states.

There are also indications in other research that the way such guidance may operate in contexts of this kind is by helping the child to make explicit and keep track of the steps involved in false belief reasoning. For instance, Stevenson & Mitchell (1992), using the UCT (Wimmer, Perner & Leekam, 1987) asked children what they thought was in the box, and then suggested to children that when they first saw the box they thought there were Smarties inside. All 3 year olds agreed including those who had failed the false belief test. To control for the possibility that children were merely answering compliantly, in a follow-up study the suggestion was changed to the child thinking that the box contained 'jelly babies'. The vast majority of children rejected this false suggestion, again including children who failed the standard false belief test. Thus it would appear that with adult guidance and support young children are able to capture elements of correct judgements that otherwise elude them. However, as seen earlier, children can also be directed to agree with false realist suggestions (Stevenson & Mitchell, 1995b). This indicates that they are prepared to accept support in disembedding any task-relevant feature, and, within those bounds, do not discriminate further in terms of actual correspondence to events. In a sense, though, their susceptibility to such errors is further evidence of the scale of impact that such external direction of attention might have.

These studies not only provide evidence that external guidance can have a positive influence on children's capability to make mentalistic judgements, but also serve to yield a glimpse of the kind of arena in which support of this type might have particular importance. One implication is that there is some element of the demands of theory of mind tests that is within children's grasp, but which they cannot quite reach on their own. In line with this possibility, Siegal & Beattie (1991, p. 10) argue that since "experimenters' concerns are scientific ones while those of children can be more pragmatic and localised, differences may arise in the interpretation of conversation". It is these, they argue, that lead to difficulties with performance on standard tests of false belief.

This signals a potential reinterpretation of children's theory of mind performance – and wider mentalising abilities – in terms of their understanding of the pragmatic aspects of language i.e. the management of topics of conversation (Siegal, 1993; Siegal & Peterson, 1994). According to this view, insight into other people's mental states is acquired through and deployed within conversational practices, and it is inappropriate to consider theory of mind as a conceptual achievement in isolation from these. This idea has important implications for the nature of false belief tasks, with critics of them claiming that young children's problems with such tasks arise because they fail to share the experimenter's scientific purpose, and to grasp the context in which the latter's questions occur (Siegal & Beattie, 1991).

For example, in framing the question, 'Where will Maxi look for the chocolate?' in the original false belief task (Wimmer & Perner, 1983), the experimenter has made an assumption that the test aspect is relevant to children, and that they will be able to follow the conversational implication that this question means ‘Where will Maxi look first?’. However, when children do not share an understanding of the scientific purpose or context of the test they fail to make this inference and take the question to mean ‘Where will Maxi have to look to find the chocolate?’ Siegal & Beattie (1991) claim that, in effect, unsuccessful children’s responses correspond to the truth, as answers of this type are relevant to the external world as it is, and it is the experimenter who has violated this relationship. “Children’s early conversational habits are consistent with the speech input of caregivers who, for the most part, have not set aside conversational rules.” (Siegal & Peterson, 1994, p. 430). In support of this view, when children were explicitly asked where a character would ‘look first’ for an unexpectedly transferred object, in Siegal & Beattie’s (1991) ‘Look first’ false belief (FB) test, even 3 year olds were found to respond correctly.

Siegal is effectively making three claims in his research (Siegal & Beattie, 1991; Siegal, 1999) none of which have remained undisputed: (1) 3 year olds have the ability to pass standard false belief tasks, but are disadvantaged by the conversational characteristics of the test situation; (2) the reason they are disadvantaged is that they are unused to violations of conversational rules of the kind employed in false belief tests; and (3) when such violations are removed, their performance shows substantial improvement. In fact, Lillard (1999) notes that Siegal fails to investigate directly whether children expect experimenters to adhere to conversational rules, or indeed do so themselves. She also criticises the claim that children’s failure on standard versions of the false belief task is the result of the experimenter's ambiguous questions, citing one of many variant versions of the UCT employing a Smarties tube containing pencils (e.g. Gopnik & Astington, 1988). In this version, children are asked what they thought was in the container when they *first* saw it. Children of 3 years are apt to respond by stating 'pencils', indicating that use of the word *first* as a device to reduce the pragmatic demands of the question is not in fact consistent across false belief tasks, and does not necessarily serve to facilitate children's performance by clarifying supposed ambiguities.

Moreover, Clements & Perner (1994) failed to replicate the 'Look first' FB test findings of Siegal & Beattie (1991). This study employed a variation of Wimmer & Perner's (1983) classic false belief test, in which *Sam Mouse* placed some uneaten cheese in a blue box and left to go to bed. A second character then moved the cheese to a red box. When *Sam Mouse* returned for the cheese the experimenter said, 'I wonder where he is going to look?', followed by a 2 second pause during which a video camera recorded children's eye movements to either the red or blue box. The experimenter then reminded the child that *Sam Mouse* was looking for the cheese, and asked them either the standard false belief question or the 'look first' question. Clements & Perner found no significant differences between pass rates to the two questions (36% and 32% respectively). They did note, though, that the oldest group of children, aged over 4 years, both looked to the correct box and answered the false belief question correctly. The middle group of children, aged 2 years 11 months to 3 years 11 months also looked to the correct box but failed to answer the false belief question correctly. The youngest group failed on both counts.

Clements & Perner interpreted children's eye movements as revealing an implicit understanding of false belief that may be in place prior to the explicit understanding exhibited through verbal responses. This came into play, it was argued, because the task narrative presented children with sufficient information to determine how the mental state of *Sam Mouse* came about, unlike the explicit false belief test used by Siegal & Beattie (1991). According to Surian & Leslie (1999, p. 150), this led Clements & Perner to infer that "children can benefit from the 'look first' question only in the explicit false belief tasks because then the 'look first' question helps them to construct the experimenter's interpretation of the vignettes".

Clements & Perner's (1994) claims have themselves been subjected to criticism, however, one alternative view being that children look to the correct box because they are simply retracing the events of the story. Surian & Leslie (1999) also argue that Clements & Perner's experiment was confusing to children because it used repetitive questioning, which may have led children in the middle group to conclude that they had the wrong answer in mind. Surian & Leslie also used an inferred false belief task of the type employed by Clements & Perner, but found significant improvements in the performance of children aged 3 years 9 months when the 'look first' question was posed. They note that such patterns of performance have been attributed to *ad hoc* processing strategies which produce 'false positive responding' among children receiving the 'look first' question, reflecting a bias toward responding to the first location irrespective of the character's belief (Yadzi, German, Defeyter & Siegal, 2006). A further *ad hoc* processing strategy might be that children hear the experimenter's words, "look first" and presume there is going to be a *first* look for the protagonist followed by a *second* or further looks. This may lead them to think 'why should there be further looks unless the first look failed?', causing them to follow a strategy of answering where the 'first look' failed.

As a check on these possibilities, Surian & Leslie (1999) used a 'seen condition' and a 'not seen condition', the protagonist witnessing the swap to the different location in the former. They report that more children pointed to the empty box in the not seen condition when asked a 'look first' question, but in the seen condition significantly more children pointed to the box that contained the object. These findings indicate that 3 year olds' responses were not in fact merely the result of ad hoc processing strategies.

Even so, Wellman, Cross & Watson (2001) in their meta-analysis of 178 false belief studies, whilst recognising that a 'look first' question was one of several factors that influenced children's success, noted that this still left a proportion of variance unexplained. Thus any implication that mode of questioning was the sole factor determining children's responses – as Siegal & Beattie (1991) appeared to be claiming – was untenable. Moreover, 'look first' questions appeared only to improve performance among the oldest children in the 3 – 5 year old age range. However, Yazdi *et al*., (2006) found that the 'look first' effect was not limited to children approaching supposed conceptual change (at around 4 years of age), but was well established in younger 3 year old children, where there were appreciable advantages for 'look first' questions over the standard form. They conclude

(p. 362), "the 'look first' question acts as a conversational aid, providing a performance resource that helps 3 year olds to process the appropriate intent behind the experimenter's question and calculate the correct response. This idea has great intuitive appeal, because the manipulation can be captured as one that 'clarifies' things for the younger child, whose conversational resources are less sophisticated than those of the older child". The benefits are accounted for via a theoretical framework that integrates both conversational and executive-inhibitory performance factors. "The 'look first' question draws attention to the first location (i.e. the target of the false belief content) and renders it more salient than it otherwise would be with standard questioning. The question format thereby tends to reduce the salience differential between true belief (default) and false belief contents. The reduced differential in turn requires less inhibition to reverse its direction, allowing children with lower inhibitory resources a greater chance of success." (Yadzi *et al*., 2006, p. 363).

Support for the Yadzi *et al*., (2006) theoretical framework is provided by research which found that the 'look first' question may reduce executive inhibitory demands present in standard versions of the task (Carlson, Moses & Breton, 2002; Leslie, German & Pollizi, 2005). Further support is provided by indications that the working memory capacity of young children may limit their ability to comprehend standard versions of the false belief task (Mutter, Alcorn & Welsh, 2006).

The picture that emerges then is that if 'look first' questions have facilitatory effects in false belief tasks, these are perhaps only partially due to disambiguation of the meaning of i.e. to conversational support in this sense. At the very least, then, Siegal's original claims regarding the nature of these effects have had to be moderated. Siegal's (1999) contention that attentive caregivers adjust their speech to suit the characteristics of young children by making their utterances shorter, truthful, clearer and relevant to the joint focus of attention, and that young children assume researchers will act as parents do, is also not undisputed. A number of researchers have noted frequent violations of these basic conversational practices. One example is claimed to be adult jokes and teasing with young children (Astington, 1999). Another may be made by adults when referring to objects or people that are <u>outside</u> the joint focus of attention (Ninio & Snow, 1996). Research by Rundquist (1992) and Brumark (2006) suggests that violations of conversational practices between adults and young children are many and varied, and quite common within family discourse e.g. at dinner table conversations. Such violations are argued to be initiated by adults because they serve social functions with young children, such as addressing eating and table manners.

In a sense, though, even if Siegal's (1999) initial position on the impact of pragmatics within false belief and other theory of mind tasks has proved not to be completely tenable, this need not be considered a matter of great moment. The broader point remains, that standard theory of mind tasks are not normal conversational occurrences for children. Unless it is argued that the underlying competences targeted by such tasks are applied by children uniformly regardless of context – a point disputed in Section 1.5 above – then there is some reason to suppose that the unfamiliar character of the exchange may give rise to a variety of unpredictable consequences. Rather than attempting to track down the source of these, it may in fact be more profitable to consider the implications of the broader thrust of Siegal's argument: that mentalistic perspective-taking is intimately bound up with conversational pragmatics, and that adults have some control over the level of pragmatic demand placed upon children. If they choose to increase the level of demand in some aspects of everyday conversational practice, there is a very real possibility that this serves to stretch and augment children's understanding of the role of minds in determining those actions that make up conversation, via the social constructivist processes defined by Vygotsky (1962; 1978). If so, it is everyday conversations that are the key arena for the development of explicit, representational theory of mind capabilities.

1.7 The development of conversational pragmatics

As will be seen, work on the development of conversational pragmatics has produced a range of evidence in support of the view that key aspects of theory of mind are acquired through social constructivist processes and guided activity between caregivers and children. To some extent, this is consistent with Siegal's (1999) claim that children's understanding of false belief is related to their conversational awareness. However, Siegal has been criticised for his vagueness with regard to which aspects of conversational pragmatics are acquired by children through means of such external guidance: e.g. "Siegal's claim that what develops in children's cognition is an interplay of *both* conversational and cognitive is stated but not documented nor elaborated" (Lourenco & Machado, 1999, p. 21); "Siegal's theory of conversational awareness in never clearly spelt out" (Astington, 1999, p. 15). Siegal's lack of clarity on the specifics of children's development of conversational abilities and adult support for the acquisition of these leaves unanswered the key question: which conversational practices lead to conversational competence in children, and in turn promote their perspective-taking abilities? This question is addressed below, after more detailed consideration of the nature of conversational pragmatics.

1.7. 1 Historical background of conversational pragmatics

The study of conversational pragmatics can be traced back to the research of Austin (1962) who first identified the functional role that utterances have in the course of human communication. Austin argued that utterances perform actions e.g. 'I promise…' 'I find the defendant guilty… 'I name this ship…', Such performative utterances can only be analysed in terms of their usage in context. Therefore they correspond to an action, changing the state of the world i.e. the ship is only named after the act of naming has taken place (Crystal, 1987). To account for the many linguistic acts that conversationalists perform, speech act theory (Austin 1962; Searle 1969) was devised to examine the effect of utterances upon speaker and hearer behaviour by means of a threefold distinction. Locutionary acts constitute the fundamental act of saying something; illocutionary acts are acts performed as a consequence of the speaker saying something i.e. promising, adjudicating and naming etc.; and perlocutionary acts are the effect the speaker's utterance has upon the listener i.e. to instil pleasure, to be forewarned, to be informed etc. Many speech acts are also indirect requests for action i.e. an imperative request to ‘Shut the door’ may instead be supplanted with the utterance ‘Would you mind shutting the door?’ to avoid conveying abruptness or rudeness on the speaker’s behalf.

According to Grice, (1957; 1969; 1982; 1989a) conversational action is rule-governed, and an exercise in mind reading involving the expression and recognition of intentions. Speakers and listeners must co-operate to make conversation meaningful and purposeful. Grice (1975, p. 45) argues for the proper conduct of a conversation when he states “make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged”.

Grice (1975) maintains that to run a conversation successfully one must adhere to a co-operative principle, which involves the application of four conversational maxims. These are:

1) <u>Maxim of Quantity</u>: make your contributions as informative as possible, but no more.
2) <u>Maxim of Quality</u>: make your contribution true; do not say anything that you believe to be false or for which you lack sufficient evidence.
3) <u>Maxim of Relevance</u>: make your contribution relevant to the aims of the conversation.
4) <u>Maxim of Manner</u>: be clear: avoid obscurity, ambiguity, wordiness and disorder in your language.

However, in the adult world this co-operative principle is often violated (Grice, 1975), requiring us to register a 'conversational implicature'; that is, to recognise that something is implied or meant beyond the utterance itself. For example, the question 'Can you pass the salt?' does not mean literally 'Do you have the ability to pass the salt?' The inference must be drawn that it is a polite request to pass the salt. Understanding of such conversational implicatures is known to be problematic for young children, and yet to become competent communicators they must learn how to decode them appropriately (Siegal & Peterson, 1994).

In an extension of the work of Grice, Sperber & Wilson (1986; 2002) argue that the goal for Pragmatics is to explain how a hearer constructs the actual meaning that the speaker intended to convey given the latent and risky nature of conversational implicatures and indirect speech acts i.e. there is no guarantee that speaker meaning will be correctly coded or inferred by the hearer simply by the latter following the best possible linguistic processing. The recovery of meaning by best possible linguistic processing involves the application of pragmatic language understanding, which is ultimately governed by the presumption of relevance (cf. Sperber & Wilson, 1986; 2002). "The utterance is presumed to be the most relevant one compatible with the speaker's abilities and preferences, and at least relevant enough to be worth the hearer's attention" (2002, p. 13). Thus the search for and recognition of relevance in communication is argued to form the crucial basis for interpretation and inference of speaker meaning.

Sperber & Wilson (1986; 2002) also argue that context and cognitive efficiency help serve to guide determination of relevance: "Relevance is a feature of all stimuli, linguistic or otherwise, with various stimuli being more relevant than others in a given context, depending on the effects they generate in that context" (2002, p 10). The effects generated by context are influenced in turn by the allocation of attentional resources to maximise the selection of relevant stimuli. This notion of a cost-benefit analysis applied to relevance is also apparent in Bezuidenhout & Sroda's statement (1998, p. 270), "In interpreting the speaker's utterance, the listener will select the first interpretation which has adequate contextual effects for no gratuitous effort".

According to Sperber & Wilson (2002), the predisposition to maximise relevance may allow individuals, at least to some extent, to anticipate the mental states of others. For example, they argue that an individual 'A' can often predict:

"(a) Which stimulus in an individual B's environment is likely to attract B's attention (i.e. the most relevant stimulus in that environment);
(b) Which background information from B's memory is likely to be retrieved and used in processing this stimulus (i.e. the background information most relevant to processing it);
(c) Which inferences B is likely to draw (i.e. those inferences which yield enough cognitive benefits for B's attentional resources to remain on the stimulus rather than being diverted to alternative potential inputs competing for those resources)".
(Sperber & Wilson, 2002, pp. 10-11).

Sperber & Wilson (2002) argue that inferential comprehension involves use of such relevance-based procedures to decode ostensive stimuli and in particular, linguistic utterances, but also provision of support by speakers for such decoding: "Communication is achieved by ostensively providing an addressee with evidence that enables him or her to infer communicator's meaning" (p. 12). Such evidence would include:

(1) 'A' attracting the attention of 'B'
(2) 'A' prompting the retrieval of certain background information from the memory of 'B'
(3) the joint processing of background information that leads 'B' to draw certain inferences intended by 'A'.

Sperber & Wilson (2002) suggest that use of relevance-guided pragmatic interpretation to identify speaker meaning may be a human adaptation, in the form of an evolved sub-module relating to a domain-specific mind reading ability. In extending their theory of pragmatic language and mind reading abilities to that of Grice and the neo-Griceans, they go on to argue (2002, p. 6) that "Grice substantially underestimated the amount of metapsychological inference involved in comprehension". Indeed, there is reason to suppose that Grice viewed the decoding of conversational implicature as an intuitive process rather than a reflective one that might require theory of mind capabilities.

Sperber & Wilson's (1986; 2002) exposition of relevance, as the prime maxim governing inferential comprehension, raises issues regarding the role that understanding of relevance might play in children's ability to infer speaker meaning, and how this may connect with theory of mind. They rightly state that most studies of children's mind-reading focus upon the ability to attribute beliefs and desires to others. However, relevance theory would take as its starting point aspects of mind-reading that involve attracting the attention of others and the more reflective evidence (outlined above) that enables a person to infer a communicator's meaning.

1.7. 2 Conversational practices: violations of maxims by caregivers in interaction with children

Viewed from this perspective, the origin of inference-based decoding of relevance may lie in child-caregiver interaction, and the naturalistic, observational research of Ninio & Snow (1996) into the early pragmatic language skills of infants supports this premise. According to these authors, the gradual ability of the child to build upon its social and linguistic experience is supported by attentive and sensitive caregivers, who guide the child toward effective perspective-taking, including shared understanding of the relevance of references. A range of adult pragmatic skills feed into this process, including turn-taking, maintenance of topic-relevance, topic initiation and topic transition, eliciting of participation, and use of repair to bring otherwise dislocated child utterances into the body of the conversation. At a broader level, attentive caregivers recognise the need to adjust their speech to suit the characteristics of young children, by complying with the social rules of conversation (Grice, 1975) and with the principle of complementarity (Goffman, 1983); i.e. they make their utterances shorter, truthful, clearer and relevant to events in the external world, and consider the assumptions, knowledge and viewpoint of the infant listener (Ninio & Snow, 1996).

However, counter to Siegal's (1999) view that caregivers do not violate conversational maxims, a number of researchers report frequent adult violations of the basic conversational maxims (e.g. Rudquist, 1992; Ninio & Snow, 1996; Astington, 1999; Lillard, 1999; Brumark, 2006). These researchers regard such violations as both a natural and necessary means of communicating with young children. Ninio & Snow (1996) argue more particularly that they serve to stretch the child's abilities, by requiring them to keep up with more opaque references. In this account, adults work to establish an expectation of relevance on the part of infants, and then gradually introduce increasingly complex violations of these kinds in turn, initially, for example, by asking the child to retrieve a toy that is outside of the current joint activity and beyond the child's immediate focus of attention. Later instances may take the form of references to a shared past event in memory, especially where this is relevant to a present activity e.g. "Do you remember where we saw a Lion like the one in this story?" The types of reference made by parents appear to shift as children get older, effectively introducing them to more demanding inferences. In this way, parents exemplify conversational practices, and help children to interpret what others may have in mind as to objects, people and events which are not immediately apparent.

At a more technical level, by doing this, they also illustrate the ways in which non-immediate relevance can be established, and how this depends on shared understanding of possible mental connections. The corollary implication is that children exhibit a growing ability to decode adult conversational implicatures, as these arise in everyday conversations and during joint activities. These may include variation in the form of requests (Babelot & Marcos, 1999), such as imperative requests i.e. direct requests (e.g. “Give me the doll”), or directive questions i.e. indirect requests (e.g. “Have you got the ball?”). The latter may be particularly problematic for young children when the toy items are outside the joint focus of attention. Here, the child must learn that something is implied or meant beyond the utterance itself i.e. it does not mean that he or she literally has the toy. They must therefore draw the inference that it is a polite request to locate and retrieve the object referred to.

Other work also points to the particular importance of early parental support for the decoding of relevance. According to Grice, (1975); Levinson, (2000); Carston, (2004), Particularised Conversational Implicatures (PCIs) relate to violations of the maxim of relevance, where making a correct inference is dependent upon particular contextual grounds provided by the speaker. In contrast, Generalised Conversational Implicatures (GCIs) are based on violations of the maxims of quantity and manner (and to a lesser extent quality), where inferences are generally derivable without the help of context. According to Borg (2005, p. 26), children have greater difficulty in the decoding of GCIs, as these “are accessible without any knowledge of the aims and intentions of the speaker and they are generated by computational, deductive processes”. Research by Noveck & Posada (2003) confirms longer time–delay variations in the recovery of GCIs for children as opposed to adults. The central point is that PCIs, which derive from violations of relevance, appear to be easier to resolve for young children, and thus are more likely to form the focus of parental support. Thus Borg (2005) argues that children’s understanding of PCIs requires early perspective-taking abilities, with gradual progression towards the recovery of GCIs demanding further language competencies: “As exposure to normal language use increases, so the drawing of implicatures based not directly on particular speaker intentions but on the characteristic use of words increases” (Borg, 2005, p. 26).

1.7. 3 Conversational practices and developmental progression in conversation

There is reason to hold therefore that the maxim of relevance (and violations of it) should be the central focus of enquiry with respect to conversational practices involving young children and the relationship of these to their perspective-taking abilities. This view is strongly supported by the literature in the area. Haslett & Samter (1997, p. 99) argue that the most fundamental skills for children in acquiring conversational competence are the understanding of intentionality and the maxim of relevance: “Intentionality presupposes the meaningful and purposeful nature of messages; while relevance mandates a search for connectedness within the ongoing behaviour and/or conversation”. This ‘search for connectedness’ may be one that is initiated by adults through violations of immediate relevance (cf. Ninio & Snow, 1996).

The observational research of Ninio & Snow (1996) goes on to draw particular attention to the crucial role that the conversational violations may play in children's developmental progression in conversation, and the move towards a degree of conversational competence. Ninio and Snow make the point that developmental progression in conversation reflects the main organising principle of the speech use system i.e. the more a conversation is built on a joint perceptual focus, the earlier it is participated in verbally by young children. Thus, children progress from discussions of topics in their external world, with parental references to immediate objects being talked about earlier than references to non-immediate objects, and parental references to shared past events being introduced when children become familiarised with talking about joint present activities. Progression then moves towards the discussion of topics concerning children's internal world, with adult references to evident and non-evident emotions and evident and non-evident beliefs i.e. young children's external world having a greater tangibility for them than their internal world.

To participate effectively in these more decontextualised conversations (i.e. those where background knowledge is not explicit, as in discussion of non-evident emotion and non-evident belief states), young children require the conversational, historical and psychological support of adults. Conversational support is provided by adults building simple child utterances into larger and more complex informational and interactive structures. Historical support derives from use of adults' and children's shared history of experiences e.g. likes, dislikes and shared topics. Psychological support comes into play when adults take children's perspective on events e.g. an adult's readiness to find monsters scary (Ninio & Snow, 1996).

Developmental progression in conversation is gradual, however. As the work of Ninio & Snow (1996) suggests, and as examination of the CHILDES Language Database (MacWhinney, 2001) will shortly show, young children do produce effective and functional utterances and are assisted in doing so by adults. Adults appear to maintain topic relevance in discussion through questions designed to elicit relevant background information, by linguistically prompting and modelling the child's knowledge of events when insufficient information has been given, and by allowing the child conversational turns. Having worked to establish the expectation of relevance on the part of infants, adults then gradually introduce violations of relevance in the following conversational domains: through their conversational support (adult references to non-immediate objects); by use of historical support (adult references to shared past events during a present activity) and via psychological support (adult references to non-evident emotion and belief states).

As we will return to later, Ninio & Snow's (1996) mapping of developmental progression in conversation leads to clear and testable predictions about the sequence in which children's understanding of these different types of reference would be expected to emerge. However, ultimately, this is not just a matter of absence or presence of an ability but of relative ease of application of such understanding in different conversational contexts, with varying degrees of other types of demand. As Ninio & Snow point out that a key aspect of acquiring conversational competence is learning to operate without these three supports, i.e. under the withdrawal of the presumption of aligned perspectives with adult partners. “Young children’s stories often align perspectives successfully only because they can rely on empathetic listeners to anticipate the desired reaction. When they can no longer assume anticipatory alignment, they have to start working on linguistic strategies for conveying the appropriate reaction” (Ninio & Snow, 1996, p. 184).

One observable consequence of the shifts that occur once support has been withdrawn and children have begun to start operating as independent conversational partners is the exercise of some degree of ‘pragmatic sensitivity’ (cf. the notion of infant rich interpretation) to the conversational demands imposed by speaker on listener, as evidenced by the modification of a response, for example according to referential context and grammatical form. At a more rudimentary level, it may be signalled by recognition that a response is required even when the answer is uncertain, and an attempt to explore what form of response might suffice. Whilst not a strategy in itself, pragmatic sensitivity of this kind – i.e. an awareness that there are conversational demands which partners expect to be met, occasioned by the reduction or withdrawal of support – may plausibly be the driver of efforts to gradually map referential forms onto intended functions, as well as being a marker that such efforts are being made. Framed in this way, the type of adjustment made (recognition that a response is required, or modification of the character of that response according to context or grammatical form) would provide an index of how far these efforts had progressed, and in this sense, the conversational maturity of the child.

Babelot & Marcos (1999) report data that indicate forms of pragmatic sensitivity are in place as young as 1:7 years, utilising a task in which requests for children to supply objects that were either part of current activity or outside of it were made using various direct or indirect grammatical forms. In this task, children typically used action responses (touching or retrieving an object) for requests relating to the immediate context (i.e. objects in view that were relevant to a current activity). In contrast, requests that did not relate to the immediate context (i.e. for objects still in view but further away that were irrelevant to a current activity) tended to result in information responses (i.e. pointing or verbal indications to the correct object) and incomprehension responses (verbal or gestured puzzlement). This pattern obtained regardless of the grammatical form of the request and the child's developmental level. As Babelot & Marcos (1999, p. 184) argue, the results indicate "the use of context as the first criterion for interpreting the utterance and then, whenever that context is unclear or not very informative, the interpretation is based on the linguistic characteristics of the utterance".

Babelot & Marcos (1999) did find effects associated with the referential form of requests, however. They used four grammatical forms: imperative requests, embedded imperatives, expression of desire, and directive questions. Imperative requests (e.g. 'Give me the doll') elicited the most correct responses, whilst directive questions (e.g. 'Have you got the doll?') appeared to be the most demanding form, containing a conversational implicature that required children to draw the inference that it is a polite request to pass the object. In particular, when a request of this form was made for a non-immediate object, they often answered literally with an indicative information response of “Yes” or “No” (that is, X is available to them but they have not physically got X). Children gave the fewest correct responses to this question. Crucially, however, Babelot & Marcos report that children were sensitive to the requirement to respond to a request even when the answer was uncertain. They found that children would more often give the wrong object or give incomprehension responses than not respond at all. The change in the form of their response to such indirect requests might also be taken to suggest at least the beginnings of an awareness of the need to tailor reactions to the character of the demands made of them – though this would be more clearly signalled in this context by responses that provided more precise verbal information about the location of the requested object.

As indicated above, putting this bias towards active responding from an early age together with Ninio & Snow’s (1996) claims about children’s developmental progression in conversation, the implication is that these types of pragmatic sensitivity to different forms of request might serve as a sophisticated index of the extent of children’s understanding of references to immediate objects, non-immediate objects, past events, non-evident emotions, and non-evident beliefs. In the context of tasks of the form used by Babelot & Marcos (1999), for instance, the greater the degree of pragmatic sensitivity shown by children, the more secure their understanding of a type of reference might be concluded to be. This is because the processing demand deriving from decoding the reference itself (i.e. the primary element of the conversational turn) would apparently be sufficiently small for attention to be paid to addressing the additional layer of implicature. Support for this kind of processing load approach is provided by Bezuidenhout & Shroda (1998), who found that 3 year olds could make use of an experimenter's restricted viewpoint to decode otherwise referentially opaque requests. They also suggest, however, that the immature working memory of younger children may be involved in them making inappropriate responses to ambiguous messages, these imposing demands that are beyond their capacity to work through. These methodological points will be returned to later.

1.8 CHILDES Language Database: The research background

In order to explore further how far young children's ability to decode relevance is both taxed and supported in everyday conversational exchanges with caregivers, extracts from the CHILDES Language Database (MacWhinney, 2001) were examined. Only conversations involving children under the age of 5 years old were considered, and the Child Language Analysis (CLAN) component of CHILDES was utilised to identify instances of 'lexical search' and 'discourse analysis/interaction' within these. These were deemed most likely to identify examples of dialogue where issues of relevance were focal i.e. where references to things, topics and objects were the subject of joint focus of attention within caregiver-child conversational exchanges.

Programmes implemented to conduct the search included 'CHAINS' (which displays 'runs' or 'chains' of speech acts); 'CHIP' (tracks imitations, repetitions, lexical overlap); 'FREQ' (tracks the frequency of each word used); 'KEYMAP' (looks at the variety of speech acts following a given act) and 'KWAL' (search for a specific word or group of words, including, in this case, words associated with references to internal states, such as 'look', 'think', 'remember', 'funny', 'sad' and 'cross'). 'KEYMAP' was used to trace the extent to which a mother's question was followed by an answer from the child. 'CHAINS' allowed tracking of between and within speaker sequences of speech acts, reference types, or topics.

The following exchange is typical of many examples in CHILDES, where parents seek to establish and maintain clear dimensions of conversational relevance during joint play by referring to objects that are in the immediate focus of attention. Here, Anne, aged 2:0 is building 'Duplo' with her mother.

<u>EXAMPLE 1</u>

MOTHER:	Have all the bits gone down the slide?
	Referring to bits of Duplo.
MOTHER:	What about this one, Anne?
ANNE:	That one go there.
MOTHER:	Chinese bridge.
MOTHER:	What do you think of that?
ANNE:	Oh.
ANNE:	And get it off.
ANNE:	Get those off.
MOTHER:	Those off?
ANNE:	Get those off.
ANNE:	Look.
MOTHER:	What?
MOTHER:	It is a bit of a crash of bricks at the bottom of that slide,
	Isn't there?
ANNE:	Oh.
MOTHER:	What has happened?
ANNE:	Oh.
MOTHER:	What has happened, Anne?

ANNE: What?
MOTHER: Tell me about it.
MOTHER: What has happened?

Manchester Language Study, CHILDES Language Database (2001)

In this example, the mother establishes what is legitimately relevant by repeatedly referring to objects that are the focus of immediate attention, that is, the Duplo bricks, a toy slide and a collapsing toy bridge. As the extract draws to a close, climaxing with the collapse of the toy bridge, the mother also repeatedly prompts Anne to speak about ‘What has happened’ and to ‘Tell me about it’. By trying to persuade Anne to reveal her perspective on events, the mother may be attempting to encourage the child’s conversational competence, by giving her conversational support.

Ninio & Snow (1996) suggest that caregivers, having worked to establish an expectation of what is relevant on the part of infants by referring to objects in the immediate focus of attention gradually introduce violations which effectively illustrate the ways in which reference to non-immediate objects can be made relevant, and how this depends on shared understanding of possible mental connections. Further examination of the CHILDES database provides examples of this.

For instance, in Example 2, John aged 2.5 is playing with toys in the presence of his mother and an investigator, Anna from the Manchester Language Study. John’s mother raises the subject of lost toys.

EXAMPLE 2.

MOTHER:	Haven't seen your tractor for a few days.
JOHN:	Where is the tractor?
ANNA:	I don't know.
ANNA:	Have you lost it?
JOHN:	Yeah.
JOHN:	Lost this man.
JOHN:	Lost this tractor.
MOTHER:	Lost.
JOHN:	This is lost this man.
JOHN:	It is lost this man.
MOTHER:	John.
MOTHER:	You could look for it.
MOTHER:	You could look for it in your red box.
JOHN:	Oh.
JOHN:	No
MOTHER:	Well have a good hunt for it.
MOTHER:	Have a root around.
JOHN:	There is a man.
ANNA:	Oh yeah.
ANNA:	He doesn't look like a farmer to me though.
JOHN:	Look
JOHN:	Man.
ANNA:	Okay.
ANNA:	I'll look after the man then.

Manchester Language Study, CHILDES Language Database (2001)

This further example is typical of a joint play situation where the mother introduces a violation of relevance to the immediate context by referring to an object (toy tractor) that is not in the immediate focus of joint attention. John indicates effective understanding of the reference, by referring in turn to the toy man that is associated with the tractor. The mother later provides a clue for the child as to where the objects might be located, and this may be an instance of historical support as the 'lost' toy may have been previously kept in the red box. John in turn connects the exchange back to the present context, confirming his perception of its relevance, by pointing out the toy man that is present. This conversation illustrates the ways in which the relevance of non-immediate references can be established by caregivers, provided there is shared understanding of the possible mental connections. Similar violations of reference only to those things that are immediately relevant can also occur on the part of young children who do not always possess the necessary linguistic skills to formally request an object of their desire that is not physically present. In these situations, parents and children must work hard to establish the identity of the object that children have in mind. Example 3, where Gary aged 1:6 is crying and pointing to a kitchen cupboard with his father and mother in close attention, illustrates this.

<u>EXAMPLE 3</u>

FATHER:	What do you want?
GARY:	Look? [Gary is crying. Trying to say something]
FATHER:	What do you want?
MOTHER:	Come here.
MOTHER:	What?
GARY:	Look.
MOTHER:	Come here.
	[Father lifts him up to cupboard. Gary takes biscuit tin out]
FATHER:	Hey Janet [Mother]
MOTHER:	Yes.
FATHER:	Look.
FATHER:	That what you wants'?
GARY:	yyy [Yes]
FATHER:	What do you want?
GARY:	That?
FATHER:	He don't take one.
MOTHER:	He always takes at least two biscuits. He took one for everyone today
MOTHER:	He got to take two!
FATHER:	Yes.
FATHER:	Alright?

Wells Language Study, CHILDES Language Database (2001)

In this example, Gary does not possess the appropriate linguistic skills to name the object of his desire (biscuits), which is also outside immediate attention. Gary attempts to direct parental attention to the biscuit tin by pointing and using the single-word directive utterance i.e. 'look'. It is apparent that the parents, mindful of the child's attempts to direct their attention towards an unknown, non-immediate object, prompt, question and physically assist Gary until the relevance of his attempted communication is established by identification of the object. They then underpin the shift in focus of attention by conducting a further exchange about biscuits – including a reference to a past event that Gary was involved in. Whilst the cognitive consequences are unclear, the exchange would seem to be a clear demonstration of the use of scaffolding (Bruner, 1975) of the child's grasp of referential possibilities.

Similar scaffolding of children's ability to understand the different ways in which an utterance can be meant to be relevant is evident in other parental references to past shared events during a present activity with their children. In Example 4, Anne aged 2:9 is being read a book by her mother.

<u>EXAMPLE 4</u>

ANNE: We read books.
MOTHER: Which one would you like to read?
MOTHER: Let us put that out of the way then. [Another book]
ANNE: Read.
MOTHER: Where did we buy this?
MOTHER: Can you remember where we bought… where David bought you that?
ANNE: In the shop.
MOTHER: In the shop.

MOTHER:	But where were we?
MOTHER:	Can you remember?
ANNE:	At the…
ANNE:	What is it?
MOTHER:	Well.
MOTHER:	I think you called it a church… didn't it?
MOTHER:	It was a big house.
ANNE:	Church.
ANNE:	Church.
MOTHER:	It was at that big house, wasn't it?

Manchester Language Study, CHILDES Language Database (2001)

The mother's reference to a past shared event (the purchase of the book) may serve to help establish an understanding on the part of the child of how such references can be relevant even if they violate the joint focus of attention on immediate circumstances, by illustrating how they can grow out of a connection to something that *is* immediately present (the book). Indeed, the mother's somewhat gratuitous introduction of this reference seems to be a clear instance of the kind of stretching of children's understanding of referential possibilities described by Ninio & Snow (1996). Again, there is effective use of historical support by the mother in providing the child with a succession of referential contexts to help the child recall the location of where the item was bought.

Parents may also introduce references to present and immediate, but non-evident emotions. In Example 5, Emma aged 1:9 is being read a story by her mother.

EXAMPLE 5

EMMA:	Lion.
MOTHER:	'Raar', yes.

MOTHER:	Ok, there's the lion again.
MOTHER:	What's a lion say?
EMMA:	‘Raar’
MOTHER:	Oh, you hear that construction noise?
MOTHER:	Is that what you hear?
MOTHER:	That big machine that we saw out on the street.
MOTHER:	‘Huh’, ‘boomamamama’.
MOTHER:	Is (laughing) that funny?

Tardiff Language Study, CHILDES Language Database (2001)

Although the mother originally supposes that Emma’s roaring noise refers to a lion in the story she soon realises that the noise is coincidental and is the child’s attempted reference to the noise of a construction machine operating on the street. As the notion of coincidences, are complex and multifaceted, the mother makes reference to her non-evident emotion of amusement that stems from her confusion of the referent of Emma’s noise.

As children grow older, caregivers’ references to non-evident emotions may also allow a shared understanding of mental connections to emotional states to be constructed. In Example 6, Jonathon who is aged 2.9 is hiding the keys to some of his toys. His Mother, who is in a different room, holds a conversation with him.

<u>EXAMPLE 6</u>

	[Jonathon is hiding the keys to some of his toys]
MOTHER:	Then I will be very, very cross.
JONATHON:	And what will you say?
JONATHON:	Then what will you say?
MOTHER:	I will say that you can't play with it any more.
MOTHER:	If you can't look after your toys properly you won't play with them.
JONATHON:	Very, very cross with you.

	Repeating what Mother said over
JONATHON:	Very cross with you.
	Repeats 15 times

Wells Language Study, CHILDES Language Database, (2001)

It seems fair to presume that this is an action that Jonathon has carried out before and one he realises will result in his mother's disapproval. The shared understanding of the relevance of this reference to a past emotional state (i.e. his mother's past annoyance at his hiding of the keys) allows Jonathon to form the mental connection that events in the world may cause emotional states in caregivers even when those caregivers are not immediately present. As he cannot rely on the visual expression of emotion from his mother, Jonathon's improving language skills and shared understanding of the relevance of the reference enable him to question and confirm his mother's emotional response to his action, and he may now comprehend that his mother's utterances are a reliable predictor of the emotion she feels.

Parents also introduce relevant references to present and immediate but non-evident beliefs. In Example 7, Aran aged 2:3 and his mother are playing with some toy vehicles and a toy town.

<u>EXAMPLE 7</u>

ARAN:	Is there enough room in that garage?
MOTHER:	No.
MOTHER:	There isn't enough room in there.
MOTHER:	You see you've got that one in… haven't you?
MOTHER:	And oh.

MOTHER: Shall we put this horse and carriage near the museum?
MOTHER: I think that is where it should go… don't you?
MOTHER: Because it is rather old fashioned.

Manchester Language Study, CHILDES Language Database (2001)

The mother establishes shared understanding of the relevance of a reference to a non-evident belief state by initially making Aran aware that she thinks that the toy horse and carriage should be placed at the museum. Subsequently, she seeks confirmation that Aran shares this thought and is thinking likewise with her question 'don't you' before finally explaining the reason in her mind as to why the toys should be placed there i.e. because of their old fashioned nature. Parents seeking to establish that young children are sharing thoughts before giving them the reason for their belief state is a common theme found throughout the database, and one that highlights the scaffolding process that parents employ to support understanding among their children of the relevance of references to non-evident belief states.

In Example 8, Evan is aged 5 years old.

EXAMPLE 8

EVAN: yeah, Mom, and you know what? on the tv you know like I was so scared when the tv was turned off, I was like, "uhoh, there he is"
MOTHER: oh, you felt you were seeing him places?
EVAN: like um, like um, like um, like um I was dreaming about a Tyrannosaurus rex, and he saw me and he came after me.
MOTHER: yeah?
EVAN: and went and didn't pick me up.
MOTHER: oh.
EVAN that like [screams]. I said, wh- wh- wh- wh- wh-where were we?

[pretending to stutter from fright]

Ninio & Snow (1996)

In this example, Evan lucidly articulates a frightened but imaginative perspective on an event. However, his mother still provides psychological support for his belief state by actively aligning perspective with him by her utterance, 'oh, you felt you were seeing him places?'. It is apparent that psychological support (cf. Ninio & Snow, 1996) within conversational exchanges continues by parents with older children i.e. those deemed to have passed the age-threshold for representational theory of mind understanding.

Study of the examples given above also reveals constraints on the subject matter that adults and young children can discuss. Childhood conversations develop around topics founded in joint activities, e.g. joint play with toys, games, book reading and fantasy themes, expanded by adult references to connected objects that may or may not represent a shared focus of attention, to connected shared past events and later to connected non evident emotion and belief states. According to Ninio & Snow (1996, p. 153), 'Young children are disadvantaged in the making of small talk by their ignorance of the standard, culturally determined list of topics i.e. the subjects that organise casual conversation for adults as well as their lack of knowledge about those topics' (cf. Kellerman, Broetzmann, Lim & Kitao, 1989). However, the ability of adult partners who have a knowledge of social rules for the conduct of conversations, coupled with a knowledge of the ways in which particular types of reference can arise in connected fashion out of an activity – in both adult and child contributions – serves as a resource to help structure interactions and foreground these referential possibilities. In this way, young children can be inducted into appropriate conversational practices, and grasp of the cognitive events which underpin these.

Thus, for instance, there are social conventions that organise casual conversation for adults, and one widely recognised custom between people involves turn taking in conversational exchanges. Levinson (1983) proposes that much conversation in the adult world is organised around the social rule of two part exchanges called adjacency pairs. These are sequences of two utterances that are: (i) adjacent; (ii) produced by different speakers; (iii) ordered as a first part and a second part; and (iv) that consist of types so that a particular first part requires a particular second (or range of second parts) e.g. greetings require greetings. Adjacency pairs are commonplace during adult conversation in such contexts as questioning/answering, informing/acknowledging and complaining/excusing.

Schegloff & Sachs (1973) also propose that adjacency pairs are governed by conditional relevance, 'Given that a speaker produced a first part, the second pair part is relevant and expectable as the next utterance'. Such conditional relevance may be a difficult concept for children to grasp, and young infants in particular will frequently not respond to adult utterances with an appropriate adjacency pair. However, young children's minimal contribution to conditional relevance seems to be compensated for by parent's extensive and enriching input, and their adherence to such conventions themselves, drawing the child into this structure of exchange, and the relevance of elements within it.

So, for instance, examination of the examples above, and the CHILDES Language Database (MacWhinney, 2001) more widely, reveals that much adult-child conversation is organised in three or more part exchanges formed around joint activities and adult references. In three or more part exchanges, adults maintain a topic connected to current activity until children respond, and generally their responses are then followed by an element of adult feedback, providing closure that signals acceptance of the child's decoding of the relevance of the adult initiation, and/or the relevance of the child's own contribution. Such initiation-response-feedback (IRF) patterns of exchange are of course a common feature of teaching situations (Sinclair & Coulthard, 1975).

According to Ninio & Snow (1996), one of the measures of mastery of the pragmatic system is the contribution of individual utterances in verbal interactions. Mothers produce a larger number of speech acts than young children do, with mothers' utterances being most prevalent when children are aged 12 months; the number of speech acts made by children increases as they grow older, and converges with mother's utterances by 32 months of age.

1.9 Outline of research and broad hypotheses

1.9.1 Outline of the present research

The application of rich interpretation accounts, social constructivist thought and developmental pragmatics suggests that children's performance on standard theory of mind tasks, and their mentalistic understanding more generally, may be bound up with their grasp of conversational practices, particularly the decoding of references that apparently violate the principle of relevance to immediate context. The research of Ninio & Snow (1996) and examination of the CHILDES Language Database (MacWhinney, 2001) suggests that caregivers, during conversation with young children, establish an expectation of relevance by converging with children on meaningful topics concerning immediate objects; present joint activities; evident emotions and evident beliefs. Caregivers then gradually introduce violations of relevance in respect of each of these topics in turn i.e. via references to (1) objects that are present but non-immediate, (2) past shared events during a present activity (3) non-evident emotions, and (4) non-evident beliefs. In order to converge on understanding of the object of these references, young children rely upon the conversational, historical and psychological support of caregivers (Ninio & Snow, 1996).

These instances serve to illustrate the ways in which the relevance of non-immediate references can be established, and how decoding of relevance depends on shared understanding of possible mental connections. This in turn locates the development of a facility with understanding the perspective of others more generally within the communicative practices of caregivers, and the way in which these stretch children's interpretative abilities to meet established communicative conventions.

Previous work has examined the relationship between language and mentalistic understanding (e.g. Dunn, Brown, Slomkowski, Tesla, & Youngblade, 1991; Meins & Fernyhough, 1999) but the focus of this has been on global associations, i.e. how the reference by parents at one point to mental and emotional states predicts the usage of such constructs by children at a later date, or performance on false belief tasks. Investigation of children's understanding of non-immediate referents including non-evident mental and emotional states, and the informal rules governing how these may be brought into conversations, would put the focus more squarely on the role of actual circumstances under which utterances have to be interpreted via a grasp of what others might plausibly have in mind. It therefore offers the possibility of greater insight into the language/mentalistic understanding relationship.

This said, the research to date on pragmatic development is largely observational and focused on performance, and the connections to performance on tasks such as those involving inference of false belief are essentially speculative. What is needed to move things forward is systematic and controlled investigations of a) children's competencies with respect to the interpretation of statements utilising each of the four types of non-immediate reference identified above, and b) the relationship between competence with respect to decoding internal references (i.e. to emotions and beliefs) and performance on standard false belief tasks. Such studies would pave the way for investigation of the relationship between the then determined competencies of children and everyday caregiver-child interaction. This would in turn help establish the extent to which pre-school children's understanding of theory of mind emerges in pragmatic contexts and situational exigencies rather than simply as an age-related conceptual shift (cf. Newton *et al*., 2000).

The present thesis reports on four such studies, each using similar materials within a request task of the form used by Babelot & Marcos (1999), and each focused on the prototypical situation of an apparent violation of relevance within a verbal request made to a child. This violation was potentially and plausibly resolvable by the child using awareness of a specific referential possibility or relevance sub-maxim (cf. Ninio & Snow, 1996) to direct their attention to available but non-obvious information. In each instance, the issue under investigation was whether children in the pre-school age range were able to make the appropriate resolution, as indicated either by their actions or a verbal response. Study 1 was concerned with understanding of references to objects that were physically present but outside of the immediate context. Study 2 addressed understanding of references to a shared past event. Studies 3 and 4 concerned understanding of references to non-evident emotions and non-evident belief states respectively. Taken together, then, the four studies addressed the apparent developmental sequence of pragmatic language competence outlined by Ninio & Snow (1996). In addition, Studies 3 and 4 included measures of standard false belief performance, in order to assess the relationship between this and understanding of references to both evident and non-evident internal states. Similarly, standardised language tests and memory tests were included in the four studies to examine whether children's competence was related to their expressive language or working memory capacities.

It was anticipated that these studies would reveal a sequence of emergence of understanding of each of the four types of reference which was consistent with that described by Ninio & Snow (1996) on the basis of naturalistic and observational research and the influence of caregiver input that they propose. However, since the kinds of support provided by caregivers were absent in the experimental setting, each study effectively constituted a test of independent competence rather than embedded performance, and it was therefore possible that apparent age of emergence might vary upwards from that reported by Ninio & Snow. The early pragmatic language skills found in children aged 1:7 years by Babelot & Marcos (1999), and the evidence of false belief understanding in 3 to 4 year old children reported by Siegal & Beattie (1991) using variants of the false belief task, set lower and upper boundaries on what might be expected. These age-groups correspond broadly in fact with the observational research of Ninio & Snow (1996): the acquisition of the four sub-maxims appears to be protracted, with discussions of non-immediate objects (cf. Study 1 in the present research) evident from around 18 months, but conversation involving beliefs (cf. Study 4 here) not appearing until around 32 months.

Children from 1.5 to 4.0 years were therefore recruited for the four studies from two university ‘on campus’ nurseries in Glasgow and six local authority/private nurseries in North Ayrshire. Each nursery was visited three times by the experimenter for two days in a row over a two–year period. In Studies 1 and 2, participating children (the same sample for each) were divided into three age groups (1.5 to 2.0 years, 2.5 to 3.0 years, and 3.5 to 4.0 years) with clear gaps between them. In Studies 3 and 4, children in the age range 1.5 to 2.0 were excluded, and only the middle and oldest groups participated. This choice of age groups reflected the need to cover a range of different developmental levels, as outlined above, from the first signs of mentalistic understanding in 1.5 year olds (Reddy & Simone, 1995), to the precursors of representational theory of mind in 3.5 year olds.

All four studies were of mixed between-subjects (age group) and within-subjects (reference type) design, and each involved 8 trials, specified by 2 contexts (violation vs. non-violation) x 2 grammatical forms (direct vs. indirect requests) x 2 instances utilising different objects. The materials for each study consisted of age-appropriate collections of soft toy animals and accompanying items. In Study 1, children had to determine whether the toy requested was in the immediate or non-immediate context. For the latter type of request, the requirement was for children to perform a search operation and locate toys that were out of their line of sight in the non-immediate context. For Study 2, they had to work out whether the toy requested was a part of the present or a past context, in the latter case suspending their focus on the here and now (a current joint activity involving novel toys) and grasping that the experimenter was referring to a past joint activity involving different toys that had taken place the day before.

In Study 3, an evidently happy context was initially established for two pairs of soft toys animals, but requests were made for both happy and sad animals. To retrieve the latter, children needed to work out that one pair of animals had become non-evidently sad (i.e. emotion felt but not expressed) due to a change in circumstances as the scenario unfolded, and determine that such requests were for members of this pair. Study 4 paralleled previously outlined false belief research (e.g. Siegal & Beattie, 1991; Clements & Perner, 1994; Surian and Leslie, 1999; Yadzi *et al*., 2006) in using simultaneous testing of both 'an evident true belief' (held by members of pair of toy characters who witnessed the swap of the contents of a well known container), and 'a non-evident false belief' (held by another pair of toy characters who did not witness the swap). Children need to distinguish between different sources of belief in order to determine accurately which pair of toys the experimenter was referring to when making a request couched in terms of character belief.

The grammatical form of the requests was varied in each study between direct requests i.e. ('Give me the X', where X refers to a toy item) and indirect requests i.e. ('Have you got the X?') following Babelot & Marcos (1999). This work indicated that these are the most and least likely, respectively, to elicit an accurate response from pre-school children. In this sense then, they provided a test of the easiest and most difficult forms the violations being tested could take. As noted earlier, the degree of pragmatic sensitivity shown in responding to the two therefore provided a sensitive index of how secure children's understanding of each form of reference was: if children could decode the source of the non-immediate reference at the same time as responding in a fashion sensitive to the form of indirect requests, decoding both forms of request, this would suggest more established competence than ability to decode the reference only where the form of the request involved no embedded implicature.

In addition to accuracy of response to the different types of request, attention was therefore also paid to the form of response, and how far this varied between action and verbal response, according to form of request. Since direct requests essentially call for an action in response, whereas indirect requests arguably call more for a verbal response indicating location, such differentiation provides the key measure of pragmatic sensitivity, which could be used to further establish children's level of competence. At a more rudimentary level, a further index of pragmatic sensitivity was also available in terms of children understanding that the different forms of request required a response, even if the answer were uncertain. If children gave an 'actively incorrect response' (e.g. by retrieving the wrong object or making a verbal or gestured response that showed no understanding of the object referred to), in contrast to not responding to the request at all, they were taken as showing this basic level of sensitivity. Incomprehension responses (facial puzzlement or uncertain verbal statements, such as 'don't know') were also categorised as 'actively incorrect responding', since they too evidenced recognition of the need to provide some form of acknowledgement of the request. The coding of children's responses was based on the methodology employed by Babelot & Marcos

(1999), expanded to meet the specific requirements of the present studies, and remained consistent throughout. Children's responses were noted during testing by the experimenter, but also recorded on videotape, as a back-up for scoring.

Finally, as noted above, Studies 3 and 4 employed standard measures of false belief understanding. For the purposes of these studies, however, it was decided to focus primarily on the 'Look first' FB test i.e. Siegal & Beattie's (1991) version of the unexpected transfer test (Wimmer & Perner, 1983). In respect of Study 4, and the understanding of false belief, the 'Look first' FB test was also evaluated with the UCT (Perner, Leekam & Wimmer, 1987) The reasoning here was based on the distinction between Generalised Conversational Implicatures (GCIs) and Particularised Conversational Implicatures (PCIs) (Grice, 1975; Levinson, 2000; Carston, 2004). More specifically, the original questions posed in both the unexpected transfer test (e.g. 'Where will Maxi look?') and in the UCT (e.g. 'What will the next child to enter the room think is inside the smarties tube?) may be argued to be GCIs, since decoding these original questions to mean 'What will Maxi do first? (cf. Siegal & Beattie, 1991) or 'What do you think was in the smarties tube when you first saw it'? (cf. Gopnik & Astington, 1988) requires knowledge that *in general* this is what a question of this form would signify, and there is no such context-specific cue given to aid this interpretation in the original standard false belief questions. In contrast, Siegal's form of the question does provide such a cue, and may therefore be legitimately regarded as a PCI. Since the focus of the present research was on competence at decoding violations of relevance that rested on context-dependent cuing, with the particular contextual grounds provided by the experimenter's references to the non-immediate context (Study 1); shared past event context (Study 2); non-evident emotion context (Study 3) and non-evident belief context (Study 4). Such references also took the form of PCIs because the context is constrained by the fact that the non-evident references utilise the same form as the evident ones, signalling that the child's search should therefore be targeted at events of the same kind, even if these are not immediately apparent. Specifically for Study 4, it was also considered that responses to the PCI question inherent in the 'Look first' FB test, would be better related to understanding references to non-evident belief given the close similarity in both form and content.

1.9.2 Broad hypotheses

In broad terms, it was anticipated that older children would show greater relative pragmatic grasp and sensitivity by 1) giving more correct responses to requests based on references to non-immediate contexts; 2) giving more correct responses to requests in the form of indirect requests; and 3) utilising response forms that were better tailored to request form (actions for direct requests vs. verbal responses for indirect requests), or at least heeded conversational requirements by taking an active form when incorrect. It was also anticipated that the oldest age group would in general show good absolute levels of performance. Support for this view was apparent in Ninio & Snow's (1996) observational research, and from examination of the CHILDES Language Database (MacWhinney, 2001) both of which indicate that as children approach their fourth birthday, they receive less conversational, historical and psychological support from caregivers during discussions. The implication of this reduced reliance on adult support is that they should be able to adapt reasonably well to the circumstances of experimental testing, where support from the experimenter may be more formal and constrained at best.

It was further predicted that, taking pragmatic sensitivity as an additional measure, across the four studies ability to respond appropriately and consistently to non-immediate references would emerge in the sequence described above. It was also expected that the evidence would be in keeping with gradual or incremental change as opposed to stage-like, age-related conceptual shifts. In statistical terms, such evidence would take the form of 'main effect only' outcomes, with an absence of interactions involving age and reference type which would indicate rapid or discontinuous developmental change in children's understanding.

As regards Studies 3 and 4 and the comparison of performance to that on tests of false belief understanding, it was hypothesised that children's competences in understanding references to non-evident internal states would be better related to Siegal & Beattie's (1991) 'Look first' FB test, as both utilised PCIs. In contrast, the standard false belief test questions rest on GCIs, which were considered to flout the maxim of manner, introducing an additional source of variance into performance. Therefore, in relation to Study 4, and given the PCI vs. GCI distinction, it was further predicted that references to evident and non-evident belief states would be better related to the 'Look first' FB test rather than the UCT (Perner, Leekam & Wimmer (1987). Given the widespread reporting of relationships between language competence and theory of mind, it was expected that accurate decoding of references would be associated with the expressive measures used here. However, it was also hoped that multiple regression analyses examining the interplay of influences on performance might help reveal more specific detail regarding the nature of such relationships, and also the role of working memory, implicated in conversational skill too, as noted earlier.

CHAPTER 2

1.5 to 4.0 year old children's understanding of references to the immediate and non-immediate context.

2.1 Introduction

To summarise thus far, as discussed in Chapter 1, research relating to lean and rich interpretations of the development of young children's understanding of conversational reference examine their abilities primarily with regard to objects that are the immediate joint focus of attention. According to Butterworth & Jarrett's (1991) lean exposition, infants aged 18 months show an awareness of embodied minds through direct perception by recognising that changes in the direction of their mother's gaze are good predictors of an object's location. A broader view, based on the research of Baldwin & Moses (1994) adopts a rich interpretation of young children's perceptual abilities, where infants are said to be performing a mentalistic analysis of others' communicative behaviour e.g. an infant will shift its gaze to the named object of the mother's line-of-regard, inferring a connection between percept and referent, facilitating the learning of new words. This is taken as reasonable grounds to conclude a fledgling insight into the minds of others on the part of young children. Consistent with this, Diesendruck *et al*., (2004) found that children aged 2 years are sensitive to the relevance of a speaker's communicative intentions under such circumstances i.e. children learned the word for an object when the context involved intended actions on behalf of the speaker but not accidental ones.

These findings are also in line with the concept of “attention-goal psychology” (Baron-Cohen, 1993) where a child’s ability to interpret the focus of an adult’s attention is taken to indicate an understanding that the caregiver is interested in the object in question. Similarly, Bretherton’s (1991) findings suggest young children make use of rich interpretation in their communications with adults. Infants at 9 months of age convey intentional messages that may presuppose a basic shared framework of mutual understanding e.g. via imitation, pointing and gaze direction, and by 18 to 20 months of age, children use single words to talk about a range of internal states, such as hunger, thirst and disgust

Despite the evidence for a rich interpretation of children’s abilities to decode both gaze and reference with regard to objects in the immediate focus of attention, however, Baldwin & Moses (1994) point out that little research has been directed towards children’s decoding of the referential intentions of adults with respect to objects that lie outside of the immediate focus of attention or to absent objects. Early findings from language studies (Baldwin, 1993b; Sachs, 1983) suggest that understanding of such references might be expected to develop from the middle of the second year. However, Baldwin & Moses acknowledge that how children might comprehend that attention and thought persist and support a reference to objects not present in the immediate surround remains largely uncharted. They suggest that the study of how children make sense of adult utterances in the absence of tangible objects may pave the way towards a better understanding of their mental state attribution: language is the compelling means of sharing information with regard to intangible objects, and investigation of children’s interpretation of adult’s communications may provide an insight as to how they decode such references more generally.

Pragmatic language research also suggests that children's ability to decode referential intentions and their mentalistic understanding more generally may be bound up with their grasp of conversational practices (Ninio & Snow, 1996; Babelot & Marcos, 1999; Diesendruck *et al*., 2004). Attentive caregivers recognise the need to adjust their speech to suit the characteristics of young children by making their utterances shorter, truthful, clearer and relevant to the joint focus of attention (cf. Grice, 1975). However, violations of this co-operative principle arise in adult, and adult to children's conversation, requiring the hearer to register a conversational implicature by recognising that something is implied or meant beyond the utterance itself. Decoding of conversational implicatures is known to be problematic for young children as the literal interpretation of a statement must be disregarded (Siegal & Peterson, 1994). Central to this achievement may be the ability to understand the different ways in which an utterance can be intended to be relevant (i.e. what the speaker might possibly have in mind) when it violates joint focus of attention. To date, however, most research has been directed at how conversational implicatures and their relevance are decoded in the adult world, with methodology being qualitative in nature and largely focused upon discourse analysis of adult conversation outside of experimental settings. (Grice, 1975; Ninio & Snow, 1996; Sperber & Wilson, 2002).

Other qualitative methodological approaches have been employed to examine children's pragmatic language skills more generally. Both the examination of the CHILDES Language Database (MacWhinney, 2001) presented in Chapter 1 and Ninio & Snow's (1996) naturalistic, observational studies suggest that caregivers, having worked to establish an expectation of relevance on the part of infants, gradually introduce violations in the form of references to objects, topics or things that lie outside joint focus of attention. By doing so, they effectively illustrate the ways in which references to non-immediate objects can be made, and the way in which these depend on particular types of shared (or shareable) mental connections.

Coupled with this finding is the fact that there are a very limited number of ways in which apparent violations of relevance can be resolved and caregivers seem to play on this. According to Ninio & Snow (1996), with young infants such violations take references to objects that are in children's immediate view as their starting point or baseline, but extend these in the first instance to objects that are behind them or hidden from view – what they frame as the first sub-maxim of reference to non-immediate objects. Grasp of this first sub-maxim may guide search strategies in young children to re-establish relevance when this appears to have been violated. From categorisation of the communicative acts apparent in 30-minute verbal interchanges in naturalistic-observational studies of mothers and children at play in their own home, acquisition of the first sub-maxim appears to be evident in children aged 18 months and over. On this basis of this evidence, they argue that apparent violations – and concomitant grasp – start with the first sub-maxim, and then extend in turn to the second, third and fourth sub-maxims of references to shared past events, non-evident emotions and non-evident belief states respectively.

2.1.1 Rationale for the design of Study 1 and related aspects of Studies 2, 3 and 4

Ninio & Snow's (1996) research has, to some extent, provided an answer to Baldwin & Moses' (1994) question about the uncharted nature of children's ability to understand references to non-immediate objects or events. However, this research was naturalistic and observational, and focused on children's performance as opposed to their competence. In contrast, the studies reported below and in the next chapter employed experimental methods to measure children's competence in a systematic fashion, via controlled investigation of children's understanding of physically and temporally remote referents. Study 1 contrasted grasp of references to immediate and to non-immediate objects, whilst Study 2 compared understanding of references to present and shared past events. As noted above, according to Ninio & Snow's (1996) developmental progression in conversation, there is the expectation that these competencies are acquired first and second respectively. To facilitate direct comparison of performance, and establish whether there was evidence of this being the case, the same sample of child participants was used for both studies, with the primer for Study 2, i.e. the focal past shared event, taking place immediately following Study 1, on the preceding day.

For both studies, children were divided according to the same age groups. The youngest age group was aged 1 year and 6 months to 2 years old; the middle age group was aged 2 years and 6 months to 3 years old and the oldest group 3 years and 6 months to 4 years old. The choice of age group reflected the need to cover the full range of observed competencies according to age as found by Ninio & Snow (1996) i.e. the appearance of understanding of the first sub-maxim in children aged 18 months and over to appearance of the third and fourth in children aged 36 months and over. The six-month gap between the three age groups was set to better highlight developmental trends.

The choice of age groups for Study 1 and 2 also covered the period of initial emergence of pragmatic language skills according to Babelot & Marcos (1999). This research reported that children aged 1 year and 7 months were capable of giving an adult a correct object in response to a request when that object was present in the immediate focus of attention. In addition, children aged 2 years and 5 months were found to be able to relate the context of a request to its form: when the object referred to was outside the immediate focus of attention, children focused on the literal meaning of indirect requests (‘Have you got…?’), taking these as requests for information. Thus although unable to decode the implicature involved, children at this age were at least sensitive to variation in referential form and context, and aware of the requirement to respond to a request even when the answer was uncertain.

It has been previously acknowledged how the methodology employed by Babelot & Marcos (1999) might be suitably utilised to examine children's pragmatic sensitivity in the context of references to objects, events, non-evident emotions and beliefs which are not immediately apparent, and provide a more fine-grained measure of competence in relation to these. In particular, it was argued that relative security of grasp of different types of reference would be revealed by the extent to which the forms of responses indicated a) recognition that a response is required, b) modification of the character of that response according to context or grammatical form. In fact, there is logically a third level of pragmatic sensitivity, in which the implicature of indirect requests is understood, and both forms of request produce accurate action responses (as distinct from action responses per se, which would just indicate an undifferentiated bias in response form). There is no evidence on the point at which this level of sensitivity emerges for any of the reference types being investigated here, but analyses examined how far it was apparent.

Utilising Babelot & Marcos' methods also provided a basis for comparing the data obtained here to an external source, yielding a means of judging its reliability. In making such comparisons, however, it is also important to bear in mind four differences between the present research and Babelot & Marcos' experiments.

First, in Babelot & Marcos' (1999) experiment, toy objects in the non-immediate context were in view but positioned behind the experimenter. Study 1 in the present research contained a more demanding non-immediate context, in which objects in the non-immediate context were positioned out of view behind the child. The object here was to create a more realistic measure of competence, by setting up a scenario in which objects would be visible to the person making the request, but out of sight to the child, requiring them to realise the need to instigate a search operation. It was considered that this would also serve to make the request to retrieve the objects a more natural one. This counteracted any tendency that might have been operating in Babelot & Marcos' task design for children to respond to requests for non-immediate objects by making verbal responses simply as a matter of convenience rather than one of understanding - an effect which might have coloured the data. By extension, Studies 2, 3 and 4 also used situations in which the referent might reasonably be regarded as evident to the requester, and retrievable by the child, but where the child would have to engage in a search of one form or another in order to do this successfully.

Second, the studies reported here used only two of the four grammatical utterances employed by Babelot & Marcos (1999) i.e. imperative (direct) requests and directive questions (indirect requests). In their research, these were found to be the most and least likely utterances respectively to elicit an appropriate response from children, and thus constituted suitable extremities of demand for ascertaining degree of competence, whilst at the same time restricting the number of trials to workable levels for the age range being tested.

Third, Babelot & Marcos' (1999) responses categories were expanded slightly. In particular, their Information Response type (where children made a verbal response to a request rather than providing the requested object), was divided into Indicative (correct verbal/gestural responses) and Non-Indicative Information Responses (incorrect uninformative verbal responses) in order to help distinguish actively incorrect responses at the verbal level as well as at the level of action.

Fourth, the third, older group of children aged 3 years and 6 months to 4 years old were included to better highlight potential developmental gaps and trends both within and across studies.

2.1.2 Summary of task and hypotheses for Study 1

To summarise, then, the basic task that children were asked to perform for Study 1 focused on a prototypical situation of apparent violation of relevance within a verbal request for an object made to a child, this being potentially and plausibly resolvable by the child directing their attention to available but non-obvious information. Children's ability to make the appropriate resolution was determined by whether they gave the mentioned object (correct action response) or provided suitable information about its location (correct verbal/gestural response). Competence on trials of this kind was compared to that on trials where requests were made for objects in the immediate focus of attention.

The form of both types of trial was further varied by the experimenter's use of either direct requests ("Give me the X") or indirect requests ("Have you got the X?"). The materials consisted of eight age-appropriate soft toy animals, four of which were placed in front of the child in the immediate focus of attention (immediate context), and four out of sight behind the child (non-immediate context). To reduce differences in the relative accessibility to the child of the two contexts, the trials took place within a room where the non-immediate toy locations were familiar to children even though out of sight, and thus might plausibly form the focus of a search for a referent. Participating children were each exposed to two instances of all four types of trial: immediate/direct request, immediate/indirect request, non-immediate/direct request, and non-immediate/indirect request.

It was hypothesised that:

1) Children in the oldest age group (3.5 to 4 years) would make more correct responses than those in the younger groups to requests for objects in both the immediate and non-immediate contexts; but that all three age groups would show some ability to decode references to non-immediate objects, based on Ninio & Snow's (1996) claims about the point at which children are introduced to these.

2) Older children would show a higher degree of pragmatic sensitivity to referential form in both contexts, albeit primarily at the level of recognition of differences in request form, among the middle age group at least, by extension from the data reported by Babelot & Marcos (1999). Therefore, older children. would give more action responses to direct requests and more verbal responses to indirect requests regardless of reference type, indicating a more secure understanding of non-immediate references as well as immediate.

3) Children's performance in terms of accuracy of response to both immediate and non-immediate references would be related to both memory and expressive language ability, reflecting the impact of processing demands, and the extent of competence to deal with these.

Beyond this, it was uncertain whether all age groups would exhibit differences in grasp of references to immediate and non-immediate contexts. On the one hand, the theoretical impact of parental support would suggest a relatively incremental process of change, rather than the sudden emergence of more general insight; but on the other, on the basis of Ninio & Snow's (1996) analysis, the oldest age group ought to have been exposed to references to non-immediate contexts for up to two years, and therefore might have attained high levels of competence, leading to performance at ceiling levels. Thus whilst main effects of age and interactions between age and request form were anticipated in terms of both accuracy and response type, as implied by hypotheses 1 to 3, no firm prediction was made as to the likelihood of main effects of reference type, or interactions between age and reference type.

2.2 Methodology

2.2.1 Design

Study 1 employed a request task of the form used by Babelot & Marcos (1999) within a mixed between-and within-subjects design. There was one between-subjects variable, age (youngest, middle or oldest age group). The within-subjects variables were reference type (to the immediate or non-immediate context) and request form (a direct or indirect request). Levels of these were crossed to give eight trials, two for each combination. The dependent variable was the coding of responses into sub-types of action, verbal and non-response (see 2.2.4). Language tests designed to measure variation in children's expressive language were also conducted with each child to examine the relationship between language competence and response accuracy. The tests used were either the 'First Word Test' from First Word First Sentence (FWFS, 1997) for children in the age range of 1 year and 6 months to 2 years old, or British Ability Scales II, Naming Vocabulary Test (BAS II, 1996) for children in the age range 2 years to 4 years old. A measure of memory ability was also gathered for each child within the context of their subsequent performance in Study 2. For ease of reporting, data regarding the relationship between this and performance in Study 1 will be considered in the context of that study.

2.2.2. Participants

Participants were 79 pre school nursery children (40 girls, 39 boys) attending two university 'on campus' nurseries in Glasgow and six local authority/private nurseries in North Ayrshire. Each nursery was visited once by the experimenter for two days in a row over a six month period. The sample was divided into 3 age groups, with 26 children in the youngest group (14 females and 12 males; age range 1 year and 6 months to 2 years with a mean age of 1 year and 10 months), 26 children in the middle group, 14 females and 12 males; age range 2 years and 6 months to 3 years, mean 2 years and 9 months), and 27 children in the oldest group (12 females and 15 males; age range 3 years and 6 months to 4 years, mean 3 years and 8 months). All children spoke English as the language of their home. Parental consent was given with respect to all participating children and children were advised that they could withdraw from the study at any time. With regards to the youngest children, who may not have understood the 'right to withdraw', the nursery assistant was responsible for withdrawing any children who were reluctant to participate either prior to or during the experiment. Relevant local authority clearance and university ethical approval was obtained. The experimenter was given Enhanced Disclosure Scotland clearance to work with young children.

2.2.3 Materials

The principal materials consisted of a collection of soft toy animals, of a form familiar to children within these age groups, and which can be purchased from most leading toy distributors and found in both nursery and home toy collections. For the purposes of administration, the animals were arranged into two groups. Group 1 animals consisted of a *Cat, Dog, Monkey and Rabbit.* Group 2 consisted of a *Bear, Frog, Lion, and Penguin.* No animal was smaller than 15 cm or higher than 18 cm. A large green cloth mat was also used to conduct the test sessions upon nursery floors, which allowed the animals to be evenly spaced apart.

Other materials included the First Word test from the First Word First Sentence (FWFS, 1997) and the British Ability Scales II Naming Vocabulary Test (BAS II NVT) (BAS, 1996). The pro-formas shown in appendices A & B were also used by the experimenter to score children's responses during the test sessions. A video camera on tripod was utilised to record children's responses, this enabling the coding of responses to be re-checked by the experimenter and then verified by an independent judge at a later date. Permission to use and record children's responses by video, in addition to their participation approval, was sought and granted from participating nurseries and parents, who were both informed that the video footage would be erased once the project was complete.

2.2.4 Procedure

Testing took place individually within each nursery, using a separate room. Children received toys and questions in one or other of two configurations, Set A or Set B (Group 1 animals in immediate context, and Group 2 in non-immediate, or vice versa), in order to counterbalance reference type and request form across toys. In each setting, approximately equal numbers of children in each age group received the two configurations. For Set A, the four soft toy animals representing the non-immediate context, the *Bear, Frog, Lion, and Penguin*, were placed to the rear of children entering the experimental room in unobtrusive but available locations and about 15cm apart. The remaining animals were positioned on the green mat in direct view. These soft toys represented the ‘immediate context’, and comprised the *Dog, Monkey, Cat and Rabbit*. For children allocated to Set B, the positioning of the toys was reversed.

Each child was accompanied into the experimental room by a nursery assistant who remained with them during the course of their session. This enabled the child to have a familiar face present and helped to put them at their ease. Each child participated in a 15 minute ‘warm-up period’ where the experimenter introduced himself and engaged in a ‘play scenario’ with the child involving the immediate toy animals. This served to build rapport with the child and to allow familiarisation with the experimental room but not the location of non-immediate toy animals.
Indeed, care was taken to ensure that the non-immediate objects would not be seen by children entering the experimental room i.e. these were always set up to the rear

and out of direct view of where the nursery assistant and child entered the room. The mat with the immediate objects was always located to the front of children entering the room, and this formed the focus of their attention at the outset and during the warm-up period. To further guard against the possibility of children sighting the non-immediate objects the experimenter always observed children entering the room. Provision had been made to end children's involvement in Study 1 if non-immediate objects were identified by children at this point. This action was not required.

At the end of the warm-up period, when the game was finished, the language tests were administered using standard procedures, children in the youngest age group receiving the First Word Test, and those in the middle and older age groups receiving the British Ability Scales II Naming Vocabulary Test. The 'First Word Test' from First Word First Sentence (FWFS, 1997) is an object-naming test that measures children's expressive language and consists of 28 coloured photographs of objects that children, aged 1 year and 6 months to 2 years old, talk about in their first one hundred words. The photographs were visible one at a time and each photograph of an object was administrated item by item by the experimenter until the child had answered for all 28 objects. The child was shown each photograph of an object and asked 'What's that?'. The first picture was a practice item and was not scored. The experimenter noted the child's reply and whether the answer was correct, incorrect or another acceptable response. If the child failed to respond to a photograph, this item was noted as incorrect and the test proceeded to the next.

The British Ability Scales II Naming Vocabulary Test (BAS II) (1996) is an object-naming task that consists of 16 coloured illustrations of objects to measure children's expressive language in the age range, 2 years and 6 months to 4 years and 5 months. The experimenter pointed to each illustration of an object and asked the child, 'What is this?' The illustrations were visible one at a time and each illustration of an object was administered item by item by the experimenter until the child had answered to all 16 objects. Here again, the experimenter noted the child's reply and whether the answer was correct, incorrect or any other acceptable response. If the child failed to respond to an illustration the experimenter then asked, 'What is this called?' If a child failed to respond again the test proceeded to the next illustration. If a child had 5 consecutive failures or no responses the test was stopped.

When the language tests had been conducted the four animals in the immediate context were re-positioned on the mat in front of children. The precise order of these animals was varied systematically for each child, but the spacing remained the same with toys set about 15 cm apart from each other. Children were then told that the experimenter wished to ask them some questions about the toy animals. Children's ability to resolve references to both the immediate and the non-immediate context was tested by examining their responses to a series of four direct ("Give me the X") and four indirect ("Have you got the X?") requests made by the experimenter. The requests were put to children one after the other with no other dialogue (see Tables 2.1 and 2.2 for the sequence of trial types for Set A and Set B children respectively), and with a 20 second interval between them, as measured by the experimenter's watch. To help establish children's confidence in responding, three out of the first four questions concerned an animal in the child's immediate view.

Children's responses were recorded at the time by the experimenter on either the Appendix A or B pro-forma, as appropriate, and also on videotape. The video recorder had been switched on after each child's warm up period had ended and just before the language tests were administered. It was decided in advance that if children became tearful or upset during the procedure or testing that their participation would be discontinued and they could return to their nursery room. Only a very few children became upset, and then only when entering the experimental room at first, so in consultation with accompanying nursery assistants, their participation was ended at that point. No child became tearful or upset following the familiarisation session or during testing. At the end of the trials, each participating child was given a happy face sticker and thanked for their participation.

Table 2.1: Order of request by reference type and request form for Set A children

REQUEST-OBJECT	REFERENCE TYPE	REQUEST FORM
1. Give me the Dog	Immediate	Direct
2. Have you got the Bear?	Non-Immediate	Indirect
3. Have you got the Cat?	Immediate	Indirect
4. Give me the Monkey	Immediate	Direct
5. Give me the Frog	Non-Immediate	Direct
6. Have you got the Rabbit?	Immediate	Indirect
7. Give me the Lion	Non-Immediate	Direct
8. Have you got the Penguin?	Non-Immediate	Indirect

Table 2.2: Order of request by reference type and request form for Set B children.

REQUEST-OBJECT	REFERENCE TYPE	REQUEST FORM
1. Have you got the Frog?	Immediate	Indirect
2. Give me the Rabbit	Non-Immediate	Direct
3. Give me the Penguin	Immediate	Direct
4. Have you got the Lion?	Immediate	Indirect
5. Have you got the Dog?	Non immediate	Indirect
6. Give me the Bear	Immediate	Direct
7. Have you got the Monkey?	Non-Immediate	Indirect
8. Give me the Cat	Non-Immediate	Direct

2.2.5 Coding of Responses and Scoring

Request task

Children's responses as recorded by the experimenter during testing were later transcribed from videotape after data collection was complete and rechecked. At this point, they were coded into one of six types, as outlined below. Four of the response types were drawn directly from the work of Babelot & Marcos (1999). These were types 1, 3, 5 and 6. Response types 2 and 4 were related to the Information Response utilised by Babelot & Marcos (1999) but sub-divided for the specific needs of this study, which required additional differentiation according to whether a child gave a response that indicated an understanding of the object referred to or not. To check on the reliability of coding of the responses an independent judge later checked response types from video footage. The responses of five children from each of the 3 age groups were selected at random for reliability testing by the judge. The 120 responses provided by these children accounted for 19% of the total. Inter-judge agreement scores on the coding of responses are given after each of the response types.

1. Correct action response: gives or actually touches the requested object (inter-judge agreement = 98%).

2. Correct verbal/gestural response: indicates an understanding of the object referred to by describing where it is or giving a gestured answer such as pointing (inter-judge agreement = 94%).

3. Incorrect action response: gives a different object to the one requested (inter-judge agreement = 100%)

4. Incorrect/uninformative verbal response: verbal response which does not give any indication of accurate understanding of the object requested e.g. simply saying or nodding ‘Yes’ or ‘No’ (inter-judge agreement = 100%)

5. Incomprehension response: verbal or gestured puzzlement (inter-judge agreement = 100%)

6. Non-response: child continues playing, and makes no verbal statement (inter-judge agreement = 94%)

The number of children's responses in each of the six categories was totalled across trials of the same reference type and request form to give a score between 0 and 2 for each. In this way, six scores, one for each response type, were derived for each of the following trial types: immediate/direct; immediate/indirect, non-immediate/direct and non-immediate/indirect.

Also, to check on order effect, the number of children's responses in each of the six categories was totalled separately across the first questions employing the two different request forms for each reference type, (e.g. direct request 1 plus indirect request 1 for the immediate context and similarly across the second questions (e.g. direct request 2 plus indirect request 2 for the immediate context and so on) again to give a score between 0 and 2 in each case. In this way, six scores, one for each response type, were derived for immediate/first requests; immediate/second requests; non-immediate/first requests; and non-immediate/ second requests.

2.2.6 Language tests scoring

In the FWFS (1997) test a child received 2 points for correctly naming a photograph of an object, 1 point for any other associated response e.g. saying *'wash'* in response to a picture of a tap and 0 points for an incorrect answer or not answering. Since there were 27 picture objects to name, a child could score a maximum of 54 points. In the BAS II NVT (1996) test a child would receive 1 point for correctly naming a picture object or for giving any other acceptable response and 0 points for incorrect answer or not answering. As there were 16 picture objects to name, a child could score a maximum of 16 points.

2.3. Results Section

2.3.1 Summary of Analyses

Before proceeding to the main analyses, preliminary analyses were conducted to check for the effects of question order (i.e. whether there was any indication of practice effects), and also whether there was any effect of the object/request set used. A 3-way mixed ANOVA (age group x reference type x order) was conducted on the scores for the immediate/first requests, immediate/second requests, non-immediate/first requests and non-immediate/second requests for each of the six response types outlined in 2.2.5. There were no significant effects or interactions found involving question order. Preliminary analyses were also performed to check for the effects of the set used. A 4-way mixed ANOVA (age group x set x reference type x request form) was conducted on the immediate/direct, immediate/indirect, non-immediate/direct and non-immediate/indirect scores for each of the six response types. Again, there were no significant effects or interactions found in relation to the set used. As neither request order nor the set used had any impact on performance these factors will not be considered further.

For the main analyses, 3-way mixed ANOVAs (age group x reference type x request form) were computed on the immediate/direct, immediate/indirect, non-immediate/direct and non-immediate/indirect scores for each of the six response types.

A 2-way ANOVA was then computed on the overall level of correct responses collapsed across correct action and correct verbal/gestural responses. This was followed with a 3-way mixed ANOVA that was conducted to examine active vs. inactive incorrect responses (age group x reference type x active vs. inactive). The results of these analyses are reported in this order in what follows.

With the main patterns of performance on the request task identified, the relationship of correct responses to language scores was explored using correlational analyses

2.3.2 Descriptive and Inferential Results for Response Codes on Request Task

Table 2.3 shows the mean scores and standard deviations for each response type, broken down by age group, reference type and request form. Appendix J presents full results for the 3-way ANOVAs conducted on these data.

TABLE 2.3: Means and standard deviations for each response code (range = 0-2) by reference type, age group, and request form

			Correct action resp		Correct verbal resp		Incorrect action resp		Incorrect verbal resp		Non-resp	
Ref type	Age group		Dir	Indir	Dir	Indir	Dir	Indir	Dir	Indir	Dir	Indir
IMM	Young	Mean	0.73	0.38	0.30	0.53	0.00	0.38	0.11	0.26	0.80	0.69
		SD	*0.82*	*0.63*	*0.61*	*0.70*	*0.00*	*0.19*	*0.32*	*0.60*	*0.89*	*0.83*
	Middle	Mean	1.23	1.00	0.30	0.42	0.11	0.19	0.15	0.26	0.11	0.11
		SD	*0.95*	*0.93*	*0.61*	*0.70*	*0.43*	*0.49*	*0.54*	*0.66*	*0.32*	*0.32*
	Oldest	Mean	1.62	1.07	0.22	0.81	0.00	0.00	0.11	0.11	0.00	0.00
		SD	*0.74*	*0.82*	*0.57*	*0.83*	*0.00*	*0.00*	*0.42*	*0.42*	*0.00*	*0.00*
	Total	*Mean*	1.20	0.82	0.27	0.59	0.03	0.07	0.12	0.21	0.30	0.26
		SD	*0.91*	*0.85*	*0.59*	*0.75*	*0.25*	*0.31*	*0.43*	*0.57*	*0.62*	*0.59*
NON IMM	Young	Mean	0.26	0.19	0.42	0.30	0.26	0.26	0.07	0.11	1.07	1.11
		SD	*0.60*	*0.56*	*0.70*	*0.61*	*0.60*	*0.60*	*0.27*	*0.43*	*0.89*	*0.86*
	Middle	Mean	0.61	0.53	0.46	0.61	0.23	0.07	0.19	0.26	0.30	0.30
		SD	*0.89*	*0.81*	*0.76*	*0.80*	*0.42*	*0.27*	*0.56*	*0.66*	*0.67*	*0.67*
	Oldest	Mean	0.59	0.40	0.74	1.11	0.18	0.14	0.22	0.18	0.14	0.11
		SD	*0.84*	*0.74*	*0.90*	*0.89*	*0.48*	*0.45*	*0.57*	*0.55*	*0.45*	*0.42*
	Total	*Mean*	0.49	0.38	0.54	0.68	0.22	0.16	0.16	0.18	0.50	0.50
		SD	*0.79*	*0.72*	*0.79*	*0.84*	*0.50*	*0.46*	*0.49*	*0.55*	*0.79*	*0.79*

Ref type = Reference type; IMM = reference to immediate context ; NON IMM = reference to non-immediate context; resp = response type; Dir = direct response form and Indir = indirect response form. Incomprehension Responses are excluded from this table due to their infrequency of use, representing only 2.5% of total responses given by children.

1. Correct action responses.

Focusing first of all on main effects, as can be seen from Table 2.3, the oldest age group gave the most correct action responses overall, followed by the middle group and then the youngest. The overall effect of age was significant, although not strongly so ($F (1,76) = 5.88, p < .05$). Tukey (HSD) post-hoc tests indicated that only the differences between the youngest and the two older age groups were reliable ($p < .05$). Children in all 3 age groups gave more correct action responses to requests for immediate objects than non-immediate, and this effect was highly significant ($F (1,76) = 45.90, p < .001$). A substantial main effect on correct action responses was also found for request form, with more responses of this type being given to an direct requests than to indirect ($F (1,76) = 18.73, p < .001$).

Within this wider pattern, there was a weakly significant interaction between reference type and age ($F (2,76) = 3.29, p < .05$) with the gap between the contexts referred to tending to narrow amongst the middle age group, but then to widen again amongst the oldest. There was also a somewhat stronger interaction between reference type and request form ($F (2,76) = 6.10, p < .05$), with the difference in incidence of correct action responses between direct and indirect requests wider for references to the immediate context than for non-immediate references.

Overall, then correct action responses increased generally between 2 and 2.5 years, and then showed a further increase only for direct requests for immediate objects, where they were in any case more likely to occur, as Babelot & Marcos (1999) found. The greater incidence of correct action responses for direct requests than for indirect was consistent with the predicted differentiation between request forms, except that it was apparent in all three age groups for immediate references, and generally absent for non-immediate references. This lack of differentiation between request forms in the context of non-immediate objects, suggests the lower incidence of correct action responses here may indicate poorer understanding of such references, regardless of age group, relative to references to immediate objects.

2. Correct verbal/gestural responses.

The patterns for correct action responses are clarified further by the analysis of correct verbal/gestural responses. In general, these tended to occur more in response to requests for objects in the non-immediate context, although the main effect of reference type was not quite significant ($F (1,76) = 3.90$, $p = .052$). Consistent with Babelot & Marcos' (1999) data, there was also a substantial significant main effect of request form, with more correct verbal/gestural responses given to indirect requests than to direct ($F (1,76) = 15.78$ $p < .001$). This contrasted with the tendency to give correct action responses more often to direct requests, as noted above.

There was also a modestly significant interaction between request form and age, ($F (2,76) = 5.38$, $p < .05$) with children in the middle and oldest groups exemplifying the tendency to give more correct verbal/gestural responses to indirect requests increasingly strongly. In contrast, the youngest group showed little differentiation between request form in the incidence of these responses. The interaction between reference type and request form was not quite significant at ($F_{(2,76)} = 3.94$, $p = .051$), but the trend was for differences in the occurrence of correct verbal/gestural responses between direct and indirect requests to be slightly greater in the immediate context.

Overall, then, there was some tendency for correct verbal/gestural responses to occur as counterparts to correct action responses, in a manner similar to that noted by Babelot & Marcos (1999). Thus they tended to be given more commonly in response to indirect requests, and where requests were for objects in the non-immediate context. This differentiation increased with age, though, and unlike correct action responses, it was absent in the youngest age group. Moreover, it was only in the oldest age group that there was rough parity between the frequency of correct action and correct verbal/gestural response (albeit on different trials). The data therefore suggest relatively clear pragmatic sensitivity at the level of differentiating request forms amongst this age group, but perhaps only a bias towards action responses amongst the youngest and middle age groups, coupled with variation in understanding of reference types and request forms.

3. Overall corrrect responses

To clarify these points further, and to address more directly the question of whether children understood better what was meant by references to the immediate context than to the non- immediate, and how far this changed with age, both forms of correct response needed to be examined together. They were therefore collapsed across request form to give a measure of overall correct reponses. Table 2.4 shows the frequency and means of these, broken down by context and age group.

Table 2.4: Total correct responses, mean scores and sd's by age group and immediate vs. non-immediate reference types

Age Group	Total Correct Responses Immediate Context	Mean Correct Responses Immediate Context	*St. Dev*	Total Correct Responses Non-Immediate Context	Mean Correct Responses Non-Immediate Context	*St. Dev*
Youngest	51	1.96	*1.68*	31	1.19	*1.64*
Middle	77	2.96	*1.42*	58	2.23	*1.72*
Oldest	101	3.74	*0 .81*	77	2.85	*1.53*
Total	229	2.89	*1.52*	166	2.10	*1.75*

In total there were 395 correct response types given out of a possible 632. Total correct response types made up 62.5% of overall responses.

From Table 2.4, it can be seen that the oldest children gave the most correct responses followed by the middle age group and the youngest age group. The results of a 2-way mixed ANOVA (age group x reference type) conducted on these data found a significant main effect of age ($F(2,76) = 10.68$, $p < .001$). Tukey (HSD) post-hoc tests indicated that the difference between the youngest and middle group was reliable at $p < .05$, and that between the youngest and oldest group at $p < .001$. Table 2.4 also shows that in all three age groups more correct responses were given to requests for objects in the immediate context, and this main effect was highly significant ($F(1,76) = 30.59$, $p < .001$). The oldest children gave near to maximum correct responses for references to the immediate context but not for those to the non-immediate context, despite performing better than younger children. Overall, then, children understood better what was meant by references to the immediate context, and though understanding of non-immediate references changed with age, there was no significant interaction between age and reference type. This indicates that children progressed incrementally, rather than showing a sudden increase in grasp of non-immediate references. Figure 2.1 gives an indication of these trends, and makes plain both the progression with age and the lack of an interaction effect.

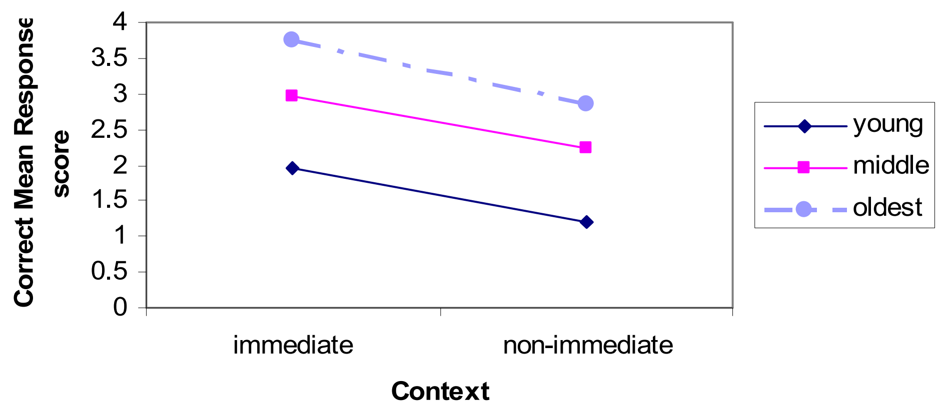


Fig 2.1: Overall mean correct responses to immediate and non-immediate contexts

4. Incorrect action responses.

Table 2.3 shows that incorrect action responses were given infrequently to requests for objects in the immediate context, and slightly less so among all groups to requests relating to the non-immediate context. This gave rise to a significant main effect for reference type ($F(1,76) = 7.18$, $p < .05$). No other effects were significant

5. Incorrect/uninformative verbal responses.

A similar pattern is evident in Table 2.3 for incorrect/uninformative verbal responses, save that these were somewhat more evenly distributed. As a result, there were no significant main effects or interactions.

6. Incomprehension Responses.

Due to their infrequency of occurrence, incomprehension responses were not subjected to statistical analysis.

7. Non-responses.

As is apparent in Table 2.3, the incidence of non-responses occurred predominantly among the youngest age group, giving rise to a strong, significant main effect of age ($F (1,76) = 20.02$, $p < .001$). Employing the Tukey (HSD) post-hoc test, differences were found to be reliable between the youngest and the middle group at $p < .001$, and between the youngest and oldest groups, also at $p < .001$. A significant main effect was also found for reference type, though ($F (1,76) = 17.28$, $p < .001$), with considerably more non-responses given to requests for objects in the non-immediate context.

8. Analysis of active incorrect vs. inactive incorrect responses.

It would appear then that when the youngest children did not understand what was being asked, they were more likely to give no response at all, whereas children in the middle and oldest age groups were more likely to give an actively incorrect response i.e. an incorrect action of incorrect/uninformative verbal response. This would suggest that the middle and older groups understood that some form of response was required, even when they did not understand what the actual referent was. In order to investigate this, an analysis was undertaken, comparing the incidence of non-responses with the combined incidence of incorrect action responses and incorrect/uninformative verbal responses as a within-subjects factor, with age group and reference type as further factors. The relevant data are presented in Table 2.5.

Table 2.5: Active incorrect vs. inactive incorrect responses by age group and reference type

Age Group	Total Inactive Incorrect Immediate Context	Mean Inactive Incorrect Immediate Context	*St. Dev*	Total Inactive Incorrect Non-immediate Context	Mean Inactive Incorrect Non Immediate Context	*St. Dev*
Young	39	1.50	*1.58*	57	2.19	*1.64*
Middle	6	0.23	*0.65*	16	0.61	*1.32*
Oldest	0	0.00	*0.00*	7	0.25	*.81*
Total	45	0.56	*1.17*	80	1.01	*1.54*
Age Group	Total Active Incorrect Responses-Immediate Context	Mean Active Incorrect Responses-Immediate Context	*St. Dev*	Total Active Incorrect Responses Non-immediate context	Mean Active Incorrect Responses Non-immediate context	*St. Dev.*
Young	11	0.42	*0.90*	19	0.73	*1.25*
Middle	19	0.73	*1.34*	20	0.76	*1.30*
Oldest	6	0.22	*0.80*	20	0.74	*1.31*
Total	36	0.45	*1.04*	59	0.74	*1.27*

In total there were 220 active incorrect and inactive incorrect types given out of a possible 632. Total active incorrect and inactive types made up 35% of overall responses.

Considering the main effects initially, a 3-way mixed ANOVA (age group x reference type x active vs. inactive) (see Appendix K) found a significant main effect of age ($F(2, 76) = 13.41$, $p < .001$), with the youngest groups giving most incorrect responses overall. A follow up Tukey (HSD) post-hoc test indicated that the significant effect of age was only reliable between the youngest and middle group ($p < .05$) and between the youngest and oldest group ($p < .001$). There was also a significant main effect of reference type ($F(1,76) = 21.11$, $p < .001$), with more incorrect responses given to requests for objects in the non-immediate context, reflecting the pattern found for correct responses. In addition, there was a significant interaction between type of incorrect response (active vs. inactive) and age group ($F (2,76) = 9.03$, $p < .001$), confirming that when the youngest children failed to understand what was being asked, they were more likely to give no response at all.

2.3.3 Language and Correct Response Scores

Because of the different language measures used with the youngest and the two older age groups, it was necessary to examine the relationship between language and correct response scores separately for these groups. Tables 2.6 and 2.7 respectively show the means and standard deviations for FWFS and BAS II NVT scores, together with total correct response scores for the relevant children across both reference types and both request forms. Given the lack of interaction between age and reference type effects, such collapsing of correct responses was considered appropriate, since in general terms the resulting total indexed performance under both conditions.

Table 2.6: Means and standard deviations for First Word First Sentence Language Test (FWFS) and correct response scores for the youngest age group.

Youngest Group	**N**	**Mean**	**Std. Dev.**
FWFS Score	26	31.9	12.4
Correct Response Score	26	3.38	3.18

Table 2.7: Means and standard deviations for British Ability Scales II Naming Vocabulary Language Test (BAS II) and correct response scores for the middle and oldest age groups

Middle and Oldest groups	**N**	**Mean**	**Std. Dev.**
BAS II NVT Score	53	14.5	2.24
Correct Response Score	53	5.9	2.61

Pearson correlation coefficients were computed between language scores and number of correct responses for the two sets of groups. For the youngest age group, FWFS scores and correct response scores were found to be highly positively correlated ($r = .69$, $p < .001$, one-tailed). An identical correlation was found between BAS II scores and correct response among the two older age groups ($r = .69$, $p < .001$, one-tailed). Thus there was a consistent and strong linear relationship between general language ability and understanding of reference right across the age range sampled.

2.3 Discussion

It was hypothesised prior to conducting Study 1 that children in the oldest age group would make more correct responses than those in the younger groups, but that all three age groups would show some ability to decode references to non-immediate objects; that older children would show a higher degree of pragmatic sensitivity and that children's performance in terms of accuracy of response to both immediate and non-immediate references would be related to both memory and expressive language ability. Firm predictions were not made about the scale of differences that would be found between grasp of immediate and non-immediate references in each age group, due to a lack of any clear basis for these.

Taking each of the predictions in turn, and firstly it was found that the oldest children did make most correct responses, as hypothesised, and that performance for both immediate and non-immediate references showed in fact a more-or-less linear increase with age. At the same time, those in the youngest age group made correct responses on more than 25% of trials involving requests for non-immediate objects, and this rose to more than 50% amongst the middle age group. There is no reason to think that chance responding affected performance on such trials, since children had to identify the location of the non-immediate objects before they could make any form of response that might prove accurate, and in fact where accuracy was lowest, amongst the youngest children, actively incorrect responses were infrequent. Thus the data are consistent with children right across this age range showing some grasp of reference to non-immediate objects, with this increasing in gradual, incremental fashion.

With regard to the second hypothesis, there was again good evidence of incremental growth in pragmatic sensitivity. As noted already, children in the youngest group made correct responses on at least some occasions to both forms of request for both types of referent. However, they showed little differentiation of response to direct and indirect forms of request, with verbal/gestural responses as likely on the whole to direct requests as to indirect, and action responses only being more likely for direct requests for immediate objects. In addition, rather than making actively incorrect responses, children in this age group tended only to make active responses when these were accurate; if they were unable to identify the appropriate referent, then they typically made no response at all. The general picture for this age group therefore is of relatively uncontrolled responding solely to those references that they understood, and no real appreciation that some form of response was required regardless of whether they knew the answer or not. For these children, then, pragmatic sensitivity seemed not to have even reached the first level.

For children in the middle age group, there remained little sign of differentiation between direct and indirect forms of request, with verbal/gestural responses only marginally more likely for indirect requests, and action responses only marginally more likely for direct requests. As with the youngest age group, action responses were somewhat more likely to requests for objects in the immediate context, but regardless of the form of request, suggesting something of a bias towards this kind of response. Broadly, then, there was still some indication of uncontrolled responding. Where they differed from the youngest age group, however was in the relative absence of non-responses, suggesting that they had at least progressed to the first level of pragmatic sensitivity by this point.

Despite the linear increase in accuracy of response between the middle and oldest groups, the pattern of responding for the latter suggests a qualitative shift in pragmatic sensitivity, and much more controlled responding in general. Here, there was a clear differentiation of response form between direct and indirect requests, with action responses around 50% more likely to direct requests, regardless of reference type; and verbal/gestural responses between two and four times as likely to indirect requests, though this effect was more pronounced where the reference was to an object in the immediate context. This pattern, coupled with the fact that non-responses were almost entirely absent, suggests that children in this age group, had achieved the second level of pragmatic sensitivity.

Thus although children in the oldest age group seemed in general not to grasp the fact that indirect requests called for an action response in the same way as direct, in general their understanding of both forms of reference and of the pragmatics of the exchange with the experimenter seemed broadly secure, something that was much less evidently true for the youngest and middle groups.

Taken overall, then, the data provide considerable support for the efficacy of using an assessment of pragmatic sensitivity alongside accuracy of responding to judge the extent of children's grasp of different types of reference. Using these indices together, it would seem clear that, as far as understanding of non-immediate references are concerned, children begin to show signs of an understanding of these at around 18 months, but that their grasp does not begin to be secure until 3.5 years, and even then it is not at ceiling.

On these data, it would seem to be beyond dispute that progress is gradual and incremental, rather than based on any sudden insight, and in this respect, the evidence would certainly seem to be more consistent with a role for external support over time, rather than an internal conceptual shift. What is less clear is why this might be such a protracted process. The strong linear relationship between grasp of immediate vs. non-immediate references and broader object naming ability – with nearly a 50% overlap in variance – suggests the presence of internal cognitive constraints as well as any putative effect of external support. This issue will be revisited in the context of Study 2, where memory data were collected on the self-same sample of children, alongside measures of their grasp of references to shared past events, allowing further investigation of the influences on Study 1 performance.

It is important to conclude by considering some methodological points, and in particular, how the results for Study 1 might have been affected by the differences from the task design employed by Babelot & Marcos (1999). In this respect, one particularly striking point was the replication of Babelot & Marcos' general finding that requests for objects in the immediate context tended to produce more action responses than those for objects in the non-immediate context. The effect was not a strong one in the present study, and yet it occurred despite the differences in the design of the contexts. In particular, as discussed earlier, the non-immediate objects here were located behind the child rather than the experimenter, so that although children potentially had to engage in a more demanding search to locate the objects, this served to create a more logical scenario in which the experimenter could see the objects, but the child was nearer to them, and thus might reasonably be asked to retrieve them. That older children in particular preferred to give verbal responses under these circumstances even when they had identified the object – and must therefore have recognised that they were nearer to it – would seem only to be interpretable as a consequence of their residual uncertainties about what they were being requested to do. In line with this, at least in part, is the fact that this tendency was stronger for indirect requests than for direct.

In other respects, certainly, it seems evident that the difference in location of the non-immediate objects did impact on performance, and that the arrangement used here provided a more suitable test of children's competences. The key pointer here is that Babelot & Marcos (1999) found that incomprehension responses were more prevalent when requests were made for objects in the non-immediate context, whilst being forced to exclude non-responses from analysis because they occurred so infrequently. This is the opposite of the pattern found in Study 1, where puzzlement was very rarely expressed, and children either made no response if they were uncertain of the referent of a request, or else made an actively incorrect response, in line with the pragmatics of the exchange.

Reviewing the nature of Babelot and Marcos' (1999) task design, it seems plausible that the high frequency of incomprehension responses among the youngest children in their research may have been the result of non-immediate objects being located out of reach of the child, and yet closer to the experimenter, leading to puzzlement on the part of the child as to perhaps how the experimenter knew they were there, and why the request was being made of them. Thus the alteration to the design of the task in Study 1 seems likely to have succeeded in rendering it more natural, and thus perhaps also less subject to intrusive effects. It might be noted that the greater prevalence of non-responses and actively incorrect responses than expressions of puzzlement is also consistent with both Robinson's (1981) and Bezuidenhout & Shroda's (1998) findings that young children do not typically ask for further clarification where a message is inadequate.

The other changes from Babelot & Marcos' (1999) design would seem also to have been broadly positive in impact. The decision to utilise only the most and least likely grammatical utterance forms to elicit correct responses appears to have served well in terms of discerning levels of pragmatic sensitivity, without evidently impeding performance. Similarly, the sub-categorisation of verbal responses into those which were indicative of understanding and those which were not seems to have served also to help clarify the nature of children's grasp of the requests being made, and the references these involved. Finally, the inclusion of a third, oldest group has allowed for the clearer identification of age trends in proficiency in pragmatic sensitivity as well as in grasp of non-immediate references.

To summarise the overall findings for Study 1, it would appear that whilst even the youngest children are showing signs of sensitivity to reference, there are clear improvements with age, with 3.5 to 4 year olds beginning to pick up references to the non-immediate context in fairly reliable fashion. These results imply that immediate references are understood at an earlier age than non-immediate references are, and this is consistent with Ninio & Snow's (1996) developmental progression in conversation that takes as its starting point references to the immediate context with understanding then progressing to references to non-immediate objects. However, one other point should be noted in this connection i.e. Ninio & Snow's (1996) work was based solely on the overt performance of children in a naturalistic and observational setting, whereas Study 1 has used experimental methods to assess the developmental progression of underlying competence. As this measure of competence has shown, the pattern of progression for children is one of gradual understanding rather than sudden insight, with that progression accompanied by growing pragmatic sensitivity, and affected by wider language competence, as well as age. Given that progression in understanding appears to be gradual, the question of the age at which children understand references to non-immediate objects seems inappropriate i.e. relative competence is the only meaningful index, with the development of competency appearing to be protracted across the age span measured.

In their research, Ninio & Snow (1996) make the claim that developmental progression in conversation proceeds from discussions relating to the first sub-maxim i.e. reference to non-immediate objects, to the second sub-maxim of references to shared past events. This being the case, Study 2 was designed to measure children's competence by comparing their understanding of references to present events and shared past events. As a result of the Study 1 findings, it will be necessary to explore several issues in Study 2. These include, in particular, the issue of whether young children continue to show signs of sensitivity to reference type and request form in this context, whether the pattern of change with age remains one of gradual progression rather than sudden insight, and how far understanding of references to shared past events is also affected by language level and working memory.

CHAPTER 3

1.5 to 4.0 year old children's understanding of references to shared-past events

3.1 Introduction

Study 1 revealed that children aged 18 months to 4 years showed increasing signs of pragmatic sensitivity to both the conversational demands implied by a request and variations in the form in which they were expressed, as understanding of reference to non-immediate objects improved. In particular, children aged 3.5 to 4 years showed consistent grasp of such references alongside clear differentiation of the form of request within which they were embedded, indicating that Ninio & Snow's (1996) first sub-maxim is fairly firmly understood by this point. Grasp of references to non-immediate objects was found to emerge gradually, however, rather than being a sudden conceptual acquirement on the part of children, and even children aged 18 months to 2 years showed some sporadic understanding.

According to Ninio & Snow's (1996) observational research, and the examination of the CHILDES Language Database (MacWhinney, 2001) presented in Chapter 1, discussions of shared past events seem to appear next in the sequence of non-obvious referents within the conversations that adults and children engage in. Ninio & Snow report that although discussions of shared past events do not commonly appear with children until 28 to 32 months of age, discussions of this type have been found among 27% of children aged 20 months, as against 51% of children aged 32 months. However, this research suggests that 20-month-old children are reliant upon, and typically only responding to, adult initiatives when discussing shared past events, or at least that this is the case to a much greater extent than with older children.

Ninio & Snow (1996) report that in general mothers do not in fact engage in such one-sided conversations or monologues with children on topics related to past events – or to other hypothetical or abstract matters that are largely beyond children's comprehension – prior to 28-32 months of age. The reason for the lack of adult engagement is, according to Ninio & Snow, that the cognitive and linguistic abilities crucial for initiating talk about past events i.e. control of past tense, future aspect and genericity markers are not yet in place with young children. Only very gradually do children become able to process utterances dealing with such topics, but as soon as they signal comprehension, mothers start including references to past shared events in their discussions. Ninio & Snow suggest that the increase in participation of children in decontextualised talk about past events at this point may in fact serve as a crucial context for the acquisition of the cognitive and linguistic abilities necessary to partake actively in such discussions.

In support of these findings, Bretherton & Beeghly (1982) and Bretherton (1991) found that by 28 months of age, children had acquired a rich enough vocabulary to discuss references to past and future anticipated inner states, including perceptions, sensations, physiological states, basic emotions (happy, sad, and scared) and moral approval/disapproval (good or bad judgements). Notably, in terms of behaviour that may seem relevant to their understanding of others as psychological beings, children of 28 months also made causal statements about internal states, indicating some awareness of temporal sequence. In the course of their third year, children made further use of this developing ability to discuss inner state causality to influence, persuade and cajole others.

Examining mother-child conversations in the home, however, Dunn (1994) found signs of this kind of ability at an earlier age. In this research, children aged 18-24 months labelled an even wider range of feeling states and made more causal statements regarding feelings to influence mothers than Bretherton (1991) had reported. The point at which children might begin to understand references to past events in connection to present experiences is therefore somewhat unclear. Work by Reese & Brown (2000) on event memory in preschool children provides support, though, for the notion that such understanding is fostered by the inclusion of these types of reference in adult-child conversations. They examined reminiscing and recounting of events between mothers and pre-school children, and found that the more memory information mothers provided during reminiscing, and the more they asked for during recounting, the more their children subsequently reported unique information about events themselves.

3.1.1 Rationale for Study 2

Despite work of the kind just outlined, little quantitative research has been undertaken in respect of references to shared past events per se in child-caregiver conversations, or with regard to children's understanding of such references. Indeed, Dunn (2006) has recently highlighted the need for more study of the conversational pragmatics of exchanges about shared past events, identifying the subject as one of four areas requiring extensive research i.e. "We know little about how often, in daily life, mothers and children engage in these conversations about past and future, about analogy, and about stressful conflicts. In which families and in which contexts do these important conversations take place?" (Dunn, 2006, p. 156).

To date, the research that has been conducted remains largely qualitative in nature (Ninio & Snow, 1996) or mainly confined to discourse analysis of adult to adult conversations i.e. where the focus is, as reported in Study 1, squarely upon the decoding of conversational implicatures and how references to past events may arise in the adult world (Grice, 1975; Sperber & Wilson, 2002). The qualitative research that has been applied to children's understanding of such references e.g. examination of the CHILDES Language Database (MacWhinney, 2001) and the observational-longitudinal research of Ninio & Snow (1996) suggests that once parents have established that children can process references to physically remote objects in a non-immediate context (as in Study 1), and show an understanding of tense markers, they then gradually introduce references to past shared events. These commonly occur during joint play activities with a child, where some aspect of the present context exhibits a connection with a shared past event, this connection becoming the subject of caregiver commentary, as in Example 4 in Section 1.8 above.

However, although the first signs of children's ability to discuss past events may be becoming evident from the middle of their second year, the *decoding* of references under this second sub-maxim may remain largely dependent for much longer upon a context-specific grasp of conversational practices, and of the ways in which such references typically arise (Ninio & Snow, 1996). As argued previously, children need to understand the different ways in which an utterance can be meant to be relevant, but the move towards discussions of shared past events may present considerable challenges, since the background knowledge on which this discourse rests is not as evident as it is in discussions concerning non-immediate objects, where referents are at least available to visual search of the environment. For this reason, discussions of shared past events may require substantial *historical* support from adults, in the sense of them drawing upon their knowledge of particularly salient past events, shared topics, likes and dislikes within a given setting (cf. Ninio & Snow, 1996). By doing this, parents may effectively illustrate the ways in which past shared events can be relevantly brought into current conversations. Extracting a more general understanding of the principles at work may take children a lengthy period of time, however. One implication of this is that in the absence of parental support, children will be slow to move from initial engagement with discussion of past events to full competence at determining referents, and may take longer than appeared to be the case for understanding of non-immediate referents.

However, given that children's understanding of references to shared past events remains largely uncharted (cf. Dunn 2006), such points are necessarily speculative. A systematic and controlled investigation of children's competencies with respect to the interpretation of such statements was therefore carried out for Study 2. The focus of the study was squarely on circumstances comparable to those under which references to past events might usually have to be interpreted, examining as it did children's understanding of an adult's request for an object jointly encountered in a past event similar to a present joint activity, but where this connection was left implicit, or assumed, as is commonly the case in everyday conversation. Following on as it did from Study 1, Study 2 also provided an opportunity to track the relative emergence of the first two sub-maxims of Ninio & Snow's (1996) developmental progression in conversation.

The basic task that children were asked to perform in Study 2 was similar in form to that used in Study 1, save that responses to requests for objects in the child's immediate field of view were contrasted with their response to requests about objects that had been encountered during a priming scenario presented the day before. In this primer, children viewed four soft toy animals who *liked* a particular toy food, as evidenced by an exemplar of it being physically attached to them. Time was taken to ensure that children learned the association of animals to these liked foods, since the request trials were couched in terms of providing the toy 'that likes (x food)'.

The following day, children then viewed four new toy animals, which *liked* four other (i.e. new) toy foods, and this represented the present event. Again, time was allowed for the children to learn the new animal to food associations. The four past event animals from the previous day's primer were then introduced *but without their* *associated foods*. A further four control toy animals, which had no *liked* foods associated with them were also included to test for random responding. A total of 12 soft toy animals were therefore in children's view, of which only the present toy animals had their four liked foods available, albeit located about 15 cm in front of them.

This arrangement was used because it was considered that removing present event toy foods out of children's view or away from their respective animals would mean that these toys would effectively become further past event items, or items in a test of recall. At the same time, detaching foods from present event toy animals and placing them a short distance to their front served to remove the possibility of children responding to requests for these items simply by scanning for the animal that had the food referred to attached to it. This made for a slightly more exacting test of children's understanding of references to the present event. It also made the form of the requests seem more natural, since it is not unusual for a toy and its accessories to be both separate and available in present joint play situations, where an adult might ask a child for the doll's dress or the driver of the toy tractor. This also ensured that the distinction between past and present events remained firmly in place i.e. children had to suspend their focus on the here and now experience of the present event toy animals, with their associated present event toy foods, to consider which past event toy animals liked an unavailable toy food – and indeed appreciate that the reference was to information encountered in the priming event.

In order to examine how far memory was a factor in performance, an overall memory test was also conducted after the request trials had been concluded (so as to avoid contaminating responses). The measure used was directly task related, and a natural part of the on-going activity, lending it greater ecological validity. Both past and present toy foods were jumbled up and the child asked to place each toy food with its appropriate soft toy animal. This made it possible to test children's performance on both present and past animal-food pairings, and the separate relationship of these to responses to requests for present and past event animals. Moreover, as the same children had participated in Study 1 the previous day, this also allowed examination of the relationship of these memory scores to children's grasp of references to non-immediate objects.

One particular consideration here was that the procedure adopted for the request task meant the toy foods associated with the present event animals were never absent from view, and that performance with respect to these items in the memory test could therefore be interpreted as a measure of working memory. There is evidence that working memory capacity in pre-school children, as measured by non-word repetition, digit span and non-word recognition tasks, constrains their ability to comprehend what they hear or read, and to learn new words (Gathercole, Willis, Baddeley & Emslie, 1994; Gathercole, Service, Hitch, Adams & Martin, 1999) There is therefore some likelihood that it also impacts on understanding of referents, especially those outside the immediate, present context, where the reference itself (and the apparent conversational objective of making it) have to be kept in mind whilst the referent is retrieved. Indeed, other research points the impact of working memory constraints on conversational references. For instance, Bezuidenhout & Shroda (1999) argue that inappropriate responses to ambiguous messages are the result of children's immature working memory (see also Carlson & Moses, 2002; Leslie, German & Pollizi, 2005; Mutter, Alcorn & Welsh, 2006).

Finally, as in Study 1, the requests within which references to the present or past event were made were varied between direct and indirect forms, suitably adapted for present purposes ("Give me the X that likes Y" and "Have you got the X that likes Y?"). The utility of the different request forms in Study 1 for discerning the emergence of children's pragmatic sensitivity alongside their grasp of non-immediate referents was considered suitable justification for retaining this element in Study 2.

All three age groups who participated in Study 1 were deemed eligible to participate in Study 2. Since Ninio & Snow's (1996) naturalistic-observational studies found that an understanding of references to shared past events was evident among a quarter of 20-month olds, albeit in a rudimentary form, the task was considered suitable for the youngest group of children in the present research, aged 1.5 to 2 years.

3.1.2 Key questions and hypotheses for Study 2

The key questions and corresponding hypotheses addressed by Study 2 were as follows:

(1) Is there an effect of context (in this case, past vs. present events) on children's understanding of referents, comparable to that observed in Study 1? On the basis of the data from Ninio & Snow (1996) regarding the emergence of conversations about past events, it was anticipated that children in all three age groups would show lower understanding of requests involving references to this context.

(2) Do children in the oldest age group pick up references to shared past events as reliably as they did references to non-immediate objects in Study 1? Again, on the basis of the data from Ninio & Snow (1996) about the ages at which children participate in conversations about past events, it was expected that 3.5 to 4 year olds would show a reasonable degree of competence in understanding references of this kind. At the same time, however, the fact that children in this age group were not at ceiling in terms of grasp of references to non-immediate objects in Study 1 made it unlikely that they would perform faultlessly here – as is implicit in Hypothesis 1. Indeed, Ninio & Snow's claims about the relative emergence of understanding of non-immediate and past event references indicated that the oldest children might be expected to perform somewhat worse with regard to the latter than they had for the former.

(3) Is the pattern of change across age groups in grasp of past event references still one of gradual progression rather than sudden insight? Since the underlying process promoting growth in understanding of past event references was held to be the same as that operating for non-immediate references (i.e. introduction of such references into conversation by parents), it was hypothesised that it would show the same

pattern of gradual change observed in Study 1. More precise predictions about the relative performance of the two younger age groups were harder to arrive at, but the incidence of use of past event references in conversations with 20 month olds reported by Ninio & Snow (1996) indicated that some grasp ought to be evident in the youngest age group. Performance of children in the middle age group was expected to be intermediate between this level and that exhibited by the oldest children.

(4) Would the effect of request form found in Study 1 (i.e. more action responses to direct requests and more verbal responses to indirect) also occur in Study 2? Given the convergence of Study 1 and the results reported by Babelot & Marcos (1999), it was anticipated that this pattern should also be present here.

(5) Would children exhibit the same pattern of emergence of pragmatic sensitivity alongside grasp of past event references as seen in Study 1 (i.e. initial awareness of the requirement to make a response, followed by differential responding to direct and indirect requests)? In view of the clear pattern of progression seen in Study 1, it was expected that this would be evident again in Study 2, and would again shed useful further light on children's grasp of past event references. Since children had not reached the point of decoding the implicature contained in indirect questions in Study 1, as would be evidenced by action responses to indirect requests, it was similarly not expected that they would have reached the level of such decoding here.

(6) Is grasp of references to shared past events affected by language and memory ability? Given the linear relationship between language ability and understanding of non-immediate references reported in Study 1, a similar pattern was expected to emerge here for grasp of past event references. External evidence on relationships between memory and conversational skill, as outlined above, led to a similar expectation that this element would be related not just to understanding of past event references, but also to the grasp of non-immediate references observed among the same sample in Study 1.

3.2 Methodology

3.2.1 Design

Study 2 employed an identical mixed between-and within-subjects design to that used in Study 1, as well as a request task of similar form, involving eight trials for each child participant. As before, these trials comprised two instances each of the combinations of two reference types (present vs. past event) with two request forms (direct vs. indirect). As in Study 1, children's responses were coded in terms of whether or not they were correct, and additionally whether they were action, verbal or non-responses. A further element of coding actively incorrect responses for which object set (present, past or control) was included to allow examination of degrees of inaccuracy. As Study 2 involved children who had participated in Study 1, data from tests of expressive language had already been collected. Data were also collected from a memory test, involving recall of associations established during preparations for the request task.

3.2.2. Participants

Participants comprised 75 (37 girls, 38 boys) of the 79 children who took part in Study 1 (four were unavailable on the day of testing). As before, the sample was divided into 3 age groups, with 22 children in the youngest group, 1.5 to 2 year olds (11 females, 11 males; mean age = 1 year, 10 months), 26 children in the middle group, 2.5 to 3 year olds (14 females, 12 males; mean = 2 years, 9 months), and 27 children in the oldest group, 3.5 to 4 year olds (12 females, 15 males; mean = 3 years, 8 months). Ethical clearance and permissions were as for Study 1.

3.2.3 Materials

The main materials consisted of a collection of soft toy animals and toy foods, of a form familiar to children within these age groups. The toy animals consisted of a *Bear, Cat, Crocodile, Dog, Duck, Elephant, Frog, Lion, Monkey, Mouse, Penguin and Rabbit.* No animal was smaller than 15 cm or higher than 18 cm. The toy food items were made of plastic and were of a standard type used by young children to represent food when playing toy *'shops'* or toy *'kitchens'* at nurseries or in the home. These items consisted of a *Burger, Cake, Chips, Chocolate, Egg, Orange, Sausage and Tomato Sauce.* No toy food item was smaller than 4 cm or larger than 6 cm. Each animal had a small patch of Velcro sewn onto it and each toy food item had an equivalent Velcro patch glued onto it, to enable the attaching of foods to animals. A large green cloth mat was also used to conduct the test sessions upon nursery floors. The experimenter scored children's responses during the test sessions using the pro-formas shown in appendices C & D. A video camera on a tripod was utilised to record children's responses, enabling the coding of responses to be re-checked by the experimenter and then verified by an independent judge at a later date.

3.2.4 Procedure

The eight participating nurseries were visited individually for two days in a row over a six month period. Study 1 always took place on the first day of each visit, followed immediately by the priming event for Study 2, and Study 2 proper always took place on the second day. Study 2 involved all children in the following four stages of activity.

Stage 1: The Past Event

Stage 1 was the priming past event, conducted individually with children the day before Study 2 proper. Each child was already present in the experimental room, with a nursery nurse still in attendance, having just completed Study 1. When all materials relating to Study 1 had been removed from the green mat, the child was asked if they would like to play a game that involved finding out *'Which foods the animals like'*. The child was then shown the soft toy animals and toy foods belonging to one of two sets, C or D (used in order to counterbalance these across past and present events). For Set C, the following four animal-food pairings were used: the Cat and the Chocolate, the Dog and the Cake, the Monkey and the Chips and the Rabbit and the Sausage. For Set D, the pairings were the Bear and the Egg, the Frog and the Burger, the Lion and the Tomato Sauce, and the Penguin and the Orange.

Each food was attached to its appropriate animal by means of the Velcro and placed on the green mat in front of the child about 15 cm apart. The child was then told that the animals *liked* the particular food that was attached to them, and that the game involved attaching and detaching each animal's *liked* food. The food items were then detached from their animal and jumbled up, and the child was asked to reattach them to the correct animal as a memory and comprehension check. As the purpose of the primer was to get children to associate each animal with its food, the game was played until the experimenter had established that each child could perform the reattachment without error. No child took longer than ten minutes to do this.

Stage 2: The Present Event

Stage 2 was always conducted the day following Stage 1 and consisted of a 'present event' of the same form as the past event. A nursery assistant accompanied the participating child into the experimental room and again remained with them during the rest of what followed. As children had participated in the Stage 1 primer the previous day, this helped the building of the rapport necessary for taking part in Stage 2. The child was asked if they wanted to play another game to find out *'Which foods some other animals like'*. With the child sat on the green mat, the four opposing animal-food pairings to those used in the previous day's primer were placed directly in front of them. To allow children to associate present event animals with their respective foods, a game identical to that used in Stage 1 was played. This ensured that procedures remained standard between past and present events.

Stage 3: Testing of understanding of Present and Past Event references

When the child had successfully grasped the present event animal-food pairings, these animals, with their foods detached but placed about 15 cm in front of them, were spread out across the green mat in a straight row in direct view of the child. Sufficient space was always left between each animal so that the past event and control animals could be interspersed between them. The experimenter then introduced the four past event toy animals that had been the subject of the child's past event primer in Stage 1. These animals did <u>not</u> have their associated foods available. A further four toy animals were also introduced, without any toy foods, to serve as control animals to test for random responding on the part of children. These toy animals were the *Crocodile, Duck, Elephant, and Mouse.* All 12 animals were evenly positioned about 5 cm apart on the green mat, with the precise configuration being varied systematically for each child. No two animals from the same present, past or control set were placed beside one another.

The video camera was now switched on, and the child was told the experimenter wished to ask them some questions about the toy animals and their foods. Their ability to resolve references to both present and past events was then tested by examining their responses to a series of eight requests made by the experimenter using two request forms, direct ("Give me the X that likes Y ", where X refers to a soft toy animal and Y refers to a toy food) and indirect ("Have you got the X that likes Y?"). The requests were put to children one after the other with no other dialogue, and spaced regularly apart with a 20 second interval between, as measured by the experimenter's watch (see Tables 3.1 and 3.2 for the sequence of requests, depending on which set of animal-food pairings was used for that child's past and present events). Children's responses were recorded at the time by the experimenter on either the Appendix C or D pro-forma, as appropriate, as well as on videotape. No child became tearful or upset at any point, necessitating termination of testing.

Table 3.1: Order of requests by object, reference type and request form for children who saw Set C as their past event primer

QUESTION- OBJECT	REFERENCE TYPE	REQUEST FORM
1. Give me the animal that likes cake (Dog)	Past	Direct
2. Have you got the animal that likes tomato sauce? (Lion)	Present	Indirect
3. Have you got the animal that likes chocolate? (Cat)	Past	Indirect
4. Give me the animal that likes burger (Frog)	Present	Direct
5. Give me the animal that likes chips (Monkey)	Past	Direct
6. Have you got the animal that likes sausage? (Rabbit)	Past	Indirect
7. Give me the animal that likes oranges (Penguin)	Present	Direct
8. Have you got the animal that likes egg? (Bear)	Present	Indirect

Table 3.2: Order of requests by object, reference type and request form for children who saw Set D as their past event primer

QUESTION- OBJECT	REFERENCE TYPE	REQUEST FORM
1. Have you got the animal that likes burger? (Frog)	Past	Indirect
2. Give me the animal that likes egg (Bear)	Past	Direct
3. Give me the animal that likes chocolate? (Cat)	Present	Direct
4. Have you got the animal that likes chips? (Monkey)	Present	Indirect
5. Have you got the animal that likes oranges? (Penguin)	Past	Indirect
6. Give me the animal that likes sausage? (Rabbit)	Present	Direct
7. Have you got the animal that likes cake? (Dog)	Present	Indirect
8. Give me the animal that that likes tomato sauce (Lion)	Past	Direct

Stage 4: Memory Testing

Once the request task had been completed, an overall test of memory for the animal-food pairings was conducted. The four past toy foods were introduced along with the four present toy foods, and jumbled up. The child was then asked to place each toy food with the appropriate animal from the 12 in view on the green mat in front of them. At the end of the placing of items children were given the opportunity to change their choice i.e. the experimenter asked children, "Are you happy that each animal has got the food it likes". No child changed their choice of foods to animals at this stage. Once the task was completed the child was given a happy face sticker and thanked for his or her participation.

3.2.5 Coding of Responses and Scoring

Request task

The experimenter observed and recorded children's responses during testing and these responses were also later rechecked and transcribed from videotape after data collection was complete. Children's responses to each request were coded using the same scheme as in Study 1, save that incorrect action responses were subdivided into three types, according to whether the mistaken object formed part of the present event, past event or control set. To check on the reliability of coding, the responses of five children from each age group were selected at random for coding from the video footage by an independent judge. The 120 responses given by these children accounted for 19% of the total. Agreement rates for each code are given below.

1. Correct action response = 97%
2. Correct verbal/gestural response = 97%
3. Incorrect action response (gives present object) = 100%
4. Incorrect action response (gives past object) = 100%
5. Incorrect action response (gives control object) = 100%
6. Incorrect/uninformative verbal response = 100%
7. Incomprehension response = 100%
8. Non-response = 93%

Scores from 0 to 2 were then computed for each coding category for each of the four reference type/request form combinations, as in Study 1. Similarly, to check on order effects, scores from 0 to 2 were also computed for each code for the first and second two present event requests, and for the first and second two past event requests (i.e. one of each request form in each case).

3.2.6 Language test scoring

Language tests for the participating children had already been scored as part of Study 1 (see Chapter 2, sub-section 2.2.6 for details).

3.2.7 Memory check scoring

Children were given one point for each correct assignment of a food to an animal, with a separate tally being kept for present and past event items, as well as an overall total.

3.3 Results Section

3.3.1 Summary of Analyses

Preliminary analyses were conducted to check for order effects in the request task data, and also whether there was any effect of the toy set used for the past and present events. A 3-way mixed ANOVA (age group x reference type x order) was performed on the order scores outlined in 3.2.5 for each of the eight response codes. No significant effects or interactions involving order were found. Similarly, a 4-way mixed ANOVA (age group x toy set assignment x reference type x request form) was conducted on the main request task scores for each of the eight response codes. Again, there were no significant effects or interactions found in relation to the set children saw for past and present events. These factors are therefore not considered further.

As regards the main analyses, 3-way mixed ANOVAs (age group x reference type x request form) were computed on scores for each of the eight response types for present/direct, present/indirect, past/direct and past/indirect requests. As in Study 1, a similar 2-way analysis was also conducted on the overall level of correct responses, collapsed across action and verbal response types. Finally, a 3-way analysis was used in order to examine the incidence of active vs. inactive incorrect responses across age groups and reference types.

Once the main patterns of performance on the request task had been identified, the relationship of correct responses to language and memory scores was examined using correlational analyses. As part of these analyses, partial correlations were used in an attempt to identify the unique contribution to performance of different sources of variance, including children's age. Finally, regression analyses were used to confirm the picture that emerged from the examination of correlations. Outcomes from these analyses are reported below in this order.

3.3.2 Descriptive and Inferential Results for Request Task.

Table 3.3 shows the mean scores and standard deviations for the five most prevalent responses given by children, broken down by age group, reference type and request form. These are examined in more detail below. Full data are available in Appendix L, which also presents results for the 3-way ANOVAs conducted on these data.

TABLE 3.3: Means and standard deviations for each response code (range = 0-2) by reference type, age group and request form

			Correct action resp		Correct verbal resp		Incorrect action resp		Incorrect verbal resp		Non resp	
Ref type	Age group		Dir	Indir	Dir	Indir	Dir	Indir	Dir	Indir	Dir	Indir
Present	Young	Mean	0.18	0.27	0.22	0.04	0.09	0.09	0.13	0.18	1.36	1.36
		SD	*0.58*	*0.63*	*0.52*	*0.21*	*0.29*	*0.29*	*0.46*	*0.58*	*0.90*	*0.90*
	Middle	Mean	0.53	0.65	0.46	0.42	0.07	0.03	0.26	0.26	0.42	0.42
		SD	*0.81*	*0.89*	*0.81*	*0.80*	*0.27*	*0.19*	*0.66*	*0.66*	*0.80*	*0.80*
	Oldest	Mean	1.44	1.33	0.18	0.29	0.00	0.00	0.14	0.18	0.07	0.11
		SD	*0.89*	*0.83*	*0.55*	*0.60*	*0.00*	*0.00*	*0.53*	*0.48*	*0.42*	*0.42*
	Total	*Mean*	0.76	0.78	0.29	0.26	0.05	0.04	0.18	0.21	0.57	0.58
		SD	*0.94*	*0.90*	*0.65*	*0.62*	*0.22*	*0.19*	*0.56*	*0.57*	*0.88*	*0.88*
Past	Young	Mean	0.04	0.04	0.04	0.13	0.04	0.04	0.09	0.13	1.59	1.63
		SD	*0.21*	*0.21*	*0.21*	*0.46*	*0.21*	*0.21*	*0.35*	*0.61*	*0.73*	*0.65*
	Middle	Mean	0.42	0.38	0.15	0.23	0.07	0.11	0.23	0.23	0.96	0.80
		SD	*0.75*	*0.75*	*0.58*	*0.70*	*0.27*	*0.32*	*0.51*	*0.58*	*0.91*	*0.98*
	Oldest	Mean	0.77	0.74	0.44	0.59	0.03	0.18	0.22	0.11	0.29	0.25
		SD	*0.84*	*0.94*	*0.80*	*0.93*	*0.19*	*0.48*	*0.64*	*0.42*	*0.66*	*0.65*
	Total	*Mean*	0.44	0.41	0.22	0.33	0.05	0.12	0.18	0.16	0.90	0.85
		SD	*0.73*	*0.77*	*0.58*	*0.72*	*0.22*	*0.36*	*0.51*	*0.49*	*0.93*	*0.95*

Ref type = Reference type (Present or past); resp = response type; Dir = direct response form and Indir = indirect response form. The following response categories are excluded from this table due to their infrequency of use by children: Incorrect action responses (gives past object) represented 3.3% of total responses; Incorrect action responses (gives control object) represented 3% of total responses; and Incomprehension responses represented 1.5% of total responses.

1. Correct action responses

Considering main effects initially, as is apparent from Table 3.3, correct action responses were given by the oldest group most often, followed by the middle group, with the youngest group giving the fewest. This pattern was supported by a significant main effect of age ($F(2,72) = 16.03$, $p < .001$). Tukey (HSD) post-hoc tests showed that these differences were reliable between the oldest and the youngest groups ($p < .001$), between the middle and the youngest groups ($p < .001$), and between the oldest and the middle groups ($p < .05$). There was also a main effect of reference type ($F(1,72) = 18.11$, $p < .001$), reflecting the fact that, as predicted, children in all 3 age groups gave more correct action responses for present event references than for past.

A weakly significant interaction of reference type x age ($F(2,72) = 3.67$, $p < .05$) was also found, which was attributable to the gap between reference types increasing with age, as the performance of the oldest children on present event trials improved more substantially over the two younger age groups than it did for past event trials. The relatively meagre incidence of correct action responses for present event requests among the youngest and middle age groups compared to their performance in Study 1 on immediate reference trials (see Table 2.3) suggests that the embedded nature of the requests used here (i.e. via reference to an associated food) made the task intrinsically somewhat more difficult. Only the oldest age group showed roughly comparable performance across immediate and present event references, hence the apparently sharp improvement. Interestingly, these children also actually gave slightly more correct action responses for past event trials than they had for non-immediate request trials in Study 1, whereas the two younger age groups continued to perform at a poorer level on past event trials.

There were no significant main effects or interactions found for request form, and as can be seen in Table 3.3, the differentiation between request forms noted in Study 1, especially among the oldest children, was largely absent here.

2. Correct verbal/gestural responses

Although there was a trend towards increased numbers of correct verbal/gestural responses with age, especially for past event requests, there was no main effect of age group, nor indeed of reference type or request form, in contrast to Study 1. A weakly significant interaction for reference type x age ($F (2,72) = 4.19$, $p < .05$) reflected the growth of correct verbal responses amongst the middle age group on present event trials, but subsequent decline among the oldest group, and contrasting growth for this age group in use of these responses on past event trials. In neither age group did correct verbal responses outstrip correct action responses on any type of trial, however, in contrast to the trend towards these dominating on non-immediate trials in Study 1. This is perhaps a reflection of the fact that the past event objects were not physically remote here, and certainly, correct verbal responses were generally less common than in Study 1.

3. Overall correct responses

The absence of effects of request form for either correct action responses or correct verbal responses indicated a general lack of this level of pragmatic sensitivity across all three age groups, in contrast to Study 1. One consequence of this was that overall patterns of correct responding were more readily discernible here. Nevertheless, analysis of total correct response was necessary to examine the extent to which children understood present event references better than past, and how this changed with age. Data was also required in this form for analysis of language (3.3.3) and memory effects (3.3.4).

The relevant data are shown in Table 3.4, which confirms that the oldest children gave the most correct responses followed by the middle group, and that the youngest group performed relatively poorly in comparison to Study 1, even on present event trials. A 2-way mixed ANOVA (age group x reference type) identified a strong main effect of age ($F (2,72) = 23.56$, $p < .001$). Tukey (HSD) post-hoc tests indicated that there were reliable differences between all three age groups ($p < .001$). Table 3.4 also shows that more correct responses were given by children in all three age groups on the present event trials, and this effect was significant ($F (1,72) = 18.85$, $p < .001$). In general, then, children understood better what was meant by present event references than by past, though both improved with age. The oldest children gave a high number of correct responses on present event trials, though not quite to the same levels as on immediate reference trials in Study 1 (see Table 2.4), perhaps reflecting differences in task difficulty. Their performance on past event trials fell somewhat short of that on present event trials, with just over half of their responses being correct. The level here was more comparable to that found for non-immediate references in Study 1, though still at a slightly lower level, as had been anticipated.

Table 3.4: Total and mean correct responses by age group for present and past references

Age Group	Total Correct Responses Present Event References	Mean Correct Responses Present Event References	*St. Dev*	Total Correct Responses Past Event References	Mean Correct Responses Past Event References	*St. Dev*
Youngest	16	.72	*1.27*	6	.27	*.88*
Middle	54	2.07	*1.80*	31	1.19	*1.54*
Oldest	88	3.25	*1.22*	69	2.55	*1.39*
Total	158	2.10	*1.77*	106	1.33	*1.61*

In total there were 264 correct response types given out of a possible 600. Correct responses made up 44% of overall responses.

No significant interaction was found between age and reference type, in contrast to those found for correct action responses and correct verbal responses, confirming that these reflected variations in response form rather than in broader understanding. As in Study 1, the lack of interaction effect is consistent with the anticipated gradual growth in understanding of references to past events, rather than a sudden increase in insight. Figure 3.1 gives an indication of these trends and the age-related progression for references to both present and past events is clear.

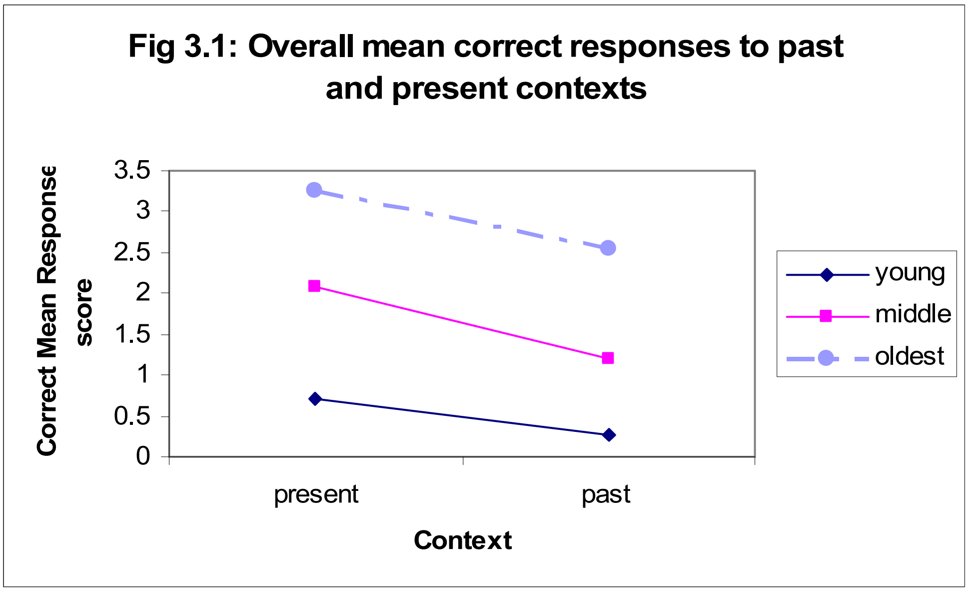


Fig 3.1: Overall mean correct responses to past and present contexts

4. Incorrect action responses (gives present object)

Incorrect action responses were generally infrequent, and as Table 3.3 shows, even the most common of these, giving a mistaken present object, was comparatively unusual. ANOVA identified no main effects in the incidence of this sub-type, not even in terms of reference type, indicating that when children gave inaccurate action responses, they tended not to discriminate between object sets – if they had, then these responses would have occurred more frequently on present event trials. There was a trend towards a reference type x age interaction, but this reflected the fact that the oldest group gave a mistaken present object more often in a <u>past</u> context, where the other two groups did not differentiate; and in any case, this did not meet conventional significance levels ($F(2,72) = 2.98$, $p = .057$). Due to their infrequency of use the two remaining sub-types of incorrect action responses were not subjected to statistical analysis.

5. Incorrect/uninformative verbal responses

The use of incorrect verbal responses by all three age groups was more frequent, but still comparatively modest, as can be seen from Table 3.3. There were no significant main effects or interactions involving this response code, suggesting the incidence of such responses was essentially randomly distributed across occasions when children could not identify a correct response.

6. Incomprehension responses

As in Study 1 (and again contrary to Babelot & Marcos, 1999), incomprehension responses were the least frequent category. Due to their infrequency of use, they were not subjected to statistical analysis.

7. Non-responses

As Table 3.3 shows, non-responses were the most prevalent category of response to the experimenter's requests amongst the youngest children, but even for the middle age groups this was more common than all the other incorrect responses put together. Only amongst the oldest age group did other forms of wrong response occur with comparable frequency, and even then, for past event trials, non-responses were the most common type of incorrect response. A strongly significant main effect of age confirmed the decline in non-responses as children became older ($F(1,72) = 25.50$, $p < .001$). Tukey HSD post hoc tests found these differences to be reliable between all three age groups. A significant main effect of reference type was also found ($F(1,72) = 12.96$, $p < .001$), with more non-responses occurring in the past event trials than in the present, confirming that for the younger two age groups at least, non-responses were effectively the default when (as was commonly the case) they did not know the answer to past event requests.

The relative prevalence of non-responses under these circumstances indicates that, in terms of pragmatic sensitivity, the majority of children were apparently not yet at the level of being aware that a response was required to past event requests, even if they did not know the answer. This pattern stands in contrast to that evident in Study 1, where this point at least was understood by the middle age group. Here, it was only the oldest age group who showed some appreciation of this requirement.

8. Analysis of active vs. inactive incorrect responses

In order to confirm the general pattern apparent in Table 3.3, an analysis of active vs. inactive incorrect responses was undertaken. Active incorrect responses combined incorrect action and verbal responses across request form, and inactive incorrect responses did the same for non-responses. The pattern of performance is shown in Table 3.5.

Table 3.5 : Active vs. inactive incorrect responses by age group and reference type

Age Group	Total Inactive Incorrect Present Event Responses	Mean Inactive Incorrect Present Event Responses	*St. Dev*	Total Inactive Incorrect Past Event Responses	Mean Inactive Incorrect Past Event Responses	*St. Dev*
Young	60	2.72	*1.77*	71	3.22	*1.26*
Middle	22	0.84	*1.59*	46	1.76	*1.83*
Oldest	5	0.18	*0.78*	15	0.55	*1.31*
Total	87	1.16	*1.75*	132	1.76	*1.83*
Age Group	**Total Active Incorrect Present Event Responses**	**Mean Active Incorrect Present Event Responses**	***St. Dev***	**Total Active Incorrect Past Event Responses**	**Mean Active Incorrect Past Event Responses**	***St. Dev.***
Young	11	0.50	*1.14*	15	0.31	*0.83*
Middle	17	0.65	*1.32*	17	0.65	*0.97*
Oldest	9	0.33	*0.96*	15	0.55	*0.97*
Total	37	0.49	*1.14*	47	0.52	*0.93*

In total there were 303 Active Incorrect and Inactive incorrect types given out of a possible 600. Total active incorrect and non-response types made up 51% of overall responses.

A 3-way ANOVA (age group x context x active vs. inactive responses) conducted on these data found significant main effects of age ($F\ (2,72) = 24.3$, $p < .001$) and reference type ($F\ (1,72) = 14.5$, $p < .001$), as would be expected (see Appendix M for full details). A strongly significant main effect was also recorded for active vs. inactive responses ($F\ (1,72) = 23.7$, $p < .001$), and an interaction between this and age group ($F\ (1,72) = 12.7$, $p < .001$), reflecting the relative overall dominance of inactive responses, but the shift in this pattern with increasing age, and the fact that the oldest group appeared to understand that some form of response was required, even when they were not sure what the actual referent was. There was also an weak interaction between reference type and incorrect response type ($F(1,72) = 6.87$, $p < .05$), with the balance between inactive and active incorrect responses being more skewed in the past event trials.

3.3.3 Language and Correct Response Scores

The relationship between the language tests used for the youngest and for the two oldest age groups respectively and their correct response scores were examined by computing Pearson correlations. No significant correlation was found between the youngest children's performance on the FWFS language test and number of correct responses ($r = .16$, $n = 22$, $p > .001$, 1-tailed). However, BAS II NVT scores and correct responses were correlated ($r = .35$ $n = 53$, $p < .001$, 1-tailed) among the middle and oldest age groups. The lack of correlation among the youngest children may reflect their apparent difficulty with the Study 2 request task, and resulting attenuation in variance.

3.3.4 Memory Tests.

Performance on the memory test

The question of whether memory was a factor in children's performance was examined in terms of their performance on both present and past event object sets.

Children's assignment of the correct food to the correct animal in the two sets is shown in Tables 3.6 and 3.7 respectively, broken down by age group.

Table 3.6 Present foods given to Present animals

Present foods to Present animals	N	Mean	Std. Dev.
Youngest Group	22	1.18	1.59
Middle Group	26	2.00	1.89
Oldest Group	27	3.40	.97

Table 3.7 Past foods given to Past animals

Past foods to Past animals	N	Mean	Std. Dev.
Youngest Group	22	0.81	1.13
Middle Group	26	1.15	1.54
Oldest Group	27	2.59	1.47

As can be seen, memory for animal-food pairings was generally more accurate for toys in the present event set, but accurate responses also increased with age, both patterns paralleling performance on the request task. A 2-way mixed ANOVA (age group x object set) conducted on these data found a significant main effect of age ($F(2,72) = 14.74$ $p < .001$). Tukey (HSD) post-hoc tests showed that these differences were reliable between the youngest and oldest group ($p < .001$) and between the middle and oldest group ($p < .05$). There was also a strongly significant main effect of object set ($F(1,72) = 23.79$, $p < .001$).

The relationship of memory scores to performance on the request task

The relationship between children's memory scores and their correct responses on the request task were examined more directly using Pearson correlations with respect to a) overall correct request responses, and b) correct responses to present and past event requests separately. Overall memory score and total correct response scores were strongly and significantly correlated ($r = .62$, $n = 75$, $p < .001$, 1-tailed). A slightly weaker correlation was found for present event requests ($r = .55$, $n = 75$, $p < .001$, 1-tailed), whilst the relationship for past event requests was between the two ($r = .59$, $n = 75$, $p < .001$, 1-tailed).

Request task performance controlling for age and memory

It is evident that there were strong correlations between children's memory scores and the accuracy of their performance on the Study 2 request task, both overall and in terms of responses to requests for present and past event objects taken separately. This raises the possibility that request task performance was in fact driven solely by children's ability to recall which food went with which animal, rather than by any conceptual difficulty with identifying that a reference to a past event had been made, and then tracking the referent in question. Poor recall would certainly constrain their ability to respond to any request, and the pattern of outcomes is consistent with this possibility, given that memory for past events was generally poorer, and that younger children had greater difficulty remembering pairings than older.

Against this, however, was the fact that the relationship between memory score and request task performance was far from perfect, with at best 38% of variance shared between the two. The implication is that there were other influences on performance beyond simple recall. In order to ascertain whether this might be the case, partial correlation coefficients were computed for the relationships between age (in months) and overall correct responses, present event responses and past event responses while controlling for memory score. If significant relationships between age and performance were found once the influence of memory was removed, this would suggest that children's performance on the request task was affected by some conceptual element as well as by recall. This turned out to be the case. There were medium-sized partial correlations between age and overall correct responses ($r = .45$, $n = 72$, $p < .001$, 1-tailed), present event correct responses ($r = .42$, $n = 72$, $p < .001$, 1-tailed), and past event correct responses ($r = .38$, $n = 72$, $p < .001$, 1-tailed). Inspection of the zero-order correlations ($r = .65$, .59 and .58 respectively) suggests that controlling for memory score had a definite, but relatively modest effect on the strength of the relationship between age and performance, with no more than about a 4% overlap between age and memory in the variance they explained in request task performance.

Examination of the obverse relationships between memory and correct responses while controlling for age revealed a similar picture for overall correct responses ($r = .42$, $n = 72$, $p < .001$, 1-tailed), present event correct responses ($r = .33$, $n = 72$, $p < .001$, 1-tailed), and past event correct responses ($r = .40$, $n = 72$, $p < .001$, 1-tailed). Again, comparison of these values to the zero-order correlations ($r = .62$, .54 and .59) suggested an overlap in explained variance of about 4%. The data confirmed therefore that whilst memory had an effect on correct responses to requests for both present and past event toys, there was a distinct effect of age over and above this.

The relationship of memory scores to Study 1 performance

Having determined that memory score was only a partial influence on request task performance, the other issue to be resolved was how the nature of this influence ought to be conceptualised. The most obvious interpretation was that it was simply a measure of specific recall. However, if memory score in Study 2 were found to be also related to performance in Study 1, where the request task was similar in form but required no element of long-term recall, this would suggest that in fact the Study 2 memory score indexed some wider competence, most obviously the management of information in working memory. The correlation between memory score and overall correct responses in Study 1 pointed strongly to the latter conclusion ($r = .48$, $n = 75$, $p < .001$, 1-tailed).

The relationship of Study 1 to Study 2 performance

If Study 1 and Study 2 performance were both affected by working memory, as the data seemed to indicate, then it would be expected that there should be some overlap in variance between the two. Computation of the Pearson correlation between overall correct responses in the two studies confirmed this ($r = .43$, $n = 75$, $p < .001$, 1-tailed). Partial correlations were also used to explore the relationship between Study 1 and Study 2 performance while controlling for memory score and age. If the relationship dropped to zero when memory score was controlled for, this would indicate that working memory was the only common influence between the two forms of the request task. On the other hand, the existence of an influence of age once the effect of memory was partialled out would indicate that Study 1 and Study 2 performance shared a conceptual dimension. The relative size of the partial correlations would give an estimate of the comparative scale of these influences.

In the event, a weak partial correlation was evident between Study 1 and Study 2 correct responses when controlling for memory score ($r = .20$, $n = 72$, $p < .05$, 1-tailed). A similar size of relationship was found when controlling for age ($r = .24$, $n = 72$, $p < .05$, 1-tailed). These results suggest that performance on both Study 1 and Study 2 was influenced by a working memory component and another distinct age-related component of similar effect size, though the total amount of shared variance accounted for by these was not especially large, and was marginally weaker in the case of the age component.

3.3.5 Regression Analyses

The partial correlations indicated that performance in Study 1 and Study 2 shared variance accounted for by working memory and by some age-related component that might reflect conceptual grasp of different forms of reference. However, they also suggested that there might be some more specific source of shared variance between performance on the two forms of request task, since the overall relationship between the two ($r = .43$, $r^2 = .18$) appeared to be only partially accounted for by memory and age, each reducing shared variance by about 5%. To further explore the relationship between Study 1 and Study 2, therefore, a forced entry regression analysis was conducted, with Study 2 overall correct score as the dependent variable and memory score, age in months and overall correct score in Study 1 as predictor variables.

This analysis produced a significant model, as expected, accounting for 50% of the variance in Study 2 scores ($F(3,71) = 25.64$, $p < .001$, adjusted R-square = .50). However, memory score (beta = .35, $p < .001$) and age (beta = .41, $p < .001$) were the only significant predictors to emerge.

Similar analyses were also conducted taking a) present event correct score as the dependent variable, and memory score, age and correct immediate score in Study 1 as predictors; and b) past event correct score as the dependent variable, and memory score, age and correct non-immediate score in Study 1 as predictors. Both analyses produced comparable outcomes to the overall analysis. For present event responses, explained variance was 41% ($F(3,71) = 18.29$, $p < .001$, adjusted R-square = .41), with memory score (beta = .27, $p < .05$) and age (beta = .39, $p < .001$) as significant predictors. For past event responses, explained variance was 43% ($F(3,71) = 19.39$, $p < .001$, adjusted R-square = .43), with significant predictors being again memory score (beta = .37, $p < .001$) and age (beta = .35, $p < .001$).

Taken together then, the correlational and regression analyses produced a consistent picture of Study 2 performance being predicted by a memory component and an age-related component, which were also influences on Study 1 performance. There appeared to be no other source of overlap between Study 1 and Study 2.

3.4 Discussion

Six predictions were made at the outset of Study 2: 1) that children in all three age groups would perform worse with respect to identifying the referents of requests for past event objects; 2) that the oldest children would nevertheless show relatively competent levels of performance under these circumstances, though not to extent that they had with references to non-immediate objects in Study 1; 3) that the evidence would favour gradual change in understanding with age rather than sudden shifts in ability; 4) that request form would affect the precise nature of children's responses, with direct requests leading to more action responses, and indirect requests to more verbal responses; 5) that children would exhibit the same sequence of emergence of pragmatic sensitivity as seen in Study 1 (i.e. recognition of the need to respond, followed by differential responding to the two forms of request), and that this would again shed light on their level of understanding of past event references; and 6) that performance would be influenced by both language and memory, with the influence of the latter extending to Study 1 performance too.

The level of correspondence between these predictions and the observed outcomes was in general very high. Children in all three age groups did indeed perform less well in responding appropriately to requests for toys that were associated with the past event than for those that were part of the present event. It must be acknowledged that there were clear signs that performance was affected to some extent by the nature of the task employed here, in particular the embedded form of the requests, which asked for animals via their associated foods. This was evident from the tendency of all age groups, but especially the youngest and middle, to do less well on retrieving objects in response to present event requests than they had to requests for immediate objects in Study 1, despite the essential similarity of the demands made by the two forms of request in terms of focus of attention on objects immediately in view (and marked out by the presence of their associated food). This style of request was considered necessary in terms of the study design, since it enabled objects to be in present view (i.e. avoiding a confound with immediacy), but only identifiable via retrieval of information either from the present context or from the past event.

In any case, the basic point remained that children at all ages found responding to past event requests to be harder than to present event ones, and there was no indication that task difficulty interacted with the difference between the two reference types. This was evident from the pattern of responses exhibited by the oldest children, which was below the level they showed in Study 1 for both immediate and non-immediate references, but to roughly comparable degrees. If there is a problem presented by the variation in task difficulty, it is that it makes it somewhat harder to discern whether children's grasp of references to shared past events is behind that of their grasp of non-immediate references, as Ninio & Snow (1996) claim ought to be the case on the basis of the sequence with which these types of reference emerge in conversation with parents. On the face of it, once task difficulty is allowed for, it would appear that 3.5 to 4 year old children's understanding of past event references was at roughly equivalent levels to understanding of non-immediate references, rather than slightly below, as had been anticipated. In contrast, though, the youngest and middle age groups' performance with respect to past event references showed a rather more sizeable decrement relative to non-immediate references in Study 1, indicating an effect of reference type beyond that of task difficulty.

The data are consistent therefore with the emergence of grasp of references to past events a little later than understanding of non-immediate references, with the former showing somewhat stronger growth after three years, roughly in line with the timing of the increased use of past event references reported by Ninio & Snow (1996). Even then, though, there is little sign of any very rapid increase in understanding at that point: instead, as in Study 1, correct responses showed a relatively gradual improvement with age – as the subsequently reported significant <u>linear</u> relationship between age in months and correct past event responses underlines. As with grasp of non-immediate references, this is much the pattern of change that might be expected if the primary influence on development were the cumulative effects of exposure to parental usage of such references over time, and their support in various ways for children's decoding of these.

If the arguments presented previously about pragmatic sensitivity as a further marker of understanding are taken seriously, then the evidence relating to this dimension of performance serves to confirm that grasp of past event references emerges later than that of non-immediate references, and indeed indicates that even the oldest age group here were still struggling somewhat with the former. The key pointers here are the lack of differential responding to direct and indirect forms of request, even among the oldest age group, in contrast to Study 1; and the much greater prevalence of non-responses, especially amongst the middle age group, but also to an extent amongst the oldest. The overall lower incidence in Study 2 of verbal responses might reflect in part the fact that all retrievable objects were immediately in front of the child, facilitating action responses relative to the conditions that obtained for non-immediate objects in Study 1, where retrieval meant physically getting up and fetching the requested toy.

However, this does not explain the lack of differentiation of response to direct and indirect requests, which suggests that either the difference in form passed unnoticed because of the demands of decoding the nature of the reference, or because the implicature contained in indirect requests was understood. The latter seems unlikely, at least among the younger age groups, given the apparent failure on 50% or more of past event trials to even recognise that a response was required at all. Even among the oldest children, non-responses were as prevalent as other forms of incorrect response, in contrast to their near-total absence in Study 1.

On balance, then, the pragmatic data indicate that none of the children had reached the level of noting the form of request alongside past event references, and only the oldest were at the level of recognising with any frequency the need to respond even when the answer was unknown. Some of the lack of differentiation of request form might be explicable in terms of the demands of the embedded nature of the requests, given that it was not evident on present event trials as well as on past. However, the frequency of non-responses to past event trials among the oldest age group, where performance was relatively competent, suggests that past event references taxed children's resources to a greater extent than the task itself did. The net conclusion then is that such references were indeed less firmly grasped among 3.5 to 4 year olds than non-immediate references. One further point to note here is that it was the self-same children who exhibited relatively high levels of pragmatic sensitivity in Study 1 who were showing a relative absence of it in Study 2. The implication is that even amongst rising fours, these pragmatic competencies show at best limited generalisation – consistent with the notion that for past event references at least, competence is largely context-specific, and perhaps dependent on conversational support much of the time, as argued earlier.

That at least part of the ability to decode references to objects or connections outside of the immediate context depends on management of attentional resources is indicated by the relationship of performance to children's memory scores. As noted in the Results section, a simple interpretation of these scores in terms of the impact of ability to recall the content of associations is ruled out by their predictive value for performance on Study 1, where such content was of no import. In the light of this, the most plausible interpretation is that they index working memory, and more particularly, the ability to hold a reference in mind whilst tracking what its referent might be, either visually or via a search of salient past events. Working memory would of course affect initial encoding of e.g. animal-food associations and their retention in longer-term memory, hence its indexing by the memory task used here. Interestingly, the slightly weaker association between memory score and Study 1 correct responses (.48) than that found in Study 2 (.62) provides some estimate of the impact of specific memory content over working memory processes per se, and this indicates that the bulk of the variance in memory scores was attributable to generic processes rather than specific content.

The Study 1 and Study 2 results both indicate that working memory affects children's ability to track conversational references, therefore, and to a roughly comparable extent regardless of reference type. This outcome is consistent with the general influence of working memory within conversations argued for by Bezuidenhout & Shroda (1998); and with evidence from Gathercole et al. (1994; 1999) that the working memory capacity of pre-school children (especially the phonological loop) limits their ability to comprehend an array of cognitive tasks that includes listening, reading, learning new words, recognition and recall-based tests. However, it was also evident that some age-related element beyond working memory ability influenced performance, and moreover that this formed part of the variance shared between Study 1 and Study 2 performances along with working memory. The nature of this age component is harder to pin down, though there is an obvious implication that it reflects some form of change in understanding or conception, and one which is not task-specific.

One evident possibility is some form of facility with language. Unfortunately, in the present studies the relationship between age and language was less amenable to the form of analysis used to investigate the relationship between age and memory score, because of the necessity of using different language tests across the three age groups. It is therefore unclear how much the age component in Study 1 and Study 2 might reflect language, though certainly language scores were age-related. Even if the age component were associated with language scores, however, this would only provide a general pointer as to its nature and the manner of its influence on request task performance: expressive language measures in this context could only themselves stand as a form of proxy for some more specific process at work in such tasks, since expressive language per se was only minimally required. Whilst the point is hard to establish with any finality, then, it would seem most likely that the age component reflected some changing ability to engage with the tasks themselves. Conceptual grasp of the nature of different forms of reference is the only element that is both sufficiently task-specific, and sufficiently general to be a common influence across Studies 1 and 2. Its age-related character would of course be consistent with the cumulative influence of parental input and support argued for in this thesis.

CHAPTER 4

2.5 to 4.0 year old children's understanding of references to a non-evident emotion state

4.1 Introduction

Study 2 found that older children (3.5 to 4 year olds) showed some understanding of references to past events. However, their apparent lack of differentiation of request form, and even partial lack of understanding of the conversational requirement to make a response was taken to indicate that their grasp of such references was less secure than their understanding of references to non-immediate contexts. The performance of the middle and youngest age groups also suggested much less grasp of references to past events. Overall, then, the data were consistent with the prediction derived from Ninio & Snow's (1996) observational work that competence with past events references emerges after that relating to references to non-immediate objects or events.

In addition to this basic pattern of performance, Study 2 also found that improvement with age in understanding of past event references was gradual, as Study 1 had indicated was the case with non-immediate references. Moreover, in both cases, observed improvements were predicted by a) performance on a task that was interpretable as an index of working memory; and b) another distinct age- and possibly language-related element that was suggestive of a conceptual component.

Whilst other interpretations may be possible, the data are certainly consistent with the developing grasp of complex conversational references having its origins in parental sensitivity to children's competence to engage with such forms (cf. the working memory element), and consequent provision of increased levels of exposure to these, along with support for decoding of referents (cf. Ninio & Snow, 1996).

Given this initial level of support for Ninio & Snow's (1996) account, Study 3 moved on to examine competencies regarding their third sub-maxim, i.e. children's understanding of references to internal, non-evident emotion states.

4.1.1 Children's understanding of emotion

Since understanding of references to emotions, especially where these are non-evident, entails at least some grasp of the nature of emotions as phenomena, it is necessary to commence by considering what understanding of emotion young children do possess. Considerable evidence has shown in fact that more-or-less from birth infants possess an innate capacity for emotion expression, and that they are responsive to and can discriminate between the basic human emotions of happiness, sadness and anger, as revealed in the facial expressions of others (e.g. Havilland & Lelwica, 1987; Izard, Heubner, Risser, McGinnes & Dougherty, 1980). Woodhead, Barnes, Miell & Oates (1995) argue that these innate capacities for expression and recognition become the building blocks for subsequent communication and learning. In particular, through *social referencing*, young children's emotional behaviour is influenced and regulated by the reactions of their caregivers. One classic example of this is Sorce, Emde, Campos & Klinnert's (1985) visual cliff experiment, where a glass surface was placed over an obvious physical drop, and infants placed on one side of this. Mothers on the opposite side of the cliff were asked to make a happy or fearful face as their 12 month old child approached to cross. No child crossed whose mother showed a fearful face, but the vast majority of children crossed when their mother showed a happy face. The clear implication is that not only were the children capable of recognising their mother's expression, but also that their behaviour was regulated by their mother's apparent emotion – and that perhaps this and similar processes provide a means by which infants begin to learn connections between specific events or precipitating conditions and the emotions that they provoke.

A range of studies of evident emotion understanding have been conducted with somewhat older children. In one classic study, Borke (1971) asked child participants to choose the emotion that characters would feel in certain situations from a selection of pictures of faces displaying afraid, angry, happy, or sad expressions. Findings revealed that children as young as 3 years old showed some awareness that, "other people have feelings and that these feelings vary according to the situation in which the individual finds herself" (Borke, 1971, p. 269). Borke's findings argued strongly against the Piagetian theory of a preconceptual period where children of this age are expressing egocentric emotion i.e. where they are merely reflecting their own distress or happiness in such situations as opposed to having genuine grasp of others' feelings and how these are connected to actual events (cf. Woodhead et al., 1995).

Further experimental studies have examined children's understanding of non-evident emotion i.e. emotion felt but not expressed e.g. (Mossler, Marvin & Greenberg, 1976; Harris, Olthof, Terwogt & Hardman, 1987; Harris, 1989) These studies have focused on children's ability to understand emotions in others as conveyed through language and within conversation, instead of through facial expressions, as in Borke (1971). This research has examined children's understanding of common emotions (including those employed by Borke), with children being asked to describe situations which provoke such emotions in others. Evidence from these studies indicates that only about 5% of 3 year olds can answer emotion questions of this type non-egocentrically, whereas a majority of 4 and 5 year olds, and all 6 year old children can answer in this way, and can therefore engage in conceptual perspective-taking regarding thoughts and feelings. This challenges Piagetian views that such inferential thinking is not in place until the child is seven. However, these same studies suggest that more sophisticated emotions such as pride, shame and guilt, which involve consideration of the approval, disapproval and expectations of others, are not understood until about 7 years old, when understanding of the fact that people can feel mixed emotions also starts to emerge.

The above research would suggest that happiness, sadness, anger and fear can be recognised fairly reliably in the facial expression of others by infants, and that they know something about the conditions that are likely to provoke these reactions. However, ability to discuss inner states in the absence of obvious expression would appear to be much more limited among younger children, and restricted primarily to answers to inquiries about their own feelings using basic emotion state expressions, including use of terms such as, afraid, angry, happy, and sad (Ninio & Snow, 1996; CHILDES Language Database (MacWhinney, 2001).

The precise point at which things might begin to change is unclear, but it appears that under more naturalistic conditions at least, it may be somewhat earlier than unsupported testing of competences seems to suggest. In order to recognise non-evident emotions, the child has to have some understanding of the causal relationship between event and emotion, so that the emotion can be inferred as a likely response to an event, even if it is not actually evident. Bretherton's (1991; 1993) work shows that children begin to have some understanding of the causes of emotions from 2 years, and thus have something of the equipment to deal with references to non-evident emotions at this stage. Bretherton's research has established that by 28 months of age, children's vocabulary is rich enough to make causal statements about a range of internal states including inner state references to the basic emotions of happy, sad, and scared terms. Dunn (1994) puts the age of onset of such understanding even earlier, finding that children aged 18-24 months labelled a wide range of feeling states and used causal statements regarding feelings to influence their mothers. Few other researchers put emergence of such discussion so young, though, and the data from Study 2 above suggest that more generally that Dunn's data relate to a particularly precocious sample. The Bretherton results are more in line with Ninio & Snow's (1996) observational data, which identify the start point for non-evident emotion conversation at 28-32 months of age, indicating this as the age at which children's understanding might first begin to become apparent in an experimental setting, a key point for the design of Study 3.

Ability to recognise and discuss the causes of emotion is not quite sufficient in itself, though. In order to make inferences about non-evident emotions – and thus be capable of understanding references to these – children also need to understand that an emotion can be felt but not expressed. This may entail at least an implicit grasp of what are called display rules (e.g. that when your granny gives you a present you don't like you don't express sadness or anger). Cole (1986) indicates that some appreciation of this begins to be evident about 4 years, though in terms of performance rather than explicit description. Moreover, even when equipped with causal understanding and some grasp of emotional display, children would still need to work out that a reference to a non-evident emotion was drawing attention to this kind of connection. In this sense then, understanding such references ought to be a more complex achievement than understanding references to past events, which are at least external in character, and in all probability are more commonly used: references to non-evident emotions would seem on the face of it to be a more specialised usage in everyday conversations – if, as Ninio & Snow (1996) claim, grasp of such references has its origin in the conversational practices of caregivers.

In Ninio & Snow's account, in the same way that children appear reliant upon the *conversational* support of adults with regard to references to non-immediate objects

(Study1), and depend on the *historical* support of parents with regard to references to shared past events (Study 2), they also need the *psychological* support of adults with respect to conversations that involve non-evident emotion and belief states (Ninio & Snow, 1996). Psychological support of emotion understanding is met for young children by an adult's willingness to take the child's perspective on events (e.g. 'toppling towers *are* funny'), helping to render aspects of this perspective explicit by encoding it in conversation. In this way, the child is provided with a resource for future conversations. However, de Rosnay & Hughes (2006) also argue that that not all psychological support from parents takes children's perspective on events. Caregiver communication to young children involving 'empathy related statements' (e.g. 'poor little girl') might help children to understand another's perspective on events. "Such comments underline the efforts that can be made to draw children into another person's point of view and raise the issue of a perspective-shift within discourse" (de Rosnay & Hughes, 2006, p. 21). These authors also point to ways in which this kind of discourse might build on causal knowledge: "Empathetic concern for others [has] proved to be strongly related to the situational determinants of emotion" (2006, p. 21).

Such research has led de Rosnay & Hughes to call for a distinction within future research on conversational practices between "interlocutors who elaborate on their own and children's current circumstances (including thoughts and feelings) and interlocutors who engage children in more *decontextualised* discussion about the perspectives of other people" (de Rosnay & Hughes, 2006, p. 21). The implication is that children's understanding of non-evident emotion might lie in the latter rather than the former.

4.1.2 Rationale for Study 3

The research findings outlined above form the basis for the rationale for Study 3. The intention was to examine children's understanding of references to non-evident emotion states as the third sub-maxim in Ninio & Snow's (1996) developmental progression. In light of the research into young children's use of inner state terms, and arguments concerning the impact of *extended* psychological support of caregivers through empathy-related statements, it seems reasonable to presume that some understanding of references of this type could be present in children from about 30 months of age onwards. Ninio & Snow's (1996) observational research of mother-child verbal exchanges regarding emotion in the home; the ever-improving skills in the use of pragmatic language by children aged 29 months found by Babelot & Marcos (1999), and examination of age-related emotion conversations between children and parents evident in the CHILDES Language Database (MacWhinney, 2001) all point in this direction. Moreover, Study 2 had shown that the youngest group's ability to understand references to shared past events was extremely limited. Since Ninio & Snow (1996) and the CHILDES Language Database (MacWhinney, 2001) both indicate that parental discussion of non-evident emotions does not occur until after discussion of shared past events – implying that grasp of references to these is a still later acquisition – there seemed no grounds to retain this age group in Study 3. A new sample of only middle (2.5 to 3.0 years) and oldest age group children (3.5 to 4.0 years) was therefore drawn. This had the additional benefit of permitting more systematic investigation of language effects and their relationship to the age component identified in the data from Studies 1 and 2, since it was possible to use a single language measure with these age groups.

The more specific objective of Study 3 was to check whether evidence on the pattern of performance was consistent with the sequence of emergence of this third sub-maxim within caregiver-child conversations described by Ninio & Snow (1996). To this end, the request task, which had worked well thus far, was retained. However, given the nature of the type of reference now under investigation, it was considered that it was not suitable to contrast competence with respect to non-evident emotions with grasp of references to immediate, present event objects, the baseline in Studies 1 and 2. Several points had an influence on this. One was a desire to avoid confounding understanding of references to non-evident emotions with non-immediate presence, whether due to physical or temporal remoteness from the immediate focus. Whilst references to non-evident emotions in remote persons are of course entirely possible, to use these as test items would have left the nature of any observed constraints on performance unclear. Both the target of such references and any baseline comparison targets therefore had to be present in the immediate context.

At the same time, a contrast between references to immediate but non-evident emotions and references to immediate objects would have served no purpose, since a) the grasp of 2.5 to 4.0 year olds with regard to the latter had already been established to be good from Study 1; b) it would have led to a certain lack of task coherence; and c) the more obvious point of contrast is between references to internal, non-evident but inferable emotions, and external, publicly-evident emotions as indicated by either expressive display (less easily manipulated when using toy animals, as was the case here for consistency at this level with the earlier studies) or explicit verbal labelling in line with conventional event-emotion pairings (e.g. parties are associated with happiness). The comparison made by Study 3 was therefore between references to a non-evident but inferable condition of sadness, and to an explicit conventional condition of happiness.

In terms of the task content within which both types of reference were framed, this drew upon a familiar area of socialisation for young children by focussing on social relations and friendships between peers. Between the ages of 2.5 and 4.0 years, associative and co-operative activities with peers increase greatly at nurseries, with friendships being founded upon the desire to play with one's friend or to sit beside them. References to both evident and non-evident emotion states were therefore predicated upon children's recognition of these based upon an understanding of how two pairs of soft toy *animal friends* would feel in certain situations (cf. de Rosnay & Hughes', (2006) argument of a relationship between empathetic concern and knowledge of the situational determinants of emotion).

Both the happy (evident-emotion state) and the sad (non-evident emotion state) were the subject matter of a story that was recited and enacted using the soft toy animals in view and hearing of each participating child. A happy emotion context was first established during the story when children were told that two pairs of soft toy animals, an *Elephant* and a *Frog*, and a *Bear* and *Rabbit*, were good friends with each other, who liked to sit together. Four toy chairs in two conjoined pairs were provided to help enact the story and create a sense of happiness based on each paired animal's *desire* to sit beside the other. A sad but non-evident emotion state for *one* of the pairs of animals was then created as the story unfolded, when both pairs of animals were invited to the party of a fifth toy animal, a *Lion.*

At this point, a table was produced and the participating child heard and saw how the *Lion seated* one pair of animals (e.g. the *Elephant* and the *Frog*) together on the long side of the table. This created an *evident* happy state since the causal association between sitting together and being happy had already been explicitly established. However, the Lion *seated* the other pair of animals separately at either end of the table, giving rise to a sad state at being seated apart which was *non-evident* since it was merely the implicit corollary of the voiding of the condition giving rise to happiness, and at no point in the story did the experimenter mention sadness. The happy and sad emotion states were counterbalanced as to which was presented first, as was the pair of animals associated with each emotion.
Various understanding checks were used during the recital and enactment of the story to ensure that children were following the events. For instance, after the animals were first *seated* children were asked if they could remember why the animals are happy, the anticipated answer being because they like to sit beside each other. If a child did not give the anticipated answer, the experimenter returned to the beginning of the story and started again. If the child failed to get this correct a

second time the child was excluded from the rest of the procedure and thanked for their participation.

A further understanding check was posed midway through the story, at the point prior to the *animals* being *seated* at the party, to ensure that children were still keeping track of events by ascertaining if children could still remember where the animals would want to sit. Again, the anticipated correct answer was ‘beside each other’ or to point to the long sides of the table, and if a child got this wrong the experimenter returned to the beginning of the story again.

After animals were seated, according to paired happy and paired sad seated positions, final understanding checks asked children if the paired animals were sitting where they would want to sit and how they felt. The experimenter noted children’s answers at this point, but gave no feedback: the story simply continued. To reiterate, at no point during the outlined procedures was the emotion state of sadness mentioned by the experimenter. Furthermore, the understanding checks, by keeping track of the animal’s desire, as the story unfolded, also served to control for the possibility that children would judge emotion states according to their own belief and desire responses in such situations, a criticism directed at the early Borke (1971) study. At the end of test questions, to check for the possibility of random responding, children were asked to explain why they thought that two animals who had been seated apart were sad.

In other respects, the request task retained the form used in Studies 1 and 2. In particular, references to happy and sad emotions were manipulated by the experimenter's use of direct and indirect requests, as before, to give a toy animal that was happy or sad. This enabled a more fine-grained assessment to be made of children's grasp of references to non-evident emotions, via their apparent degree of sensitivity to the pragmatic demands imposed by the different request forms. Children's responses were categorised according to response types similar in form to those used in Studies 1 and 2, adapted to record whether children gave an incorrect response type that involved giving a different non-emotional state toy than the one requested.

In one further departure from the methodology employed in Studies 1 and 2, no test of working memory was used in conjunction with the emotion request task, since it did not lend itself well to any form of probe that was task-related in the same way as that used for the past event task. Children were however asked to complete Siegal & Beattie's (1991) version of the false belief test (i.e. reporting on where a character would look '<u>first</u>' for an unexpectedly transferred object) following the request task, in order to determine how far children's request task performance might be related to this, as well as to language ability and age. The research reported by Ninio & Snow (1996) indicates that understanding of references to non-evident emotions precedes understanding of references to non-evident beliefs (as a more complex cognitive condition with less certain causal precursors). However, given obvious structural correspondences between the inferences required by these two types of reference, it was considered important to use Study 3 to explore relationships between understanding of references to non-evident emotions, and a basic grasp of false belief (cf. the discussion in Chapter 1 with regard to the Siegal & Beattie,1991 ‘Look first’ FB test). As de Rosnay & Hughes (2006, p. 17) argue, “children’s understanding of mind [is] indexed by false belief and emotional understanding tasks”. The underlying point here is that whilst actual understanding of references to non-evident beliefs (and hypothetically related ability on false belief tasks – see Chapter 1) may come later, conversational support for grasp of non-evident emotions may have some facilitating effect on this, since it should tend to have an impact on locating appropriate targets for a similar, if perhaps simpler, process of inference.

4.1.3 Key questions and hypotheses for Study 3

The central question addressed by Study 3 was whether children between 2.5 and 4.0 years understand references to non-evident emotion states; and more specifically, whether the degree of understanding exhibited by these children indicated that grasp of such references emerges later than understanding of references to shared past events. On the basis of Ninio & Snow’s (1996) developmental progression, it was hypothesised that some understanding of references to non-evident emotion states would be apparent amongst participating children, but that this would be at best very limited among 2.5 to 3.0 year olds, and only intermittent amongst 3.5 to 4.0 year olds. In line with this, it was further expected that where children were able to give correct responses there would be little or no sign of sensitivity to the form of request that had been used, and that even amongst 3.5 to 4.0 year olds, there would be widespread non-responses, indicating that grasp of this form of reference was insufficiently secure to be accompanied by much awareness of the pragmatics surrounding its use.

It was also predicted that children's ability in this specific conversational area would be related to expressive language ability and to performance on the 'Look first' FB test (Siegal & Beattie, 1991). Finally, it was expected that the relationship to language and also to age would be linear in character, consistent with developmental change with respect to grasp of references to non-evident emotion being gradual, as with understanding of references to non-immediate objects and to shared past events.

4.2 Methodology

4.2.1 Design

Study 3 employed a mixed between-and within-subjects design involving eight trials for each child participant. There was one between-subjects variable; age (either the middle or oldest age group). The within-subjects variables were context (evident vs. non-evident emotion) and request form (direct vs. indirect). As in Study 2, children's responses were coded in terms of whether or not they were correct and also whether they were action, verbal or non-responses. To allow for the examination of degrees of inaccurate responding, the coding of actively incorrect responses for evident and non-evident emotion was undertaken. A language test i.e. the British Ability Scales II, Naming Vocabulary Test (BAS II, 1996) was also conducted on each child. Participating children were also tested using the 'Look first' false belief FB test (Siegal & Beattie, 1991).

4.2.2. Participants

Study 3 employed a new sample of children from those who had jointly participated in Study 1 and 2. The participants were 60 pre-school nursery children (31 girls, 29 boys). The sample was divided into 2 age groups i.e. a middle and an oldest age group. The label 'middle' applied to the younger of the two age groups and was retained for the purposes of comparison with Studies 1 and 2. There were 30 children in the middle group, 2.5 to 3 years old (15 females, 15 males; mean age = 2 years, 9 months), and 30 children in the oldest group 3.5 to 4 years old (16 females, 14 males; mean age = 3 years, 9 months). Parental consent was given with respect to all participating children in Study 3 and children were advised that they could withdraw from the study at any time. All children spoke English as the language of their home. Ethical clearance and local authority permission had remained valid from Studies 1 and 2.

4.2.3 Materials

The main element of materials consisted of three toy collections. The first and second toy collections were items relating to the Study 3 story and questions. The first was a collection of soft toy animals and the second was a toy table and toy chairs and both collections were of a form familiar to children within these age groups. The animals consisted of a *Bear, Elephant, Frog, Lion and Rabbit.* No animal was smaller than 15 cm or higher than 18 cm. These soft toy animals were chosen according to the '*neutrality of their facial expressiveness'* i.e. these animal's facial appearance did not indicate either happiness or sadness and this was verified during pilot testing with a different sample of 18 children in the age range of 2.5 to 4 years old. All 18 children in the pilot test were shown each of the 12 animals one at a time and then asked to decide whether an animal was happy, sad or neither happy nor sad. Only animals that children chose that were neither happy nor sad were utilised.

Both the toy table and toy seats were made of plastic and were of a standard type used by young children to represent dining furniture when playing toy *'kitchens*' or toy '*picnics'* with soft toy animals or other toy dolls. The table measured 49 cm in length, 23 cm in width and when upright was 15 cm high. All four toy seats were identical and were of a type that had a seated base and a high back. Each chair measured 12 cm in length, 6 cm in width and from base to high back measured 24 cm. The seats dimensions fitted comfortably with both the size of the table and also with *seated* soft toy animals i.e. from base to seat measured 12 cm high thus ensuring that children could easily view each soft toy animal when it was *seated* at the table. The dimension of the table also allowed for the positioning of two seats beside each other along the table's length and also for a single chair at either end of the table along its width.

The third toy collection related to items required for the 'Look first' FB test (Siegal & Beattie, 1991); two soft toy animals, a large soft toy *cat* measuring 25 cm and a small soft toy *cat* measuring 15 cm; two toy cupboards, identical in dimension, 12 cm x 7cm x 11cm, each with a sliding drawer, one coloured blue and the other coloured green; one toy food, *chocolate* that fitted neatly into the cupboard's drawers when these were closed.

A large green cloth mat was also used to conduct the test sessions upon nursery floors. The experimenter scored children's responses during the test sessions using the pro-formas shown in appendices E & F. A video camera on tripod was utilised to record children's responses, thus enabling the coding of responses to be rechecked by the experimenter and then verified by an independent judge at a later date. Permission to use and record children's responses by video, in addition to their participation approval, was sought and granted from participating nurseries and parents, who were both informed that the video footage would be erased once the project was complete.

4.2.4 Procedure

Each nursery was visited on separate occasions over a nine month period. Testing took place within nurseries in three stages.

Stage 1: Warm-up period, Language Testing.

Nursery assistants escorted each child into the experimental room and remained with them during the course of their questions thus enabling children to have a familiar face present, which also helped to put them at their ease. A 15-minute 'warm-up period' then took place where the experimenter introduced himself and, to build rapport with children, engaged in a 'play scenario' involving the soft toy animals i.e. the *Bear, Elephant, Frog, Lion and Rabbit*. At the end of the warm-up period, when the game was finished, the British Ability Scales II Naming Vocabulary Test (BAS II) (1996) was administered using standard procedures as described in Study 1.

Stage 2: Tests of understanding of references to non-evident emotions

When the British Ability Scales II Naming Vocabulary Test (BAS II) (1996) was completed children viewed the experimenter '*pairing off*' the four soft toy animals in two groups of two in order to establish friendships between them. This involved the experimenter placing Set E soft toy animals i.e. the *Elephant* and *Frog* beside each other on two toy chairs and also Set F soft toy animals i.e. the *Bear* and *Rabbit* together on two other toy chairs. To retain the distinction of the paired friendships, Set E and Set F animal pairs and their joint seating arrangements were placed some 40 cm apart on the green mat in full view of participating children. Children saw all animals, but which pair was designated as targets for references to evident and non-evident emotions was counterbalanced, as was the order in which the evident and non-evident emotion targets were presented. One soft toy animal, the *Lion,* which would be produced later during the procedure to children, was not allocated to a set.

All children were told that they were going to hear the following story about *'Animals who are best friends that like to sit beside each other'*. For the purposes of keeping the following outline succinct, procedure from this point forward will be described as encountered by those children for whom the Elephant and the Frog had evident emotion states. The words of the experimenter are in inverted commas and italics, with numbered points and procedure in standard font.

The experimenter introduced the story to children using the following title,
"Animals who are best friends that like to sit beside each other".

(1) *"This is Elephant and this is Frog and these are their seats"*. The animals and two seats were introduced to the green mat in full view of the child.

(2) "*Elephant and Frog are really good friends and they are such good friends that they always like to sit beside each other"*. Elephant and Frog were placed beside each other on each of the seats in full view of the child.

(3) *"Sitting together makes Elephant and Frog feel very happy*".

(4*) "This is Bear and this is Rabbit and these are their seats"*. These animals and two seats were introduced to the green mat in full view of the child.

(5) *"Bear and Rabbit are also really good friends and they are such good friends that they always like to sit beside each other too"*. Bear and Rabbit were placed beside each other on their two seats in full view of the child but with a distance of about 40 cm between the sets of animals/seats.

(6) "*Sitting together makes Bear and Rabbit feel very happy too".*

A series of understanding checks was included from this point forward in the procedure in order to confirm that children were following the events of the story and understanding the evident happy or non-evident sad emotions that these should produce in the toy animals. For example, 'Understanding check, emotion, 1st pair' refers to children understanding that the *Elephant* and *Frog* are happy because they like to sit beside each other.

(7) 'Understanding check, remember, 1st pair'; Child is asked, *"Do you remember why the Elephant and Frog are happy now?* If children did not give an answer such as, 'They are happy because they like to sit beside each other' the experimenter returned to point 1 and started again. If a child did not get this correct the second time he or she was excluded from the rest of the procedure and thanked for their participation. (Two middle group children were excluded from testing at this point having failed again to give a correct answer to, 'Understanding check, remember, 1st pair'. Two other children were recruited to the middle group to replace them.

(8) 'Understanding check, remember, 2nd pair'; Child is asked, *"Do you remember why the Bear and Rabbit are happy now?* The same procedure was applied as in (7) where initial responses were incorrect.

(9) With the understanding of happiness ensured for both sets of paired and seated animals, and with these still in children's view, the child was then told, *"Lion is having a party and he has invited the pairs of friends."* Lion and a table were introduced with the long side of the table in direct view of the children.

(10) At this point evident and non-evident emotion state was varied systematically i.e. half the children would be presented with the circumstances producing a evident happy emotion first, and the other half the non-evident sad emotion first, to control

for practice or recency effects. (Procedure continues to be outlined below for a child who receives the evident happy emotion state first involving the Elephant and the Frog).

(11) *"Lion is going to tell Elephant and Frog where to sit at the table*".

(12) 'Understanding check, seating, 1st pair'; the child is asked, "*So where would Elephant and Frog want to sit at the table?"* If a child failed to give an answer such as 'Beside each other' or to point to the long side of the table, the child's response was noted and the child remained in the test, and procedure continued as follows, *"Lion wants Elephant and Frog to sit beside each other at the table"* (Elephant and Frog, in their seats, are placed beside each other on the long side of the table, in direct view of the child).

(13) 'Understanding check, emotion, 1st pair'; The child is asked, "*Are Elephant and Frog sitting where they want to sit at the table?" "So how do they feel?"* At this point the experimenter gave no verbal or non-verbal indication to the child as to whether their response was correct and the child's response was simply noted.

Twenty-eight out of the overall sample of thirty older group children (93%) gave a reply that indicated that they were keeping track of events e.g. "Yes" and "Happy" respectively. Twenty-five out of the overall sample of thirty middle group children (83%) also replied in this way. Answers that varied from those outlined included, 'They shouldn't be allowed to sit together in case they carry on': 'They are not happy because there is no food':'Lion is cross because they have no presents for him': The seven children who evinced difficulty were not excluded from the study because they had at least shown some grasp of what the scenario was about.

(14) *"Lion is now going to tell Bear and Rabbit where to sit at the table"*

(15) '<u>Understanding check, seating, 2nd pair';</u> the child is asked, "*So where would Bear and Rabbit want to sit?"* See (12) 'Understanding check, seating, 1st pair' (above) for procedure if a child failed to give an appropriate answer.

(16) *"Lion wants Bear and Rabbit to sit apart from each other at the table"* (Bear and Rabbit, are separated from their paired seating, and are placed at the opposite (width) ends of the table in view of the child).

(17) '<u>Understanding check, emotion, 2nd pair';</u> The child is asked, *"Are Bear and Rabbit sitting where they want to sit at the table". "So how do they feel?"* Again, at this point the experimenter gave no verbal or non-verbal indication to the child as to whether their response was correct and the child's response was simply noted.

(18) *"Lion goes away to prepare for the party*". Lion is removed from testing
<u>At no point during the procedure, story or understanding checks is 'sadness' mentioned by the experimenter.</u>

Twenty-seven from the sample of thirty oldest group children (90%) gave a response that indicated that they were keeping track of events e.g. ‘No’ and ‘Sad’ or ‘Not very happy’ respectively. Twenty-six of the sample of thirty middle group children (86%) also gave a response that indicated that they were following events by giving a “No” and a ”Sad” or “Not very happy” answer. Answers that varied from those outlined included, ‘No’ to the seating arrangements but that the animals were ‘Cross’ (angry) at not sitting beside each other and ‘Yes’ but that ‘The animals were scared at being away from each other’. On the face of it, these appear, to be reasonable responses, and not entirely inconsistent with being able to understand subsequent references to sad animals, since they recognise they would be experiencing a negative emotion, and for children of this age the positive/negative distinction is typically the one that carries most weight (Tolmie, 1991). Again, the seven children who evinced difficulty here were part of a non-overlapping set and they had exhibited at least a partial understanding of the grounding scenario. All 60 children’s data were included in the subsequent analysis on the basis that those 14 children evincing difficulty got at least one understanding check correct.

As was the case for the Study 1 and 2 procedure, children’s ability to decode references, in this instance to a evident happy state and a non-evident sad state was then tested by examining their responses to a series of requests made by the experimenter using two different request forms, i.e. direct request, (“Give me the X”, where X refers to a toy item) and an indirect request, (“Have you got the X?”. Children were told that the experimenter wished to ask them some questions about the story and the toy animals. The questions (see Tables 4.1 and 4.2) were put to children in the manner described in Studies 1 & 2. However, unlike Studies 1 & 2, where each question permitted only one response, here the task allowed either animal from a pair to be identified in response to each of the experimenter's requests. Since correct responses still required accurate identification of the referents, though, it was felt this created no material difference. Children's responses were recorded at the time by the experimenter on either the Appendix E or F pro-forma, as appropriate, and also on videotape. The video recorder had been switched on after each child had heard the story and understanding checks had been administered.

At the end of the trials each participating child was given a happy face sticker and thanked for their participation.

Table 4.1: Order of request by reference type and request form for Set E children

REQUEST-OBJECT	REFERENCE TYPE	REQUEST FORM
1. Give me an animal that is happy (Elephant or Frog)	Evident emotion	Direct
2. Have you got an animal that is sad? (Bear or Rabbit)	Non-evident emotion	Indirect
3. Have you got an animal that is happy? (Elephant or Frog)	Evident emotion	Indirect
4. Give me an animal that is happy (Elephant or Frog)	Evident emotion	Direct
5. Give me an animal that is sad (Bear or Rabbit)	Non-evident emotion	Direct
6. Have you got an animal that is happy? (Elephant or Frog)	Evident emotion	Indirect
7. Give me an animal that is sad (Bear or Rabbit)	Non-evident emotion	Direct
8. Have you got an animal that is sad? (Bear or Rabbit)	Non-evident emotion	Indirect

Table 4.2: Order of request by reference type and request form for Set F children

REQUEST-OBJECT	REFERENCE TYPE	REQUEST FORM
1. Have you got an animal that is sad? (Elephant or Frog)	Non-evident emotion	Indirect
2. Give me an animal that is happy (Bear or Rabbit)	Evident emotion	Direct
3. Give me an animal that is sad? (Elephant or Frog)	Non-evident emotion	Direct
4. Have you got an animal that is happy? (Bear or Rabbit)	Evident emotion	Indirect
5. Have you got an animal that is happy? (Bear or Rabbit)	Evident emotion	Indirect
6. Give me an animal that is happy ? (Bear or Rabbit)	Evident emotion	Direct
7. Have you got an animal that is sad? (Elephant or Frog)	Non-evident emotion	Indirect
8. Give me an animal that is sad (Elephant or Frog)	Non-evident emotion	Direct

After the above questions were put to children it was necessary to check for the possibility of random responding. Therefore, with the table and respective evident happy and non-evident sad seated animals still in view, children were asked to justify why they thought that the animals were happy or sad. Justification questions involved pointing to each animal and asking the child e.g. 'Why is the frog sad'? Twenty-nine from the overall sample of thirty older group children (96%) gave a response that justified that animals were happy because they were 'Sitting together'. Twenty-eight from the overall sample of thirty middle group children (93%) also justified happiness by the animals seating arrangements. One child said the animals were happy because they were, 'Asked to the party'. One child gave no-response and another child stated, 'Don't know'. Twenty-seven from the sample of thirty older group children (90%) gave responses that justified why the animals were sad e.g. ‘Because they were not sitting together’, or ‘Not sitting beside each other’ or were sad because ‘Lion kept them away from each other at the table’. Twenty-five from the sample of thirty middle group children (83%) also gave justification responses identical to those just outlined. Answers that varied from those outlined included,’ Because they are cross’. ‘That was not nice of Lion’ i.e. although these responses are not as clear as one might wish, they were certainly not inconsistent with understanding the scenario. There was also a comment of ‘Don’t know’ by one child and two children gave no verbal responses.

Stage 3: The ‘Look first’ false belief (FB) test.

The ‘Look first’ FB test, as utilised in the research of Siegal & Beattie (1991) was then conducted in a counter-balanced design i.e. half of the children from each of the two age groups received the test before Stage 2 and the remaining half of children from each of the two age groups received the test following Stage 2. The test was administered using standard procedures as established by both Wimmer & Perner (1983) and Siegal & Beattie (1991) as follows.

All other soft toy animals were removed from the green mat and two soft toy Cats, one large and one small, and the cupboards were introduced to testing. The child was recounted the following scenario with the words of the experimenter in inverted commas and procedure in brackets.

(1) “This small Cat is called ‘Kitty’ and this big Cat is ‘Kitty’s Mum”.

(2) “ These are two kitchen cupboards, one is blue and one is green and both have drawers” (both cupboards were shown to the child with empty drawers open and then closed and these were placed on the green mat beside each other with children having a clear view of the drawers and the front of each cupboard).

(3) “Kitty’s Mum leaves the kitchen” (and Kitty’s Mum was then temporarily removed from testing and placed in the experimenter’s holdall).

(4) “Kitty likes chocolate” (and toy food chocolate was introduced to testing).

(5) “ But Kitty wants to go out and play”.

(6) “ So Kitty puts her chocolate in the blue cupboard drawer” (and the chocolate is placed in the blue cupboard drawer and the drawer is closed).

(7) “Kitty goes out to play” (and Kitty is temporarily removed from testing i.e. placed out of the child’s sight in the experimenter’s holdall).

(8) “Kitty’s Mum comes into the kitchen and moves Kitty’s chocolate from the blue cupboard drawer to the green cupboard drawer” (Both cupboard drawers are then closed).

(9) "Kitty's mum then leaves the kitchen again" (and is removed from testing).

(10) "Kitty comes back into the kitchen and wants her chocolate".

(11) (The following question is asked of the child),
"In which cupboard will Kitty look <u>first</u> for her chocolate"?.

4.2.5 Coding of Responses and Scoring

The experimenter observed and recorded children's responses during testing and these responses were also later rechecked and transcribed from videotape after data collection was complete.

Request task

Children's responses to the request task were coded into one of six types, as described below. Three of the response types were drawn directly from the work of Babelot & Marcos (1999) and have been outlined in Study 1 and 2. These were types 1, 5 and 6. Response types 2 and 4 were related to the Information Response utilised by Babelot & Marcos, (1999) and again, these are described in Study 1 and 2. One of the Response types, Types 3, an Incorrect Action response (gives different emotion object) or (GDEO), was devised specifically for the coding needs of Study 3 in order to identify when a child gave the wrong emotional state object. (Interjudge agreement rates are again given below).

1. Correct action response = 97%

2. Correct verbal/gestural response = 97%

3. Incorrect action response (gives different emotion object) (GDEO) = 100%

4. Incorrect/uninformative verbal response = 93%

5. Incomprehension response = 100%

6. Non-response = 93%

The number of children's responses in each of the six categories was totalled across trials of the same reference type and request form to give a score between 0 and 2 for each. Accordingly, six scores, one for each response type, were derived for each of the following trial types: evident happy direct, evident happy indirect, non-evident sad direct and non-evident sad indirect. Also, to check on order effects, the number of children's responses in each of the six categories was totalled separately for the first two and then the second two requests for each reference type, regardless of request form, to give a score between 0 and 2 in each case. In this way, six scores, one for each response type were derived for evident happy first questions; evident happy second questions; non-evident sad first questions; and non-evident sad second questions.

4.2.6 Language test scoring.

In the BAS II NVT (1996) children responses were scored in the standard fashion to give a total out of 16.

4.2.7 'Look first' false belief (FB) test scoring

Children giving the answer 'blue cupboard' received one point as it indicated that they could represent a character's belief of where they would 'look first' for an unexpectedly transferred object and thus retain a distinction from what they knew had occurred in current reality. Children answering 'green cupboard' received 0 points as this indicated they had not represented the characters belief and judged that the item would be placed where it currently was.

4.3 Results Section

4.3.1 Summary of Analyses

Three preliminary analyses were conducted on the request task data: firstly, to check for the effects of question order (i.e. whether there was any indication of practice effects); secondly, to check whether there was any effect of emotion order by way of primacy or recency effects (i.e. as children had received either set E or F for counterbalancing purposes); and thirdly whether there was any effect of question set/sequence i.e. which set the children were in, and thus which sequence of questions they received for the request task. The results of these analyses were as follows. A 3-way mixed ANOVA (age group x reference type x order) was performed on the scores for the evident emotion first questions, evident emotion second questions, non-evident emotion first questions and non-evident emotion second questions for each of the six response types outlined in 4.2.5. No significant effects or interactions were found involving question order. A 4-way mixed ANOVA (age group x emotion order x reference type x reference form) was conducted on the scores of evident emotion/direct, evident emotion/indirect, non-evident emotion/direct and non-evident emotion/indirect by the six response types outlined in 4.2.5. No significant effects or interactions were found in relation to the emotion order used. A 4-way mixed ANOVA (age group x set x reference type x reference form) was conducted on the scores of evident emotion/direct, evident emotion/indirect, non-evident emotion/direct and non-evident emotion/indirect by the six response types outlined in 4.2.5. No significant effects or interactions were found in relation to the question set/sequence used. These factors will not be considered further.

The main analyses were then computed: 3-way mixed ANOVAs (age group x reference type x reference form) on scores for evident emotion/direct, evident emotion/indirect, non-evident emotion/direct and non-evident emotion/indirect requests for each of the six response types.

In keeping with previous studies, a 2-way ANOVA was then computed on the overall level of correct responses collapsed across correct action and correct verbal/gestural responses. This was similarly followed with a 3-way mixed ANOVA that was conducted to examine active vs. inactive incorrect responses (age group x reference type x active vs. inactive). The results of these analyses are reported in this order in what follows.

With the main patterns of performance on the request task identified, the relationship of correct responses to language scores was explored using correlational analyses. It was then necessary to examine whether overall correct responses, correct evident and correct non-evident emotion responses were related to the 'Look first' FB test (Siegal & Beattie, 1991) and this was computed using point-biserial correlation co-efficients. Further examination of the relationship between the 'Look first' FB test, age and correct responses was then explored using chi-square analysis. Regression analyses were then employed to further examine a potential relationship between the 'Look first' FB test, language level and conceptual grasp and the request task.

4.3.2 Descriptive and Inferential Results for the Request task

Table 4.3 shows the mean scores and standard deviations for the response types used by children according to age group, reference and request form. Appendix N presents full results for the 3-way ANOVAs conducted on these data.

TABLE 4.3: Means and standard deviations for each response code (range = 0-2 by reference type, age group and request form

			Correct action resp		Correct verbal resp		Incorrect action resp		Incorrect verbal resp		Non resp	
Ref type	Age group		Dir	Indir	Dir	Indir	Dir	Indir	Dir	Indir	Dir	Indir
Happy (evident belief)	Middle	Mean	0.73	0.60	0.26	0.53	0.63	0.33	0.16	0.10	0.26	0.30
		SD	*0.78*	*0.77*	*0.58*	*0.81*	*0.76*	*0.60*	*0.46*	*0.30*	*0.58*	*0.65*
	Oldest	Mean	1.20	1.26	0.30	0.43	0.50	0.30	0.00	0.00	0.00	0.00
		SD	*0.80*	*0.73*	*0.59*	*0.67*	*0.73*	*0.53*	*0.00*	*0.00*	*0.00*	*0.00*
	Total	Mean	0.96	0.93	0.28	0.48	0.56	0.31	0.08	0.05	0.13	0.15
		SD	*0.82*	*0.82*	*0.58*	*0.74*	*0.74*	*0.56*	*0.33*	*0.21*	*0.43*	*0.48*
Sad (non-evident belief)	Middle	Mean	0.93	0.90	0.26	0.40	0.33	0.30	0.26	0.16	0.26	0.16
		SD	*0.82*	*0.80*	*0.63*	*0.67*	*0.60*	*0.65*	*0.63*	*0.53*	*0.63*	*0.53*
	Oldest	Mean	1.33	1.00	0.30	0.56	0.26	0.23	0.00	0.03	0.06	0.10
		SD	*0.80*	*0.87*	*0.65*	*0.77*	*0.63*	*0.56*	*0.00*	*0.18*	*0.36*	*0.40*
	Total	Mean	1.13	0.95	0.28	0.48	0.30	0.26	0.13	0.10	0.16	0.13
		SD	*0.83*	*0.83*	*0.64*	*0.72*	*0.61*	*0.60*	*0.46*	*039*	*0.52*	*0.46*

Ref type = Reference type (Happy (evident belief) or Sad (non-evident belief); resp = response type; Dir = direct response form and Indir = indirect response form. Incomprehension responses' represented 1.2% of total children's responses are excluded from this table due to their infrequency of use by children.

1. Correct action responses

Considering the main effects and interactions for each between and within-subjects factor in turn and Table 4.3 indicates that the oldest group of children outperformed the middle group by giving the most correct action responses overall. This finding is confirmed by the significant main effect of age recorded ($F(1,58) = 6.80$ $p < .05$). There was no main effect of reference type recorded for correct action responses and this indicates that children understood what was meant by experimenter references to both evident happy and non-evident sad contexts. However, a weakly significant interaction for correct action responses for reference type x age group ($F(1,58) = 4.30$, $p < .05$) shows that the oldest group gave marginally more correct action responses to evident happy than to non-evident sad context (oldest group evident happy mean score = 2.46, sd = 1.13; oldest non-evident sad mean score = 2.33, sd = 1.53). This was in contrast with the middle group of children who gave more correct action responses to non-evident sad context rather than to the evident happy context (middle group evident happy mean score = 1.33, sd = 1.13; (middle non-evident sad mean score = 1.83, sd = 1.46). Furthermore, this finding was contrary to predictions that children would find references to non-evident emotions harder to decode, it shows that there were no differences amongst older children's correct action responding to evident and non-evident references: with middle children's performance actually better in terms of non-evident sad references. It also signifies that the gap between the middle and older group's understanding of non-evident sad emotion may be narrower than their understanding of evident happy emotion. Although, there were no main effects or interactions involving reference form there was a very marginal trend for the both groups to give more correct action responses to direct rather than indirect requests (middle group direct mean score = 1.67, sd = 1.40; middle indirect mean score = 1.50, sd = 1.30) and (oldest group direct mean score = 2.53, sd = 1.43; oldest indirect mean score = 2.27, sd =1.41).

2. Correct verbal/gestural responses

Although the trend was for slightly more correct verbal/gestural responses to be given by the middle group to evident happy rather than non-evident sad context the reverse held true for the oldest group where minimally more correct verbal/gestural responses were given to the non-evident sad context. However, there were no significant effects and there were no main effects or interactions involving age, reference type or reference type x age. As Table 4.3 reveals, more correct verbal/gestural responses were given by both the middle and oldest groups to indirect requests rather than to direct requests and this is supported by the significant main effect of reference form ($F(1,58) = 6.75$, $p < .05$). This may indicate that children were regarding the implicature aspect of indirect requests as different types of request rather than different forms of the same request. In terms of the three levels of pragmatic sensitivity defined in the earlier chapters (i.e. sub-section's 1.7.3; 1.9 or 2.1.1) this would place these children at level 2. As with correct action responses there were no differences between evident happy and non-evident sad references despite the strong prediction of such. However, children did respond differently here in comparison to Study 2 where there was a relative bias toward correct action responses with the effects of indirect requests on response form being largely absent in Study 2. This suggests that children were more on top of the reference element in Study 3, and they were able to pay more attention to the pragmatic subtleties of the request. There were no interactions involving reference form.

3. Overall correct responses

As was the case with findings in Study 1, Study 3 found a significant effect of reference form with more correct verbal/gestural response types given to indirect requests than to direct requests and again, more correct action responses given to direct than indirect requests by both groups. However, in contrast with Study 1, the request form effect for correct action responses in the current study did not meet with conventional significance levels. Given the lack of effects of reference type for correct action and correct verbal responses taken independently, it was necessary to confirm that this was the overall pattern when correct responses in general were considered, by adding these together (as in Table 4.4.).

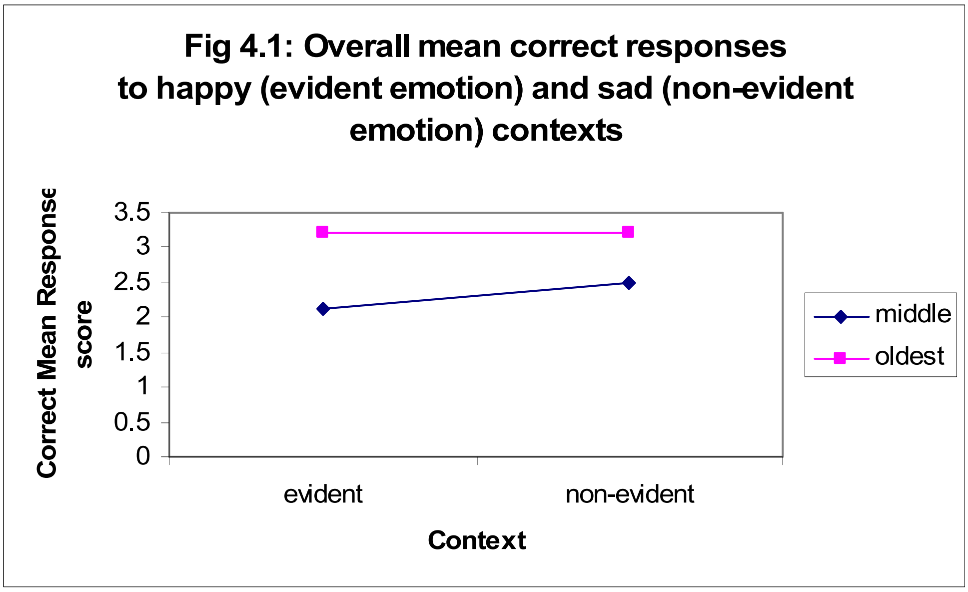
Table 4.4: Total correct response scores by age group for evident happy and non-evident sad references with mean scores and st. dev's

Age Group	Total Correct Responses Evident Happy	Mean Correct Responses Evident Happy	*St. Dev*	Total Correct Responses Non-evident Sad	Mean Correct Responses Non-evident Sad	*St. Dev*
Middle	64	2.13	*1.45*	75	2.50	*1.65*
Oldest	96	3.20	*1.03*	96	3.20	*1.29*
Total	160	2.66	*1.36*	171	2.85	*1.51*

In total there were 331 correct response types given out of a possible 480. Total correct response types made up 69% of overall responses.

Table 4.4 confirms that the most overall correct responses were given by the oldest children. The results of a 2-way mixed ANOVA (age group x reference type) conducted on overall correct responses confirms a significant main effect of age ($F (1,58) = 7.37, p < .05$). Despite no main effects or interactions involving context the trend was for the middle group to give more overall correct responses to non-evident sad context. In contrast the oldest group of children gave equivalent number of correct responses to both evident happy and non-evident sad reference types. The lack of significant effects for overall correct responding according to reference types indicates that all children understood comparably the evident happy and non-evident sad references respectively.

However, given the possibility of choosing 2 out of 4 animals at random, and still apparently giving a correct response, it is necessary to consider how the response rates compare with chance. It is apparent that the middle age group is not performing at much better than chance (2.30 overall correct responses out of 4) whereas the older age group are doing much better, and are not too far off ceiling (3.20 correct responses out of 4). Although this does not mean that the middle group was responding randomly, it does suggest that their grasp was not that good. Figure 4.1 shows these trends



4. Incorrect action responses (gives different emotion object) (GDEO)

On examination of Table 4.3 the tendency appears to be for the middle age group to make more incorrect action responses (GDEO) than the oldest group. However, there were no significant main effects of age recorded in this regard. Giving a different emotion object was more likely to occur when a child is asked for a evident happy animal with a weakly significant main effect of reference type ($F(1,58) = 4.91$, $p < .05$). Similarly, a weak significant main effect ($F(1,58) = 4.36$, $p < .05$) confirms that incorrect action responses (GDEO) were more likely to occur to a direct request rather than an indirect request. An interaction of reference type x reference form for incorrect action responses (GDEO) was not significant at conventional levels ($F(3,88) = 4.36$, $p = 0.53$) with the effect of request form only apparent to the evident happy context.

5. Incorrect/uninformative verbal responses

As is evident from Table 4.3, the middle group of children made incorrect/uninformative verbal responses to a very modest extent while the oldest children's use of this response type was negligible. A weak significant main effect of age ($F(1,58) = 4.64$, $p < .05$) confirms this finding. The tendency was for this response type to be given more in the evident happy context rather than non-evident sad context but the main effect of reference type did not meet conventional significance levels at ($F(1,58) = 3.81$, $p = 0.56$). It is also apparent that incorrect/uninformative verbal responses were made much less frequent than incorrect action responses (GDEO), reflecting the bias to action among correct responses. There were no other main effects or interactions for this response type.

6. Incomprehension responses

Due to their infrequency of use, Incomprehension Responses were not subjected to statistical analysis.

7. Non-Responses.

Looking at Table 4.3, it is clear that non-responses were used by the middle group to a larger extent than the oldest group. However, a main effect of age did not meet conventional significance levels ($F(1,58) = 3.68$, $p = .06$). Similarly, an interaction effect for reference type x age failed to meet conventional significance.

Nonetheless, the tendency was for middle group children to give more non-responses in the evident happy context than in non-evident sad context. By contrast, the oldest group of children did not use non-responses in the evident happy context and their use in the non-evident sad context was quite minimal.

8. Analysis of actively incorrect response types vs. inactive incorrect responses

The relative frequency of active incorrect response types (action and verbal) and inactive incorrect responding (non-responses) was first raised in Study 1 and 2 where it was used to examine a level of pragmatic sensitivity in children to the experimenter's request i.e. the awareness that a request requires a response even when the answer is uncertain. Table 4.5 provides descriptive data in relation to these.

Table 4.5: Active incorrect vs. inactive Incorrect responses scores by age group, and context, including means and st. dev's

Age Group	Total Inactive Incorrect responses Evident Context	Mean Inactive Incorrect responses Evident Context	*St. Dev*	Total Inactive Incorrect responses Non-evident Context	Mean Inactive Incorrect responses Non-evident Context	*St. Dev*
Middle	17	0.56	*0.16*	13	0.43	*1.13*
Oldest	0	0.00	*0.00*	5	0.16	*0.74*
Total	17	0.28	*0.86*	18	0.30	*0.96*
Age Group	Total Active Incorrect responses Evident Context	Mean Active Incorrect responses Evident Context	*St. Dev*	Total Active Incorrect responses Non-evident context	Mean Active Incorrect responses Non-evident context	*St. Dev.*
Middle	37	1.23	*1.22*	32	1.06	*1.38*
Oldest	24	0.80	*1.03*	16	0.53	*1.13*
Total	61	1.01	*1.14*	48	0.80	*1.28*

In total there were 144 Active Incorrect and Inactive Incorrect types given out of a possible 480. Total active incorrect and non-response types made up 30% of overall responses.

Considering Table 4.5 initially, it is apparent that the middle group of children gave the most active and inactive incorrect responses overall. A 3-way mixed between and within subjects ANOVA (age group x context x active vs. inactive) (Appendix P) confirms a significant main effect of age ($F(1,58) = 8.05$, $p < .05$). Table 4.5 also shows that the use of actively incorrect response types was much more prevalent than the use of inactive incorrect responses by both age groups. This pattern was substantiated by the significant main effect of response type that was recorded ($F(1,58) = 8.05$, $p < .05$). This finding shows that there is a modest tendency for both the middle and older groups of children to actively respond, albeit incorrectly, instead of simply not responding at all. This implies an understanding that some form of response was required, if not what the actual referent was to evident happy and non-evident sad references.

To summarise, the main points to emerge from the request task confirms the general lack of an effect of reference type, coupled with good levels of correct responding. There were modest effects of age and a broad lack of interactions between age and reference type. Consideration of these findings show that they are counter to predictions. However, there was some differentiation between request forms as to action vs. verbal responses, but on the whole a bias towards action responses, even when incorrect; with low levels of non-responses. These latter findings suggest a good level of pragmatic sensitivity within the context of this task, consistent with ease of decoding the references.

4.3.3 Language and correct response scores

It was necessary to consider the relationship between the language test used and correct response scores given. Table 4.6 shows the means and standard deviations for BAS II NVT (1996) language test scores vs. correct response scores for both middle and oldest groups collectively.

Table 4.6 Means and standard deviations for British Ability Scales II Naming Vocabulary Language Test and correct response scores for Study 3 for the middle and oldest groups collectively

Middle and Oldest groups	**N**	**Mean**	**Std. Dev.**
BAS II NVT Score	60	15.53	0.89
Correct Response Score	60	5.51	2.65

A Pearson's Correlation coefficient was computed between BAS II language scores and correct response scores for Study 3. Language score and number of correct responses were found to be significantly positively correlated, as predicted, ($r = .51$, $p < .001$ (1-tailed).

The relationship between language test and correct response scores in Study 3 would appear to indicate a moderate to strong linear relationship across the age range sampled.

4.3.4 'Look first' false belief test and correct responses

As the 'Look first' FB test (Siegal & Beattie, 1991) was conducted in counterbalanced order with half of each of the middle and oldest age groups completing the test prior to the request task (first), and the other half of both age groups undertaking the test following the request task (last) analyses were conducted to check for differences in performance. Independent-Samples t-tests confirmed there were no differences in performance for those in either age groups who undertook the 'Look first' FB test first or last.

In order to establish the relationship between understanding of references to evident and non-evident emotions and understanding of false belief, the relationship of the latter to each child's overall correct response score, correct responses to evident and non-evident emotions was examined. Table 4.7 shows the means and standard deviations for the 'Look first' FB test (Siegal & Beattie, 1991) scores and overall correct scores, correct evident and non-evident emotion scores.

Table 4.7: Means and standard deviations for 'Look first' false belief test (Siegal & Beattie, 1991), overall correct response score, correct evident response score and correct non-evident response scores for Study 3 for the middle and oldest groups collectively

Middle and Oldest groups	N	Mean	Std. Dev.
'Look first' FB Test Score	60	0.41	0.49
Overall Correct Response Score	60	5.51	2.65
Correct Evident Response Score	60	2.66	1.36
Correct Non-evident Response Score	60	2.85	1.51

A further examination of 'Look first' FB test responses and whether there were differences between the age groups in the number of children who got the correct answer was undertaken. Table 4.8 shows the correct and incorrect false belief test scores according to group.

Table 4.8: Correct and Incorrect 'Look first' False Belief test scores (Siegal & Beattie, 1991) for Study 3 for the middle and oldest groups

Middle and Oldest groups	N	Correct FB test	Incorrect FB test
Middle Group	30	7	23
Oldest Group	30	18	12
Total	60	25	35

Point-biserial correlation coefficient's for 'Look first' FB test (Siegal & Beattie, 1991) scores and overall correct response scores, correct evident and correct non-evident response scores found significant correlations ($r = .36$, $p < 0.01$, 2-tailed); (r

= .30, $p < 0.05$, 2-tailed); ($r = .35$, $p < 0.01$, 2-tailed) respectively. The relationship between the 'Look first' FB test (Siegal & Beattie, 1991) scores and overall correct response scores, correct evident and correct non-evident response scores signify modest linear relationships across the age range sampled.

A chi-square analysis (Yates' Correction for Continuity) computed on the data indicated that there was a tendency for the oldest group to give more correct answers on the 'Look first' FB test (Siegal & Beattie, 1991). This confirmed a significant association between age group and responses ($\chi^2 = 6.85$, df 1, $p < 0.05$).

4.3.5 Regression analyses

Regression analyses were conducted in order to assess the relative contribution of language level and conceptual grasp (via the 'Look first' FB test), (Siegal & Beattie, 1991) to children's understanding of references to evident and non-evident emotions.

Firstly, with overall correct response score as the dependent variable, and language test scores and false belief test scores as predictor variables, and using the SPSS 'forward entry' method, a significant model emerged ($F(2,57) = 14.50$, $p < .001$, Adjusted R square = .33). Both predictors were related to the dependent, language (beta = .46, $p < .001$) and false belief (beta = .27, $p < .05$).

Overall correct evident response score was then taken as the dependent variable and language test scores and false belief test scores were predictor variables. Using the 'forward entry' method a significant model emerged ($F (2,57) = 16.51$, $p < .001$, Adjusted R square = .34). Both predictors were related to the dependent, language (beta = .52, $p < .001$) and false belief (beta = .21, $p < .05$)

Overall correct non-evident response score was then taken as the dependent variable with language test scores and false belief test correct responses, as predictor variables. The 'forward entry method' was used and a significant model emerged ($F (2,57) = 15.80$, $p < .001$, Adjusted R square = .20). Both predictors were related to the dependent, language (beta = .33, $p < .05$) and false belief (beta = .29, $p < .05$).

In general, the Regression analyses indicate that overall correct responses, correct evident responses and correct non-evident responses are related to both language and false belief test performance. However, and despite the fact that there was little difference in the level of correct responses according to reference type, the regression analyses indicate that decoding of references to evident emotions was more strongly related to language, whereas for references to non-evident emotions there was more of a balance between the influence of language and that of conceptual grasp. This signifies that despite the similarities, the basis for understanding the two reference types may be somewhat different.

4.4. Discussion

Given the position of references to non-evident emotions in Ninio & Snow's (1996) developmental progression, it was hypothesised that children would show poor levels of understanding of these, and levels that were poorer than those found in Study 2 for references to shared past events. However, in fact, performance on the request task with respect to non-evident emotions was generally good, especially amongst the oldest age group, and no differences of any scale were apparent between grasp of references to evident and non-evident emotions.

The general ease of decoding both reference types was further indicated by the high levels of pragmatic sensitivity exhibited by children. There were few non-responses, with even incorrect responses showing a bias towards active pick-up of the experimenter's requests. There was some slight tendency for variation in response type according to request form, with direct requests somewhat more likely to elicit action responses, but overall, there was a marked bias towards action responses, suggesting that not only did children appreciate the need to respond to requests, they generally recognised that both direct and indirect requests required action responses i.e. they were capable of decoding the implicature involved in the indirect. In terms of the levels of pragmatic sensitivity outlined previously, this means that the majority of children were operating at the third level predominantly, regardless of reference type. On the argument that exhibited levels of pragmatic sensitivity are a reflection in part of task demands, the implication is that the request task in this version was relatively undemanding, and much less so than understanding of references to past events, where pragmatic sensitivity was found to be low. Moreover, this pattern was broadly the same for both age groups, though it was stronger for the older.

There are however signs that the observed levels of performance may have been influenced in part by task artefacts, and so may not give a true picture. One indication of this is the fact that there were age differences in identification of correct referents despite the lack of effect of reference type. The implication of this pattern is that children essentially found the two main task conditions (reference to evident and non-evident emotions) to be equivalent in difficulty. Thus the middle age group were performing at not much better than chance, but did so to much the same extent regardless of reference type. By the time children had reached 3.5 years, performance had improved – up to a level close to ceiling – but again regardless of reference type. Either children of these ages in general find references to non-evident emotions to be as easy to decode as references to evident – which seems implausible, and certainly at odds with the notion that references to non-evident states require additional layers of inference – or the way in which the task conditions were set up acted to make the decoding of these references easier than would usually be the case.

On reflection, there are two ways in which this might have been possible, and both may potentially have been a factor. The first is the possibility here of making a correct response by choosing either of two objects, where Studies 1 and 2 required children to choose one correct object out of four. In other words, chance responding was substantially more likely to lead to a correct response in Study 3. Two points make this less likely, in fact, though it cannot be ruled out. First, if chance responding were inflating children's scores, then this ought to be the case for both reference types, so any differential grasp of references to evident and non-evident emotions ought to be largely preserved, as long as performance is not at ceiling. This was not the case for the middle age group. Second, the whole notion of chance responding influencing performance is rather undermined by the grasp of which toys were happy – and sad – exhibited by children in response to the comprehension checks employed in the request task. This in turn points to where the task influence may really have lain – in the fairly elaborate attempt to make sure children did exhibit comprehension of the scenario's implications for the toys' emotions, which may well have had the effect of turning the non-evident emotion of sadness into an evident one, hence the equivalence in grasp of the two reference types. In terms of the data generated by the task, this is unfortunate, since it obscures to some extent the relative ease with which young children can decode references to non-evident emotions. At the same time, however, it perhaps serves to some extent to illustrate the kind of process by which children are provided with psychological support by parents in decoding references of this type. Indeed, the fact that such support for inferences could have the effect of reducing the task conditions to equivalence might be argued to be substantial direct proof of the power of this kind of mechanism – consistent with Ninio & Snow's (1996) position.

At the same time, there are signs that some difference in decoding of references to evident and non-evident emotions was maintained, regardless of the effects of this support – as ought in fact to be the case, given that even with support, decoding of references to non-evident emotions requires an inferential step that references to evident emotions do not, since the sources of the emotion are more readily 'read' directly off the situational characteristics as de Rosnay & Hughes (2006) maintain. Although the differences are not large, the fact that decoding of references to non-evident emotions was more strongly related to the conceptual grasp signalled by success on the 'Look first' FB test (Siegal & Beattie, 1991) and less strongly related to expressive language ability per se than was the case for evident emotion references is consistent with the maintenance of this inferential difference. Even with support then, the implication is that better conceptual grasp of internal hidden mental states was associated with greater likelihood of responding correctly to references to non-evident emotions.

The level of children's performance on the 'Look first' FB task in Study 3 is also consistent with the results from previous research in which the 'Look first' test was conducted (i.e. Siegal & Beattie, 1991; Surian & Leslie, 1999; Yadzi et al., 2006) where a majority of 3.5 to 4 year old children (60% in Study 3) gave a correct answer to the 'Look first' FB question. Some 2.5 to 3 year old children (24% in Study 3) did also exhibit success. With regression analyses establishing that request task performance, and the grasp of non-evident emotion in particular, was related to false belief test performance in Study 3 provides evidence that conversational support may have facilitating effects for understanding non-evident emotion. That conversational support by way of request task performance has such an impact on the understanding of non-evident belief (as Siegal & Beattie would argue), and where a more difficult process of inference may be required, will be addressed in Study 4.

The main themes that Study 4 will examine focus upon children's grasp of the fourth sub-maxim (Ninio & Snow, 1996) which concerns references to non-evident belief states. In a sense, all three tasks to this point have required some understanding of references, in as much as they share a common demand on the participating child to work out what the experimenter has had in mind beyond what is immediately evident in the present context. However, in the past event and non-evident emotion task, the actual reference to be decoded was not in fact to the experimenter's own state of mind, but effectively to something in somebody else's mind i.e. the child's in the case of the reference to a shared past event, and the toy characters' in the reference to non-evident emotion. For this reason, it was determined that the investigation of children's understanding of references to non-evident beliefs should continue to follow this same structure of reference to the mental state of a third character, paralleling the organisation of the emotion reference task (Study 3). The similarity between this task, and two forms of false belief tests (i.e. the unexpected contents task (UCT), (Perner, Leekam & Wimmer 1987; and the 'Look first' FB test, Siegal & Beattie, 1991) were also deliberate, since this permitted investigation of children's grasp of conversational references around events of the type employed in standard theory of mind tests. Comparisons between the two theory of mind tasks and Study 4 evident and non-evident belief references will also allow for an evaluation concerning the distinction between Generalised Conversational Implicatures (GCIs) and Particularised Conversational Implicatures (PCIs) (Grice, 1975; Levinson, 2000; Carston, 2004) raised in Chapter 1. Therefore, Study 4 was concerned with violations of references to non-evident internal belief states (i.e. those which require reasoning about others' interpretation of events). Children's ability to process language referring to a true or mistaken belief state was tested.

A combination of tasks will also make it possible to examine not only how far understanding of reference to non-evident beliefs improved with age, but also how far this improvement was predicted by language and working memory ability, and by conceptual elements. The issue of whether children show pragmatic sensitivity to reference type, reference form and experimenter requests in the concluding stage of Ninio & Snow's (1996) developmental progression in conversation remains prominent in Study 4.

CHAPTER 5

2:5 to 4:0 year old children's understanding of references to a non-evident belief state

5.1 Introduction and summary of conclusions to this point

5.1.1 Sequence of emergence

Whilst the sequence for references to non-present objects and shared past events is as predicted, going by Studies 1 and 2, the position with regard to non-evident emotions is unclear, because of the influence of task artefacts, as discussed in the final section of Chapter 4. What is apparent is that children in the middle age group performed at little better than chance level on the emotions task, regardless of reference type, but did show good signs of pragmatic sensitivity. The older age group did better on both counts, and the broad comparison to Study 2 on both dimensions would thus suggest that understanding of references to emotions precedes understanding of references to shared past events, counter to Ninio & Snow's (1996) sequencing of developmental progression in conversation. The difficulty, though, is that the impact of task support on the two dimensions of performance is unknown, so this may not be an accurate picture.

5.1.2 Pragmatic sensitivity

Putting this difficulty to one side for the moment, the key element of the above claims for the emergent sequence of growing competence found in Studies 1-3 derives from the apparent pattern with regard to pragmatic sensitivity. Contrary to what might be assumed pragmatic sensitivity is apparently not stable across the three studies and the different tasks they employ, but seems to vary as a function of level of understanding of task and reference type. These dimensions seem to be positively correlated, though far from perfectly, which may suggest a connection in terms of working memory that is consistent with the effects reported in Studies 1 and 2. One underlying factor in the improvements in children's performance with age may be that as understanding of references to non-evident objects or conditions improves, the processing demands of dealing with these reduce, allowing greater capacity for attending to the conversational implicatures in which the basic references are situated. This remains a matter for further research, however. For now, the hypothesised and confirmed covariation between understanding of reference type and pragmatic sensitivity makes it possible to consider the latter as a secondary measure of performance. On these grounds, pragmatic sensitivity has been taken to be an index of competence in given area of conversational ability allowing understanding of reference to emotion to be placed in the developmental sequence (Ninio & Snow, 1996), despite the reported lack of context effect.

This finding relating to pragmatic sensitivity is consistent with Siegal & Petersen's (1994) account of early difficulties with at least some forms of the false belief task:

where a grasp of the basic principle at work is uncertain, children trip up on the pragmatics of the test question, and get the answer wrong. When pragmatic demands are reduced they do better but performance is by no means 100%. When the basic grasp is better, the demands of the test question cease to be relevant: the pragmatics are understood, and do not impede performance. The general value of pragmatic sensitivity as an index is particularly apparent when considering children's performance on the past event and emotion tasks, since although there was some difference in performance, it was hard to differentiate solely in terms of understanding of the reference types involved especially in terms of underlying predictors of that performance.

5.1.3 Influences from changes in conceptual grasp

As well as working memory, another partial predictor of performance with regard to references to past events, was an age-related component which was loosely interpreted as being conceptual in character. As previously outlined no test of working memory was used in conjunction with the emotion reference task since it did not lend itself well to any form of probe that was task-related in the same way as that used for the past event task. It was therefore not possible to differentiate separate components of influence in the same way. However, understanding of emotion reference was predicted by age-related change in performance on Siegal & Beattie's (1991) pragmatically simplified 'Look first' false belief (FB) test, again indicating the impact of some underlying conceptual development.

5.1.4 Overall position

To summarise, performance on the tasks used in Studies 1 to 3 has proved to be assessable in terms of two dimensions, understanding of reference type and pragmatic sensitivity; on this evidence, the sequence of emergence of competence seems to be references to non-immediate objects, to non-evident emotions, and to shared past events; and improvements in understanding appear to be related to improvements in working memory and some underlying conceptual grasp, as well as to expressive language ability.

5.1.5 Rationale for Study 4

These points regarding the sequence of emergence and the underlying mechanisms at work set the scene for consideration of children's grasp of the fourth sub-maxim, concerning reference to non-evident belief states. In a sense, all three tasks to this point require some understanding of such reference, albeit in terms of perceptual content of different kinds, rather than belief states, in as much as they place a common demand on the participating child to work out what the experimenter has in mind beyond what is immediately evident in the present context. Moreover, in the past event and emotion tasks, the actual reference to be decoded was not in fact to the experimenter's own state of mind, but effectively to something in somebody else's mind i.e. the child's in the case of the reference to a shared past event, and the toy characters' in the reference to emotion. This of course parallels the structure of standard false belief tasks, where a scenario is described by the tester, and a test question is asked which requires inference of the mental state of one of the characters in that scenario; in other words, to respond appropriately, the children must understand the (embedded) reference made by the tester to the mental condition of that character. For this reason, it was decided that the investigation of children's understanding of references to non-evident beliefs should continue to follow this same structure of reference to the mental state of a third character, paralleling the organisation of the emotion reference task (Study 3). Study 4 also paralleled the false belief research discussed earlier in the literature review (i.e. Siegal & Beattie, 1991; Clements & Perner, 1994; Surian & Leslie, 1999; Yadzi *et al.*, 2006) in using simultaneous testing of both a true and false belief. For this purpose, a scenario was set up involving toy animals of the kind employed in the previous studies, in which two characters witnessed the swap in contents of a well-known type of container and therefore had a true belief about what it contained, based on line of sight evidence that the child saw them obtain; in this sense, these characters had an evident belief state.

A further two characters were subject to the same kind of swap in contents, but did not witness this. One might reasonably infer that their belief about what was in the container would therefore be based on its stereotypic contents; in this sense, though, it was not only likely to be false, but was also non-evident, since it could only be arrived at by a chain of inferences rather than by directly witnessing them acquiring pertinent information. Understanding checks were retained within the procedure as it was crucial to ascertain that children were following the basic conditions of the task. However, the focus of understanding checks for the non-evident beliefs was simply on the informational circumstances (i.e. their lack of witnessing the swap) rather than making any allusion to the inferences that might have been drawn from this bearing in mind the potential problems this created in Study 3. It is worth noting that similar understanding check questions are typically involved in false belief tasks and do not seem to promote better performance in that context e.g. Clements & Perner (1994) as previously discussed in the literature review (Chapter 1, page 20).

In order to test understanding of references to evident and non-evident beliefs, requests were made for these toys, couched in terms of beliefs about contents i.e. "Give me an animal who thinks it got X', where X is either a true belief not based on stereotypic contents item or a false belief based on a stereotypic contents item. The two different grammatical forms of request employed in the previous studies were retained i.e. direct requests and indirect requests as a continued means of assessing pragmatic sensitivity. As observations from Ninio & Snow's (1996) research had confirmed the later emergence of a grasp of this sub-maxim, the same two age groups were retained as in the emotion reference task, but a new sample of children were recruited.

The parallels between this task and the well-established unexpected contents task (UCT) (Perner, Leekam & Wimmer, 1987) were deliberate, since this permitted investigation of children's grasp of conversational references around events of the type employed in a standard theory of mind test. In addition to the reference task itself, though, data was also collected on a general measure of language ability; as before on memory for toy-food item associations in similar fashion to Study 2; and on two forms of standard theory of mind task, the UCT itself and the Siegal & Beattie (1991) ‘Look first’ FB test.

Comparisons between the two theory of mind tasks also allowed for an evaluation concerning the distinction between Generalised Conversational Implicatures (GCIs) and Particularised Conversational Implicatures (PCIs) (Grice, 1975; Levinson, 1983; Carston 2004) raised earlier in the literature review. To reiterate, PCIs are considered a more appropriate cross-referent than GCIs are, as PCIs provide a context-specific cue to aid children’s interpretation of a false belief question e.g. ‘where will a character look *first*’ vs. standard false belief test questions such as, ‘where will a character look’ which are regarded as GCIs. It is argued that GCIs are more difficult for children to decode (Noveck & Posada, 2003; Borg, 2005) as they must be deduced without the help of context and the knowledge and aims of the speaker which PCIs provide in conversational practices. This combination of tasks also allowed for the prediction that the ‘Look first’ FB test (Siegal & Beattie, 1991) would be better related to the understanding of reference to non-evident beliefs compared to the UCT (Perner, Leekam & Wimmer, 1987). Further evaluations included how understanding of references to non-evident belief improves with age, and is predicted by language and working memory ability, and by conceptual elements i.e. an age related component above and beyond the influence of language, working memory and false belief task performance, that might be attributed to some further conceptual shift.

5.1.6 Summary of task

Study 4 addressed the last aspect of the sequence of Ninio & Snow's (1996) developmental progression in conversation where adults continue to be co-operative in conversation by providing children with the psychological support necessary for understanding a final, but nonetheless crucial matter in their internal world i.e. non-evident belief. As with Study 3, children's knowledge of conversational references to their internal world was anticipated to have arisen during structured interaction with caregivers and the conversational practices they provide. Therefore, Study 4 examined the evidence on children's understanding of references to non-evident beliefs to ascertain whether two predicted corollaries of Ninio & Snow's (1996) account of this obtain: a) that grasp of such references will be late occurring in relation to other types of non-evident reference; and b) it will relate to performance on standard Theory of Mind (ToM) tasks since the necessary competences are held to have their origin in the conversational support of Ninio & Snow.

Study 4 was concerned with understanding of references to non-evident belief states (i.e. those which require reasoning about others' interpretation of events). Children in the age range of 2 years and 6 months to 3 years (i.e. middle age-group) and 3 years and 6 months to 4 years (i.e. oldest age-group) ability to process language referring to a true or mistaken belief state was tested. The scenario involved one toy animal (a Lion) giving four other toy animals in pairs (an Elephant and Frog and a Bear and Rabbit) going-home presents after a party. The former pair were to receive 'play-doh' containers with play-doh inside, and the latter pair were to receive pencil cases with a pencil inside. However, the Lion played a trick and replaced the contents i.e. play-doh was removed from both play-doh containers and crisps placed inside, and the pencils were removed from both pencil cases and chocolate mini eggs placed inside. Two animals e.g. the Elephant and Frog observed the Lion replacing the contents of their presents and this served to establish a true belief state not based on stereotypical contents. The two other animals e.g. the Bear and Rabbit did not witness the replacing of their presents and thus had a false belief state based on stereotypical contents. Children, having witnessed these events, were asked to provide the experimenter with different animals upon request, as in Studies 1 to 3, this time by referring to them via their implied belief states. A typical request was, "Give me the animal that thinks it got crisps" to refer to a true belief state or "Have you got an animal that thinks it got a pencil?" to refer to a false belief state. The requests sought to establish whether children could distinguish between different sources of belief and use these to decode references couched in those terms.

A language test (i.e. BAS II NVT, 1996) preceded the request task. Following the request task a short memory test took place with children asked to recall which animals received particular containers and presents. Both the 'Look first' FB test (Siegal & Beattie, 1991) and the UCT (Perner, Leekam & Wimmer, 1987) were utilised for Study 4. All children participated in both false belief tests. However, the conducting of the 'Look first' FB test was counterbalanced i.e. half of children from both groups took part in this test *first* i.e. prior to undertaking the request task. The other half of children undertook the 'Look first' FB test *second* i.e. following the request task/memory check. All children also took part in the UCT test. This test was not counterbalanced and was always conducted *second* i.e. following the request task/memory check. This was because of the similarities between the UCT and the request tasks, as described above in the rationale, and the requirement not to contaminate the request task itself.

5.1.7 Hypotheses for Study 4

The hypotheses addressed by Study 4 were as follows:

(1) That references to evident belief states will be understood better than references to non-evident belief states.

(2) The understanding of both references to evident and non-evident belief states will improve with age.

(3) The older group will still be some way from ceiling levels of performance with respect to references to non-evident beliefs, as indicated by both accuracy of response and degree of pragmatic sensitivity exhibited.

(4) The performance of both age groups will be below that found for children of these ages with regard to understanding of references to shared past events.

(5) Children's performance with regard to understanding of references to evident belief states will be related to language ability, but performance with respect to non-

evident beliefs will be additionally related to memory test score and performance on the false belief tasks.

(6) In terms of the relationship to false belief tasks, the 'Look first' FB test (Siegal & Beattie, (1991) will be more strongly related than the UCT (Perner, Leekam & Wimmer, 1987) because of the PCI/GCI distinction (Grice, 1975; Levinson, 1983; Carston, 2004).

5.2 Methodology

5.2.1 Design

Study 4 employed a request task of mixed between-and within-subjects design involving eight trials for each child participant. There was one between-subjects variable; age (either the middle or oldest age group). The within-subjects variables were reference type (evident vs. non-evident beliefs) and request form (direct vs. indirect requests). Children's responses were coded in terms of whether or not they were correct and also whether they were action, verbal or non-responses. This made it possible to examine degrees of incorrect responding as well as correct i.e. in terms of whether such responses were active or not. A Language test designed to measure variation in children's expressive language was conducted with each child during Study 4 testing. The test used was the British Ability Scales II, Naming Vocabulary Test (BAS II, 1996) for children in the age range 2 years to 4 years old. Children were also tested in a short memory task involving recall of associations established earlier during the request task. Two false belief tests were also conducted on each participating child i.e. the 'Look first' FB test utilised from Siegal & Beattie (1991) and the (UCT) devised by Perner, Leekam and Wimmer (1987).

5.2.2. Participants

Study 4 employed a new sample of children from those who had jointly participated in Study 3. The participants were 60 pre-school nursery children (27 girls, 33 boys), who had been attending either university ‘on campus’ nurseries in Glasgow or local authority/private nurseries in North Ayrshire. The sample was divided into 2 age groups, with 30 children in the middle group (12 females and 18 males; age range 2 years and 6 months to 3 years, mean age, 2 years and 9 months), and 30 children in the oldest group (15 females and 15 males; age range 3 years and 6 months to 4 years, mean age 3 years and 8 months). All children spoke English as the language of their home. Relevant local authority clearance and university ethical approval was obtained. Parental consent was given with respect to all participating children and children were advised that they could withdraw from the study at any time.

5.2.3 Materials

A large green cloth mat was also used to conduct the test sessions upon nursery floors. The main element of materials consisted of five toy collections. The first, second and third toy collections were items relating to the Study 4 story and questions. The first was a collection of soft toy animals. The animals consisted of a *Bear, Elephant, Frog, Lion and Rabbit.* No animal was smaller than 15 cm or higher than 18 cm. The second toy collection consisted of a toy *wall* that had been adapted for the study. A rectangular piece of plastic measuring 50 cm in length, 5 cm in width and which was 20 cm high when fixed upright by a support had been covered in brick-effect paper of a standard type that can be purchased in model shops and used to cover toy buildings such as dolls' houses or train set accessories. The dimension of the toy wall allowed children, when seated on the green mat, to have an unobstructed, 'bird's eye view' of events taking place concerning the animals and other materials on either side of the wall. The third toy collection consisted of, 2 play-doh containers and 2 play-doh materials; 2 pencil cases and two pencils; 2 small transparent bags containing 2 lots of potato crisps; two other small transparent bags containing two lots of chocolate mini eggs: these latter food items were also of a size that had been chosen to fit neatly into the play-doh and pencil case containers when their usual contents were removed.

Verification of children's unobstructed view of events when the wall was positioned, and their understanding of and naming of the materials took place in a pilot test at a different nursery one week prior to testing. This involved a pilot sample of 14 children in the age range of 2 years and 6 months to 4 years old. For each of the 14 children, the experimenter positioned the wall so that children could see events taking place on either side of it. The experimenter then pointed to the wall and asked each child, 'What is this? All children confirmed either 'a wall' or a 'brick wall'. To confirm children could recognise items, five animals, one play-doh container and one play-doh material; one pencil case and one pencil; one set of potato crisps and one set of chocolate mini eggs were in full view of children. Each child was asked to name each of the soft toy animals, the containers, their visible contents and the actual food items i.e. the experimenter pointed to each of these items in turn and asked, 'What is this?' Each child correctly named the Bear, Elephant, Frog, Lion and Rabbit; the play-doh container and the play-doh; the pencil case and a pencil; the potato crisps and the chocolate mini eggs. Naming of animals and materials also served to confirm that children had indeed a ‘bird’s eye view’ of events taking place on both sides of the wall. As the consequence of children successfully completing the pilot test, the experimenter then made two marks on the green experimental mat to assist in subsequent testing proper at a later date. One mark, a line was drawn that indicated where the wall had been suitably placed, and another mark an X, identified where children had been sitting on the mat that had accorded them the best view of events taking place on either side of it.

The fourth toy collection related to items required for the ‘Look first’ FB test (Siegal & Beattie, 1991). This toy collection consisted of a *large* soft toy *Cat* measuring 25 cm and a *small* soft toy *Cat* measuring 15 cm; two toy cupboards, identical in dimension, 12 cm x 7cm x 11cm, each with a sliding drawer, one coloured blue and the other coloured green; one toy food, Milk Chocolate that fitted neatly into the cupboard’s drawers when these were closed. The fifth toy collection related to items required for the unexpected contents test (Perner, Leekam and Wimmer, 1987). This toy collection consisted of a large Smarties tube, measuring 18 cm in length and two crayons.

The experimenter scored children’s responses during the test sessions using the pro-formas shown in appendices G & H. A video camera on tripod was utilised to record children’s responses, thus enabling the coding of responses to be re-checked by the experimenter and then verified by an independent judge at a later date.

Permission to use and record children's responses by video, in addition to their participation approval, was sought and granted from participating nurseries and parents, who were both informed that the video footage would be erased once the project was complete.

5.2.4 Procedure

Each nursery was visited on separate occasions over a 9 month period. Testing took place within nurseries in 4 stages.

<u>Stage 1:</u> Warm-up period, Language Testing.

The procedure for the warm-up period and language testing i.e. the British Ability Scales II Naming Vocabulary Test (BAS II NVT) (1996) was the same as for Study 1 and standard procedures were used for this as described in sub-section 2.2.4. As with the previous studies the video recorder was switched on after each child's warm up period had ended.

<u>Stage 1a</u>: Counterbalanced 'Look first' FB test.

Half of participating children from both age groups took part in the 'look-first' false belief test (Siegal & Beattie, 1991) which was conducted in counterbalanced order with Stage 4. See Stage 4 (below) for full details of both false belief testing and procedures.

Stage 2: Tests of understanding of references to evident and non-evident beliefs

When Stage 1 i.e. the British Ability Scales II Naming Vocabulary Test (BAS II) (1996) and Stage 1a (if applicable) was completed children were shown the toy animals in their paired friendships. This involved the experimenter holding up one pair of animals to children in turn and stating, 'this is Elephant and Frog and they are best friends with each other'. The same procedure then applied to the other pair i.e. the Bear and Rabbit. Therefore, children viewed both the animal pairs but which pair was designated as targets for references to evident and non-evident beliefs was counterbalanced, as was the order in which the evident and non-evident belief targets were presented. One soft toy animal, the Lion*,* was produced later during the procedure to children and was not allocated a pairing.

At each participating nursery half the children of each age group were allocated to Set G (see Table 5.1 below) where the Elephant and Frog paired friendship represented the evident belief scenario (conducted first) and the Bear and Rabbit paired friendship represented the non-evident belief scenario (conducted second). The other half of children were allocated to Set H (see Table 5.2 below) where for counterbalancing purposes the Elephant and Frog paired friendship represented the non-evident belief scenario (conducted first) and the Bear and Rabbit now represented the evident belief scenario (conducted second).

For the purposes of keeping the following outline succinct, procedure from this point forward will be described as if for child who received the Elephant and Frog paired friendship/evident belief scenario first and the Bear and the Rabbit paired friendship/non-evident belief scenario second. The subsequent words of the experimenter are in inverted commas and italics, with numbered points and procedure in standard font.

The experimenter said to children, *"This is a story about Lion's Party".*

'Lions Party'

<u>Stage 2a:</u> Evident belief scenario (first).

'True belief not based on stereotypical contents'

(1) Each child is seated on the green mat and the toy wall positioned according to marking as defined in pilot testing.

(2) The child is told, *"Lion has had a party. It is the end of the party and Lion has a 'going away' present for each of the four animals that came to the party in their pairs i.e. Elephant & Frog, Bear & Rabbit".*

(3) The experimenter placed the presents on the child's right hand side of the dividing wall. The Lion, Elephant and Frog and the Bear and Rabbit were then placed on the child's left hand side of the dividing wall. Children had a clear view of presents and animals on either side of the wall.

(4) The experimenter points to the presents collectively i.e. 2 x tubs of play-doh, 2 x play-doh materials inside; 2 x pencil cases, 2 pencils.

(5) "*Lion then says to Elephant & Frog, I am away to get your presents, wait here, behind the wall, until I return with them".*

(6) Lion goes behind the wall.

(7) The child is told, *"Lion wants to play a <u>trick</u> on Elephant & Frog and <u>change</u> the contents of their presents".*

(8) At this point the experimenter moves the Elephant & Frog behind the wall and makes the following statements the child.

(9) *"Elephant & Frog cannot wait to see what presents Lion has for them and they have stepped behind the wall".*

(10) The experimenter emphasizes Elephant & Frog's secret plan by positioning the animals at the top the wall in a location where the child participant can clearly see the Elephant & Frog observing the Lion's changing the contents of the presents on the other side.

(11) <u>Understanding check 1: Evident belief</u>

The experimenter shows the child the play-doh tubs and says,

"These are what Lion is going to give to Elephant & Frog".

"What are these"?

The anticipated answer was that children would state, 'play-doh'. All participating children answered this way.

(12) The experimenter then says to the child,

"*Elephant and Frog can see the Lion taking the play-doh out of their tubs and putting crisps in instead".*

(13) The tangible play-doh material itself is discarded. It plays no further part in procedure. Crisps are placed inside the Elephant and Frog's play-doh containers.

(14) The child is told,

"Look, Elephant and Frog are moving back to the other side of the wall".

(15) Lion gives Elephant and Frog their presents.

(16) <u>Understanding check 2: Evident belief</u>

(a) *What containers did Elephant and Frog get given*?

(b) What is inside Elephant & Frog's containers?

(c) *Did Elephant and Frog see play-doh being changed for crisps?*

(17) Elephant & Frog and their presents are moved to a neutral area on the mat.

Twenty-seven out of thirty middle group children (90%) and the entire oldest group (100%) gave a response to the understanding checks that indicated that they were keeping track of events. Two middle group children gave incorrect responses at understanding check 2(b) stating 'play-doh'. However, these two children subsequently responded correctly to understanding check 2(c) which is regarded as the most crucial question since it establishes the evident belief, by making the content and the characters' knowledge of the content explicit. Therefore, these children had at least a partial understanding of the grounding scenario and were allowed to continue. The other middle group child failed to give any answer at understanding check 2b. This child became distressed at this point and was excluded from the test: another middle age group child who understood the checks participated as a replacement.

<u>STAGE 2b:</u> Non-evident belief scenario (second).
'False belief based on stereotypical contents'.

(1) *"Lion now approaches the other pair of animals, Bear and Rabbit*".

(2) *"Lion has also got 'going away' presents for the Bear & Rabbit"*.

(3) The presents were positioned on the child's right hand side of the dividing wall. The Bear and Rabbit were in place on the child's left hand side of the dividing wall. Children had a clear view of presents and animals on either side of the wall.

(4) The experimenter pointed to the presents collectively i.e. 2 x pencil cases; 2 x pencils.

(5) Lion then says to the Bear and Rabbit.

*"*I am away to get your presents, wait here, behind the wall, until I return with them".

(6) Lion goes behind the 'wall'.

(7) The child is told, "*Lion wants to play a trick on Bear & Rabbit and change the contents of their presents".*

(8) At this point the experimenter emphasises that the Bear & Rabbit stay behind the wall by making the following statements to the child.

(9) "*Bear and Rabbit are staying behind the wall, where they cannot see what presents Lion has for them".*

(10) The experimenter emphasizes Bear and Rabbit's *obedience* by positioning the animals close to the middle and left hand side of side of the wall i.e. in a location where the child participant can clearly see that the Bear and Rabbit are unable to observe the Lion's changing the contents of the presents.

(11) <u>Understanding check 3: Non-evident belief</u>

The experimenter showed the child the pencil cases and pencils and said,

"These are what Lion is going to give to Bear and Rabbit"

"What are these"?

The anticipated answer was that children would state, 'pencil cases and pencils'. All participating children answered this way.

(12) The experimenter then says to the child.

"*Bear and Rabbit cannot see the Lion taking the pencils out of the pencil cases and putting chocolate mini eggs in instead".*

(13) The tangible pencils are discarded. They play no further part in procedure. Chocolate mini eggs are placed inside Bear and Rabbit's pencil cases.

(14) Lion returns from behind the wall.

(15) Lion gives Bear and Rabbit their presents.

(16) <u>Understanding check 4: Non-evident belief.</u>

(a) What containers did Bear and Rabbit get given?

(b) What is inside Bear and Rabbit's containers?

(c) Did Bear and Rabbit see pencils being changed for Chocolate mini eggs?

In contrast to understanding check 2(c)-evident belief, this question (i.e. understanding check 4(c)-non-evident belief) leaves the contents of the character's belief about contents implicit.

(17) Bear and Rabbit and their presents are moved to the neutral area on the mat beside Elephant and Frog and their presents.

(18) All presents were then removed from proceedings and the four animals were placed directly in a row in front of the participating child. Positioning of the four animals, from left to right animals was varied systematically for each incoming child.

Twenty-six out of thirty middle-aged group children (86%) and twenty-nine out of thirty oldest children gave responses to understanding checks 3 & 4-non-evident belief that indicated that they were keeping track of events. Four middle group children and one oldest group child gave incorrect responses at understanding check 4(b). However, all of these children subsequently responded correctly to understanding 4(c). Children who showed problems here were not the same as those children who had difficulty with the evident belief scenario and therefore showed sufficient understanding of the scenarios overall for their data not to be discounted.

For the purposes of counterbalancing half of the middle and oldest age groups of children were allocated to Set H (see Table 5.2 below) and saw the non-evident belief scenario first and evident belief scenario second. However, the order in which the animal pairs were employed remained constant. This also ensured the counterbalancing of swapped presents and stereotypical contents. Therefore, in contrast to children allocated to Set G (described in Procedure above), a participating child allocated to Set H would be led to think that the Elephant and Frog were expecting play-doh inside their play-doh container i.e. a false belief based on stereotypical contents. Furthermore, children allocated to Set H would also understand that the Bear and Rabbit expected chocolate mini eggs in their pencil case and therefore had a true belief based on witnessing the swap from stereotypical contents.

As was the case for Study 1 to 3 procedures, children's ability to decode references, in this instance to an evident belief and a non-evident belief state, was then tested by examining their responses to a series of requests made by the experimenter using two different request forms. These were a direct request ("Give me the animal that thinks it got X", where X refers to a toy food or item), and an indirect request ("Have you got the animal that thinks it got X?"). Children were told that the experimenter wished to ask them some questions about the story and the toy animals. The questions (see Tables 5.1 and 5.2) were put to children in the manner described in Studies 1 to 3. However, the Study 4 task paralleled that of Study 3 in the sense that it allowed either animal from a pair to be identified in response to each of the experimenter's requests. Correct responses still required the accurate identification of the referents and though consistent with Study 3 it was felt that this created no material difference. Children's responses were recorded at the time by the experimenter on either the Appendix G or H proforma, and on videotape. At the end of trials children were given a happy face sticker and thanked for participating.

Table 5.1: Order of request by reference type and request form for Set G children.

REQUEST-OBJECT	REFERENCE TYPE	REQUEST FORM
1. Give me an animal that thinks it got crisps. (Elephant or Frog)	Evident belief	Direct request
2. Have you got an animal that thinks it got a pencil? (Bear or Rabbit)	Non-evident belief	Indirect request
3. Have you got an animal that thinks it got crisps? (Elephant or Frog)	Evident belief	Indirect request
4. Give me an animal that thinks it got crisps. (Elephant or Frog)	Evident belief	Direct request
5. Give me an animal that thinks it got a pencil (Bear or Rabbit)	Non-evident belief	Direct request
6. Have you got an animal that thinks it got crisps? (Elephant or Frog)	Evident belief	Indirect request
7. Give me an animal that thinks it got a pencil. (Bear or Rabbit)	Non-evident belief	Direct request
8 Have you got an animal that thinks it got a pencil? (Bear or Rabbit)	Non-evident belief	Indirect request

Table 5.2: Order of request by reference type and request form for Set H children

REQUEST-OBJECT	REFERENCE TYPE	REFERENCE FORM
1. Have you got an animal that thinks it got play-doh? (Elephant or Frog)	Non-evident belief	Indirect request
2. Give me an animal that thinks it got mini eggs. (Bear or Rabbit)	Evident belief	Direct request
3. Give me an animal that thinks it got play-doh? (Elephant or Frog)	Non-evident belief	Direct request
4. Have you got an animal that thinks it got mini eggs? (Bear or Rabbit)	Evident belief	Indirect request
5. Have you got an animal that thinks it got mini eggs? (Bear or rabbit)	Evident belief	Indirect request
6. Give me an animal that thinks it got mini eggs? (Elephant or Frog)	Evident belief	Direct request
7. Have you got an animal that thinks it got play-doh? (Elephant or frog	Non-evident belief	Indirect request
8. Give me an animal that thinks it got play-doh (Elephant or Frog)	Non-evident belief	Direct request

Stage 3: Memory check

Following the request task a memory check took place. The Elephant, Frog, Bear and Rabbit were placed beside one another. The play-doh containers and the pencil cases were positioned away from the animals with their lids and zips open respectively: their contents were clearly visible to children i.e. the play-doh containers had crisps on view inside and the pencil cases had chocolate mini eggs on view inside.

The child was asked to place containers with contents visible in front of their respective animals who got them as presents i.e. the experimenter said each of the following questions two times, '*Which animal got crisps to eat?' 'Which animal got chocolate mini eggs to eat?'*

Stage 4: False belief tests

All 60 children participated in the two false belief tests i.e. the 'look-first' false belief test, Siegal & Beattie (1991) and the (UCT) (Perner, Leekam & Wimmer, 1987).

However, the order in which children undertook the 'Look first' FB test was counterbalanced i.e. half of participating children from both groups i.e. 30 children had undertaken the 'Look first' FB test first i.e. before the request task prior to Stage 2. The other 30 participating children from both groups took part in the ‘look-first’ test at this point in Stage 4. The ‘look-first’ test Siegal & Beattie (1991) was conducted to establish the relationship between the understanding of false belief and references to non-evident belief states. The test was administered using standard procedures as described for Study 3.

All 60 participating children in both groups then took part in the UCT (Perner, Leekam & Wimmer, 1987) at this point in Stage 4. The UCT was not counterbalanced due to its similarities with the request task and the need not to contaminate the latter task. This false belief test was also conducted in order to establish the relationship between the understanding of false belief and references to non-evident belief states. It was administered according to the following UCT standard procedures as described by Perner, Leekam & Wimmer (1987).

Children were shown a ‘Smarties’ tube and asked, *‘What do you think is inside the tube?”* The anticipated answer from children was either ‘sweets’ or ‘smarties’. The experimenter then opened the tube and revealed to children that the contents were in fact two crayons. Children were then asked, *‘What do you think that the next child* [*waiting outside*] *will think is in this tube?*

5.2.5 Coding of responses and scoring

Request task

The experimenter observed and recorded children's responses during testing and these responses were also later rechecked and transcribed from videotape after data collection was complete. Children's responses to each request were coded using the same system as described in Study 3. The incorrect action response (gives different belief object or GDBO) was devised specifically for the coding needs of Study 4 in order to identify when a child gave the wrong belief state object. To check on the reliability of coding, the responses of five children from the two age groups were selected at random for coding from the video footage by an independent judge. The 80 responses given by these children accounted for 17% of the total. Agreement rates for each code are given below.

1. Correct action response = 93%
2. Correct verbal/gestural response = 93%
3. Incorrect action response (gives different belief object) = 100%
4. Incorrect/uninformative verbal response = 93%
7. Incomprehension response = 100%
8. Non-response = 97%

As in Studies 1 to 3 the number of children's responses in each of the six categories was totalled across trials of the same reference type and request form to give a score between 0 and 2 for each. Also, to check on order effect, the number of children's responses in each of the six categories was totalled separately across the first and second questions employing the two different request forms within each reference type to give a score between 0 and 2 in each case.

5.2.6 Language test scoring.

In the BAS II NVT (1996) children responses were scored in the standard fashion as described in sub-section 4.2.6.

5.2.7 Memory check scoring

Children received one point for correctly assigning a container with food inside to the appropriate animal. There were no points for an incorrect assignment.

5.2.8 'Look first' false belief test scoring

Scoring for 'Look first' FB test (Siegal & Beattie, 1991) remained the same as for that of Study 3 as described in sub-section 4.2.7.

5.2.9 Unexpected contents test scoring

Children giving the answer 'smarties' or 'sweets' in the UCT (Wimmer, Leekam & Perner, 1987) received one point as this indicated that they could represent another child’s (who had not witnessed the unexpected contents) belief that the tube would contain stereotypical contents. Children answering ‘crayons’ or ‘pencils’ received 0 points as this indicated they had not represented the next child’s belief state of stereotypical contents.

5.3 Results Section

5.3.1 Summary of Analyses

Three preliminary analyses were conducted on the request task data: firstly, to check for the effects of question order (i.e. whether there was any indication of practice effects); secondly, to check whether there was any effect of belief order by way of primacy or recency effects (i.e. as children had received either Set G or H for counterbalancing purposes) and thirdly, whether there was any effect of question set/sequence i.e. which set the children were in, and thus which sequence of questions they received for the request task. The results of these analyses were as follows. A 3-way mixed ANOVA (age group x reference type x question order) was performed on the scores for the evident belief first questions, evident belief second questions, non-evident belief first questions and non-evident belief second questions for each of the six response types outlined in 5.2.5. No significant effects or interactions were found involving question order. A 4-way mixed ANOVA (age group x belief order x reference type x request form) was then conducted on the scores of evident belief/direct, evident belief/indirect, non-evident belief/direct and non-evident belief/indirect by the six response types outlined in 5.2.5. No significant effects or interactions were found in relation to the belief order used. A 3-way mixed ANOVA (age group x set x reference type) was conducted on the scores of evident belief/direct, evident belief/indirect, non-evident belief/direct and non-evident belief/indirect by the six response types outlined in 5.2.5. No significant effects or interactions were found in relation to the set/question sequence set used. These factors will not be considered further.

The main analyses were then computed: 3-way mixed ANOVA's (age group x reference type x reference form) were conducted on scores for evident belief/direct, evident belief/indirect, non-evident belief/direct and non-evident belief/indirect for each of the six response types.

A 3-way mixed ANOVA was also used to examine correct action responses vs. correct verbal/gestural responses (age group x reference type x response type). This relationship was then further explored using correlational analyses. As with the previous studies, a 2-way ANOVA was then computed on the overall level of correct responses collapsed across correct action and correct verbal/gestural responses. A 3-way mixed ANOVA was conducted to examine active vs. inactive responses (age group x reference type x active vs. inactive). The results of these analyses are reported in this order in what follows.

With the main patterns of performance on the request task identified, the relationship of correct responses to language and memory test scores was explored using correlational analyses. Prior to the correlational analyses of correct responses and memory test scores, a One-way ANOVA was also computed to examine age group and memory test performance. The relationship between age group and the two false belief tasks i.e. the UCT (Perner, Leekam & Wimmer, 1987) and the 'Look first' FB test (Siegal & Beattie, 1991) was examined using the Fisher exact test. One-way ANOVA's were also conducted separately on language, memory and request task in respect of the UCT and the 'Look first' FB test. Regression analyses were then employed to further examine a potential relationship between the 'Look first' FB test and the predictors of age, language, memory and request task.

5.3.2 Descriptive and Inferential Results for Request task

Table 5.3 shows the mean scores and standard deviations for the response types used by children according to age group, reference type and request form. Appendix Q presents full results for the 3-way ANOVAs conducted on these data.

TABLE 5.3: Means and standard deviations for each response code (range = 0-2) by reference type, age group and request form

			Correct action resp		Correct verbal resp		Incorrect action resp		Incorrect verbal resp		Non resp	
Ref type	Age group		Dir	Indir	Dir	Indir	Dir	Indir	Dir	Indir	Dir	Indir
Evident Belief	Middle	Mean	0.13	0.20	0.93	1.00	0.63	0.50	0.07	0.07	0.23	0.23
		SD	*0.43*	*0.48*	*0.83*	*0.87*	*0.76*	*0.63*	*0.25*	*0.36*	*0.63*	*0.63*
	Oldest	Mean	0.70	0.57	0.83	1.00	0.43	0.43	0.00	0.00	0.00	0.00
		SD	*0.88*	*0.82*	*0.83*	*0.91*	*0.63*	*0.63*	*0.00*	*0.00*	*0.00*	*0.00*
	Total	Mean	0.42	0.38	0.88	1.00	0.53	0.47	0.03	0.03	0.12	0.12
		SD	*0.74*	*0.69*	*0.82*	*0.88*	*0.70*	*0.62*	*0.18*	*0.26*	*0.45*	*0.45*
Non-Evident Belief	Middle	Mean	0.17	0.13	0.97	0.90	0.53	0.60	0.03	0.07	0.30	0.33
		SD	*0.46*	*0.43*	*0.89*	*0.76*	*0.78*	*0.72*	*0.18*	*0.25*	*0.70*	*0.71*
	Oldest	Mean	0.60	0.47	0.97	1.17	0.43	0.40	0.00	0.00	0.00	0.00
		SD	*0.81*	*0.73*	*0.85*	*0.75*	*0.68*	*0.56*	*0.00*	*0.00*	*0.00*	*0.00*
	Total	Mean	0.38	0.30	0.97	1.03	0.48	0.50	0.02	0.03	0.15	0.17
		SD	*0.69*	*0.62*	*0.86*	*0.76*	*0.72*	*0.65*	*0.13*	*0.18*	*0.51*	*0.53*

Ref type = Reference type = resp = response type; Dir = direct request form and Indir = indirect request form. Incomprehension responses' were never given by children and are therefore excluded from this table.

1. Correct action responses

There was a main effect of age group ($F(1,58) = 10.07, p <. 05$). Older children gave more correct action responses than the middle age group did, but there was no effect of reference type or request form, though there was a marginal tendency to give these responses less often to indirect requests. There were no further significant effects involving correct action responses.

2. Correct verbal/gestural responses

There were no significant effects involving correct verbal/gestural responses. Both age groups gave these responses with approximately equal frequency, regardless of reference type, and though older children gave these slightly more often to indirect requests (where correct action responses were fewer), the interaction was not significant.

3. Comparison of correct response forms

As Table 5.3 shows, correct verbal/gestural responses were given more often than correct action responses, especially by middle age group where they were approximately five times as frequent. This effect was tested by a 3-way mixed ANOVA (age group x reference type x response type) and found to be highly significant ($F(1,58) = 22.10, p < .001$) and interacted with neither age group nor reference type.

Both age groups appeared to default to correct verbal/gestural responses regardless of reference type and request form, though in fact the older outperformed the middle primarily in terms of correct action responses: it was these where the difference between the two age groups lay mostly. On the face of it, it appears curious that the older age group should show this almost exclusive increase in correct action responses, regardless of any obvious task characteristic. In fact, closer inspection indicates that the mean values are misleading in implying a greater response mix as children became more adept. These appear actually to be different individual response styles, given the high standard deviations and relatively sizeable negative correlations between correct actions responses and correct verbal/gestural responses ($r = -.56$, $p < .001$, 1- tailed, for the evident context; $r = -.36$, $p = .002$ 1- tailed) for the non-evident context). These signify that if children gave correct action responses they did not tend to give correct verbal/gestural responses and vice versa, producing high variance in the scores for each measure. The difference between age groups would therefore seem to be essentially one of an increase in those who were prepared to give correct action responses or at least in those who were correct when they did so.

4. Correct responses

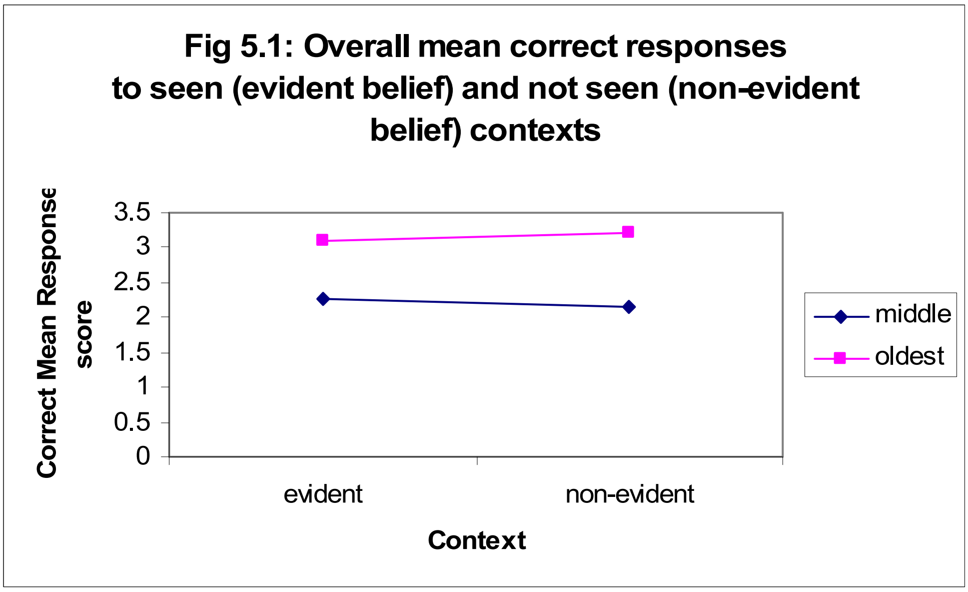
Table 5.4 shows the total number of correct responses by children in each age group.

Table 5.4: Total correct response scores by age group for seen (evident belief) and not seen (non-evident belief references) with mean scores and st. dev's

Age Group	Total correct responses: Seen (evident belief)	Mean correct responses: Seen (evident belief)	*St. Dev*	Total correct responses: Not seen (non-evident belief)	Mean correct Responses: Not seen (Non-evident belief)	*St. Dev*
Middle	68	2.27	*1.46*	65	2.16	*1.57*
Oldest	93	3.10	*1.06*	96	3.20	*1.24*
Total	161	2.69	*1.33*	161	2.68	*1.50*

In total there were 322 correct response types given out of a possible 480. Total correct response types made up 67% of overall responses.

In relation to correct responses there was a significant main effect of age group ($F(1,58) = 12.86$, $p = .001$), but no other effects were found. In general, then, children improved with age in their ability to decode references to belief states, those in the middle age group performing at roughly chance level (one response out of two correct, with two types of toy items to choose from) but those in the older group performing above chance, though not apparently approaching ceiling. There was no effect of reference type and none of request form. Figure 5.1 shows these trends.



4. Incorrect action responses

There were no significant effects for incorrect action responses, though the frequency of these responses was slightly less amongst the older age group, reflecting their higher number of correct responses.

5. Incorrect verbal responses

These were only given at all by the middle age group, and even then in such small numbers that there was no significant effect of age group, or any other factor.

6. Non-responses

Similarly, these were only given by the middle age group, but in this case in rather greater number, producing a main effect of age group ($F(1,58) = 5.89$, $p < .05$). There were no other effects.

7. Active incorrect vs. inactive incorrect responses

Table 5.5 compares the incidence of active incorrect responses (i.e. incorrect action responses and incorrect verbal responses) and inactive incorrect (i.e. non-responses) to the experimenter's requests, by age group and reference type.

Table 5.5: Mean number of active incorrect responses and inactive incorrect responses by age group and response type

Age Group	Active Incorrect seen	Active Incorrect not seen	Inactive Incorrect seen	Inactive Incorrect not seen
Middle	1.27	1.23	0.47	0.63
SD	*1.28*	*1.33*	*1.22*	*1.40*
Oldest	0.87	0.83	0.00	0.00
SD	*1.01*	*1.05*	*0.00*	*0.00*
Total	1.07	1.03	0.23	0.32
SD	*1.16*	*1.21*	*0.89*	*1.03*

A three-way ANOVA (age group x reference type x active vs. inactive responses) was conducted (see Appendix R) and found a highly significant main effect of active incorrect vs. inactive incorrect responses ($F(1,58) = 15.91$, $p < .001$), and of age group ($F(1,58) = 13.57$, $p < .001$). There were no other effects.

It would appear that older children made fewer incorrect responses overall, as would be expected, regardless of reference type. There were, as already noted, no effects of request form, but active incorrect responses were predominant over inactive, indicating that there was at least this degree of pragmatic sensitivity. Curiously, though, action responses were much more frequent than verbal responses when answers were incorrect, in contrast to correct responses, where the pattern was the exact opposite. One construction that might be put on this is that verbal responses were simply more likely to be judged correct than action responses, especially amongst the middle age group. However, bearing in mind the point already noted about apparent differences in response style, an alternative interpretation is that those who chose to use correct verbal/gestural responses tended to have a better grasp of the task (i.e. of references of this kind) than those who made correct action responses. It may also be that a further factor is influencing children's higher use of correct verbal/gestural responses in Study 4 i.e. this response type is a more polite way for children to conduct themselves due the experimenter's lack of psychological support in the formal setting of a request task and this argument will be taken up in the discussion section.

5.3.3 Effects of language

Table 5.6 shows the means and standard deviations in relation to language scores and correct responses.

Table 5.6: Means and standard deviations for British Ability Scales II Naming Vocabulary Language Test and correct response scores for Study 4 for the middle and oldest groups

Middle and Oldest groups	N	Mean	Std. Dev.
BAS II NVT Score	60	15.55	0.87
Correct Response Score	60	5.40	2.35

The number of total number of correct responses and BAS language ability scores were found to be marginally correlated (r = .22, p = .046, one-tailed).

5.3.4 Memory test performance

Table 5.7 indicates memory test performance according to groups.

Table 5.7: Memory test performance by groups (range = 0-4)

Correct Memory Score	N	Mean	Std. Dev.
Middle Group	30	2.73	1.23
Oldest Group	30	3.47	0.90

Performance on the memory test was generally good, but not perfect, with older children doing significantly better ($F(1,58) = 6.95$, $p = .011$) than those in the middle age group. Given that this was essentially a test of working memory in the sense of keeping track of associations, this raised the possibility that it might have impacted on performance in a manner similar to that observed in Studies 1 and 2. In fact, memory test performance and overall number of correct responses were correlated somewhat more strongly than was the case for language ability ($r = .36$, $p = .002$, one-tailed), though the relationship was again not a sizeable one.

5.3.5 Theory of mind test performance

As the 'Look first' FB test (Siegal & Beattie, 1991) was conducted in counterbalanced order with half of each of the middle and oldest age groups completing the test prior to the request task (first), and the other half of both age groups undertaking the test following the request task (last) analyses were conducted to check for differences in performance. Independent-Samples t-tests confirmed there were no differences in performance for those in either age groups who undertook the ‘Look first’ FB test first or last.

Table 5.8 shows children’s performance on the two theory of mind tasks, broken down by age group, together with the joint contingency between performance on the two tasks.

Table 5.8: Performance on the unexpected contents test (Perner, Leekam & Wimmer, 1987) and ‘Look first’ false belief (FB) test (Siegal & Beattie, 1991).

	Unexpected contents test		**‘Look first’ FB test**		**Joint contingency**		**Unexpected contents test**	
Age Group	Incorrect	Correct	Incorrect	Correct			Incorrect	Correct
Middle	17	13	24	6	**‘Look first’ FB test**	**Incorrect**	25	12
Oldest	19	11	13	17		**Correct**	11	12

As can be seen, there was little difference between the two age groups on the unexpected contents task (UCT) (Perner, Leekam & Wimmer, 1987) with the preponderance of responses incorrect, and the older age group doing marginally worse than the middle. There was no association between age group and performance, nor between performance on this task and the ‘Look first’ FB test (Siegal & Beattie, 1991) (Fisher exact ns in both cases). However, there was a clear trend towards older children doing better on the ‘Look first’ FB test (Fisher exact $p = .004$, one-sided). In general, it would appear that the UCT provided a poor measure in this instance, perhaps because its very similarity with the reference task that it always followed causing confusion. A further explanation points simply to one of fatigue, given that it was always conducted at the end of a long session of testing.

Unsurprisingly, then, performance on the UCT (Perner, Leekam & Wimmer, 1987) was found to be unrelated to language score, contrary to consistent evidence in the literature. The 'Look first' FB test (Siegal & Beattie, 1991) was associated with language in the usual direction, though, with those who passed the task having higher language scores than those who failed it (mean = 15.83 vs. 15.38). However, this did not quite meet conventional significance levels at ($F(1,58) = 3.92$, $p = .052$). Similarly, performance on the UCT was found to be unrelated to score on the memory test, but there was a highly significant relationship between performance on the 'Look first' FB test and memory score ('Look first' FB test pass mean = 3.65, fail mean = 2.76; ($F(1,58) = 10.30$, $p < .001$). The latter relationship, whilst of interest in itself for confirming the relative status of the two theory of mind tasks, is of further significance for the support it offers in construing the memory test as an index of working memory, given its ability to predict performance on another task. Finally, performance on the UCT was also unrelated to performance on the request task, but the 'Look first' FB test was strongly associated with it (mean correct responses for those passing the 'Look first' FB test = 6.43, mean for those failing it = 4.70; ($F(1,58) = 10.06$, $p < .001$).

5.3.6 Regression analyses

Given the apparently interrelated nature of performance on the request task and 'Look first' FB test (Siegal & Beattie, 1991), language scores, memory test scores and age, regression analyses were used to determine the relative importance of effects, and the extent of overlap in variance between different predictors. Before doing so, however, it was noted that whilst there were no differences of any kind in the mean number of correct responses to the evident and non-evident references, the two were not in fact significantly correlated at conventional levels ($r = .21$, $p = .052$, one-tailed). In other words, whilst the overall levels were the same, the data suggest that it was to some extent *different* children giving these responses.

For this reason, it was considered appropriate to derive separate regression models for correct responses to requests for the evident belief toys and the non-evident belief toys, with age, language score, memory score and 'Look first' FB test (Siegal & Beattie, 1991) performance as potential predictors. A further regression model was constructed for performance on the 'Look first' FB test, with the same predictors, except that request task performance replaced theory of mind score, as a check on reciprocity of effects. All analyses were single-level forced entry. The outcomes of the regression analyses are shown in Table 5.9.

Table 5.9: Outcomes of regression analyses for request tasks evident vs. non-evident contexts, correct responses and performance on 'Look first' FB test (Siegal & Beattie, 1991).

	Evident belief items correct		**Non-evident belief items correct**		**'Look first' FB Test**	
Predictors	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**
Age	.32	.028	.31	.03	.29	.048
Language	.00	ns	.06	ns	.09	ns
Memory	.19	ns	.04	ns	.21	ns
'Look first' FB Test	.01	ns	.17	ns	-	-
Evident belief correct	-	-	-	-	.01	ns
Non-evident belief correct	-	-	-	-	.16	ns
Model details	Adj $R^2 = .14$ $F(4,55) = 3.38$, $p = .015$		Adj $R^2 = .16$ $F(4,55) = 3.85$, $p = .008$		Adj $R^2 = .23$ $F(5,54) = 4.54$, $p = .002$	

Ostensibly, the model for each dependent is the same: a significant influence of age and no other factor. However, since the effect of age must be supposed to be mediated by other processes unless pure maturation effects are inferred, the point of interest is the nature of the subsidiary influences, and what happens to these when age is removed as a predictor. Table 5.10 shows the results of the regression analyses with age removed as a predictor.

Table 5.10: Outcomes of regression analyses for request tasks evident vs. non-evident contexts, correct responses and performance on 'Look first' FB test (Siegal & Beattie, 1991) with age removed as predictor.

	Evident belief items correct		**Non-evident belief items correct**		**'Look first FB test**	
Predictors	**Beta**	**p**	**Beta**	**p**	**Beta**	**p**
Language	.04	ns	.10	ns	.12	ns
Memory	.26	.059	.11	ns	.27	.040
'Look first' FB test	.12	ns	.28	.043	-	-
Evident belief correct	-	-	-	-	.08	ns
Non-evident belief correct	-	-	-	-	.24	.056
Model details	Adj R^2 = .07 F(3,56) = 2.60, p = .060		Adj R^2 = .10 F(3,56) = 3.24, p = .029		Adj R^2 = .18 F(4,55) = 4.40, p = .004	

For evident belief item-correct responses, the only other predictor of any note is memory score. This remains ns when age is excluded (no significant model emerges in fact), but it does come close (beta = .26, p = .059). In contrast, for non-evident belief item correct responses, the other predictor of note is performance on the 'Look first' FB test (Siegal & Beattie, 1991), and when age is excluded this effect becomes significant (beta = .28, p = .043). For performance on the 'Look first' FB test itself, *both* the subsidiary predictor for the evident belief items and the reciprocal effect of the non-evident belief items are secondary predictors; when age is excluded memory becomes significant (beta = .27, p = .04) and non evident belief item-correct responses is near-significant (beta = .24, p = .056).

The implication is that although both types of request task response were made correctly to about the same extent, and both are predicted by increased age first and foremost, the underlying nature of the process through which age has an impact is different for evident and non-evident belief items. For the former, the primary element appears to be the ability to keep track of what has been seen (cf. the age/working memory relation in Study 2); for the latter, the key predictor seems to be a conceptual advance that is held in common with performance on the Siegal & Beattie (1991) ‘Look first’ FB test. Performance on the last is also age-related, but the underlying influence appears to be from both processes. It is therefore not the case that performance on the non-evident request task items is simply a proxy for performance on the ‘Look first’ FB test.

5.4 Discussion

With regard to the first three predictions (sub-section 5.1.7), correct responses to requests for items via reference to evident and non-evident belief states did increase with age in line with prediction 2, with the middle age group performing at or around chance level, and the oldest age group doing better, but not yet having reached ceiling, in line with prediction 3. However, contrary to prediction 1, yet again, as in Study 3, no difference was found between correct responding to the two reference types being investigated. Given the care taken to avoid the kinds of support that comprehension might have provided in Study 3, the lack of context effect here is not readily explicable in these terms, and seems at first sight to be somewhat puzzling.

Putting this to one side for the moment, in terms of level of pragmatic sensitivity, non-responses (or inactive incorrect responses) were low in the middle age group, and non-existent in the oldest (much the same as in Study 3). By comparison, active incorrect responding was also found to be employed to a greater extent by all children indicating some degree of pragmatic sensitivity to the conversational requirement consistent with Study 3. However, there was only marginal differentiation of response according to request form and only in the oldest age group, with such differentiation also being slightly less than was apparent in Study 3. This was not consistent with full decoding of the implicature contained in the indirect request, either, since that ought to lead to undifferentiated *action* responses, and not predominantly the verbal responses that ensued. Not only were the oldest age group somewhat short of the ceiling in terms of accuracy of response then, they were also below ceiling in terms of pragmatic sensitivity, in line with prediction 2. Overall performance in these respects was broadly comparable to that for references to emotions in Study 3, or, if a little lower, within the bounds of variation that might be expected to be found for different samples. There seems little to choose between grasp of references to emotion and to belief states, then, contra Ninio & Snow (1996), with both emerging in fairly consistent fashion in the third year. Note that despite this discrepancy with Ninio & Snow, in fact emergence at a similar point in time would be in line with evidence on children's own references to emotions and mental states appearing at much the same time (e.g. Bretherton & Beeghly, 1982). Also contra to Ninio & Snow, and counter to prediction 4, there was no evidence here to support the notion that understanding of references to belief states is harder for children of these ages than understanding of references to shared past events; as comparison with the data from Study 2 plainly shows, even the oldest age group of children performed at a notably lower level for past event references in terms of both accuracy of response and pragmatic sensitivity.

Such inferences regarding comparative difficulty and point of emergence rest on the present data being reliable, of course. The lack of difference between grasp of references to evident and non-evident beliefs may be regarded as a potential challenge to this, but in addition there are other aspects of the response patterns found here which might also be taken as indications of the presence of task artefacts. The most notable of these is the apparent bias here towards verbal responses, where the bias in Study 3 (which was very similar in design in many respects) was towards action responses, the more natural default, given that the task was fundamentally about requests for objects. This is compounded by the fact that amongst the oldest age group, where action responses appeared more commonly, these seemed to be made primarily by a sub-group of children distinct from those making verbal responses, who were frequently inaccurate in their choice of toy (more than 30% being wrong) whilst the verbally responding children were wholly accurate. The source of this individual variation is hard to discern on the available data, but the bigger point in any case is why the shift to verbal responses at all within what is ostensibly the same task structure as Study 3.

One argument may be that the procedure in this case, by avoiding the kind of psychological support present in Study 3, seemed more formal to children because it provided fewer understanding checks. The experimental setting in Study 4 may also have been rendered more formal because of necessary limitations to the warm-up period. In Study 4, although children were introduced to, and got to play with all the toy animals, the containers and contents (gifts and food items) could not be brought into play and shared in the warm-up period because this might have led to later contamination of the request task. Therefore, it cannot be discounted that politeness became a barrier to children reaching for the objects requested. The lack of contact with some aspects of the materials may have implicitly signalled that they were the experimenter's and thus not to be handled by the child him/herself without permission.

This may have set a tone for the response bias in which the more distant verbal responses seemed more natural than actually handing over objects, which perhaps requires a degree more familiarity of manner for children to feel at ease in making. Children of 3 years of age are only likely to have recently been introduced to these polite forms of request, and so may well be unsure of what they mean (cf. Nippold, Leonard & Anatopolous, 1982).

Sample differences might be another reason for differences in response mode to Study 3 though there was little evident difference in background or language levels. So too, though, is the possibility of specific task demands which might have coloured the data in other ways. Ultimately, then, the possibility that the form of task used here (and in Study 3) leads to greater and more hidden artefacts than was previously recognised cannot be discounted, with these leading in turn to a lack of difference between evident and non-evident referents and/or the decoding of these, as well as instabilities in the nature of individual response patterns.

There is another possibility, though, under which these other variations can be seen simply as secondary in character and not especially unusual in the context of work in this area: quite simply, that once references to emotions and beliefs are grasped, it matters little whether the targets of these are evident or not – in the sense that they are manipulated here, at least – since such grasp is implicitly linked to a process of inference connecting inner state with outer event, which is activated regardless of the type of reference. Note that it is in fact *this* that is the most parsimonious explanation of the progression with age coupled with absence of reference type effects observed now in both Studies 3 and 4. Under this account, when a request is made, it instigates a search for a toy that fits to the witnessed scenario in terms of the outcome belief (or emotion) state referred to, identification being made via the inferred connection – providing the child can make this inference. Since the evidence available to children to support the inference process is the same in both instances, on this analysis the evident vs. non-evident distinction is reduced now to at most one instance of prior activation of this inference process. For instance, in the evident belief scenario (where the characters saw the swap, and therefore have line of sight knowledge of the contents of their present), children may have made a spontaneous inference as to their belief state on witnessing this. In the Study 3 emotion task, the corresponding scenario was accompanied by an explicit act of labelling the resulting emotion, but it was never particularly supposed even then that children's responses would be made simply on the basis of this labelling: identification would require reactivation of the inference process, since the referent is *always* intrinsically non-evident in the sense of not being physically available.

In fact, in Study 4, there was one other distinction between the referents under the evident and non-evident contexts, that should have been controlled for since it represented a potential confound: the evident referents were always true beliefs and the non-evident were always false, and there is good evidence that children in the age groups employed here find true beliefs easier to understand than false. In the light of the above argument, there might actually be better reason to think that *this* distinction would be more likely to make a difference to responses than the one intended except that of course there were no differences. However, the evidence in favour of the true belief vs. false belief distinction typically comes from more standard theory of mind tasks, where, as well as making an inference of mental state, children have to attribute a consequent choice or some form of action on behalf of another character based on that mental state. Whilst it is unclear exactly how much this adds to the task demands, as will be discussed shortly there is evidence from the present study alone that it does add something. In the context of the request task, then, where only the inference and a personal response is required, there is some reason to suppose that the true/false belief distinction might be less. Nevertheless, there is a general consensus that inference of false belief under conditions like these requires some level of discounting or inhibiting of personal knowledge which true belief does not, creating an expectation that some difference should nevertheless remain. On closer analysis, though, it becomes easier to see why this did not manifest. As far as the middle age group were concerned, performance overall was around chance level, as already noted, suggesting that the majority of children in this group lacked the understanding to make the inference of either true or false belief – in the same way as this age group generally appeared to lack the understanding to infer either emotion in Study 3. In contrast, the majority of those in the oldest age group passed the 'Look first' FB test (Siegal & Beattie, 1991) indicating that they were capable of making inferences of false belief – and thus were also likely to be capable of inferring true belief states too. In short, then, there are grounds for thinking it was characteristic of the children in *this* sample to be able to make both inferences or neither. The one point against this is the non-significant correlation between correct responses to evident and non-evident items. However, the value of this correlation was in the right direction, and was in fact not far from significant at .052. A certain amount of noise in the data produced by chance responding on the part of the younger children may have been sufficient to reduce the observed value to this point. The correlation may also have been suppressed by the performance of that subset of children in the oldest age group who made action responses, given that these were frequently - and unsystematically - incorrect, introducing a further element of noise.

The connections to performance on the ‘Look first’ FB test (Siegal & Beattie, 1991) are of wider significance too, of course. In general, this test seems to have worked well as a triangulating device in both Studies 3 and 4, pointing up amongst other things the similarity of the inferential requirements of understanding of references to emotion and to belief states. In contrast, the version of the UCT (Perner, Leekam & Wimmer, 1987) employed here appeared to have functioned poorly, failing to discriminate different ages or to predict performance elsewhere. This may perhaps simply reflect its constant position at the conclusion of a long series of tests for children of this age. In a sense, of course this confirmed prediction 6, but for spurious reasons, and further adjudication on the PCI/GCI distinction (Grice, 1975; Levinson, 1983; Carston, 2004) is therefore not possible.

The ‘Look first’ FB test (Siegal & Beattie, 1991) was robust in discriminating ages, though, as noted already, as well as in its consistent, coherent pattern of relationship to performance elsewhere. As far as the latter is concerned, the relationship between request task and ‘Look first’ FB test performance in particular is arguably the most important piece of evidence confirming the essential validity of the former, and the relative lack of artefactual contamination. The same point applies by extension to the emotion version of the task in Study 3. The memory measure, unavailable in Study 3, but of considerable predictive value in Studies 1 and 2, was of additional use in helping to disentangle the various strands of influence. In particular, although the effect of age was predominant for both the request task and the 'Look first' FB test, examination of the subsidiary influences revealed a common factor across the evident – or true – belief items, and the FB test in the impact of working memory, whilst – as was the case in Study 3 – there was a stronger, more direct relationship between the non-evident – or false – belief items and 'Look first' FB test performance. The first connection might perhaps be seen as reflecting the information management capacity necessary to make true belief inferences (request task) as well as holding false belief inferences in mind whilst attributing action to another character ('look first' test). The second connection suggests – crucially – some shared conceptual element between understanding references to false beliefs and the capacity to make unsupported inferences of this kind. This is in line with supporting the initial conjecture of a direct linkage between the two, although it is plain that this is by no means a one-to-one correspondence: as already noted, there are other influences at work with regard to performance on false belief tasks. The repetition of the form of relationship here perhaps also sheds light on the nature of that apparent in Study 3: the non-evident emotion there, on the current analysis, was characterised less by being non-evident than by being the product of a small element provoking sadness within a wider happy context, and thus in some sense requires a kind of discounting similar to that required in rejecting a focus on the ‘true’ state of affairs in favour of the more specific state of one (or two) individual characters.

Of this, only the ‘look first’/non-evident item relationship is actually in line with prediction 5. The predicted relationship of performance on non-evident items to working memory, based in part on Study 2’s results, was not apparent, and its absence is curious, in fact. One possibility is that the ‘look first’/non-evident item relationship takes up variance that might otherwise be attributable to working memory, because of the ‘look first’/memory relationship, suppressing evidence of the effect. It certainly seems unlikely that evident items carried working memory loads that were absent for non-evident items. Also absent were the predicted effects of language on both types of item, even when age was excluded as a predictor, contra Study 3. There would seem to be two possible explanations here: first, language effects are actually a proxy for working memory effects, and the inclusion of a direct measure of memory therefore wipes out the apparent effect of language; or second, the task involved in Study 4 is different in nature to that used in Study 3. Given other obvious points of correspondence, the latter does not seem a likely explanation, whereas the former has a degree of plausibility that merits further investigation.

Chapter 6

6.1 Conclusions and future directions for research

6.1.1 Degree of support for main hypotheses

The major aim of the present research was to redress the lack of clarity on the specifics of children's development of conversational abilities, and on if and how adult support might play a role in the acquisition of these. It will be remembered that Ninio & Snow (1996) argue that a central aspect of pragmatic ability is the capacity to understand conversational references to objects and events outside the immediate context. They also claim that this pragmatic ability is acquired by parents systematically bringing into dialogue with young children what effectively amount to violations of Grice's (1957) maxim of relevance, by using four types of reference of this kind – those to non-immediate objects, shared past events, non-evident emotions, and non-evident beliefs. In the introduction to this thesis, it was posited that such violations, together with the different forms of support employed by parents for their interpretation, not only help children to grasp the meaning of these specific types of reference, but encourage them more generally to take the perspective of others by interpreting what the speaker has in mind when conversing on abstract, remote and non-evident topics. Social support for pragmatic development was therefore argued to be the basis for wider mentalising abilities.

At root, then, the research reported here was an attempt to identify support for a social constructivist account of the development of theory of mind and related abilities. Its key objectives were therefore 1) to map the emergence of children's competences with regard to decoding each of Ninio & Snow's (1996) four types of reference, by devising tasks that would help shed light on the relationship of these competences to their reported observations of parent-child interaction; and 2) to establish further the relationship of these competences (especially understanding of references to internal states) to performance on false belief and similar theory of mind tasks.

These objectives were met by conducting four studies employing related methodologies, one for each reference type, which tested children's grasp of these when embedded in both direct and indirect requests for different objects. In each case, their grasp of these non-evident references was compared to their ability to decode similar requests for more immediately relevant objects. The principal hypothesis tested across the four studies was that the sequence of emergence of competences with regard to each reference type would correspond to the order in which Ninio & Snow (1996) observed parental support for each to appear, in line with the social constructivist account. The use of different request forms made it possible also to gauge how far grasp of these reference types was affected by the introduction of other kinds of pragmatic demand, and whether children could deal with different demands simultaneously. In addition, the last two studies, those addressing references to non-evident emotions and non-evident beliefs, included theory of mind measures of ostensibly varying degrees of difficulty, in order to ascertain whether performance on these was related to understanding of related types of reference, as had been hypothesised should be the case.

In general terms, the four studies established that children came to the experimental settings with varying degrees of understanding of the different reference types, and also that these varied with age. Thus the general thrust of the notion that performance would differ according to reference type was supported. The oldest children, aged 3.5 to 4 years, approached ceiling in one study, performed generally well in two others, but relatively poorly in the fourth. The middle age group of children, 2.5 to 3 years, performed at or just above chance level in all but one study, whilst the youngest children, 1.5 to 2 years, performed less well on both of the studies where they were included. Beyond this, though, whilst bearing in mind the qualifications identified with regard to the later studies in particular (discussed further below), the sequence of emergence of competence with regard to each reference type appeared to be first non-immediate objects; then emotions and beliefs (these being more or less concurrent); and then objects within shared past events. With regard to the impact of varying the form of requests, as is discussed in more detail later, this differed according to both task and age group, and appeared on balance to be a further indicator of children's grasp of the reference type, rather than creating any fixed additional demand. As far as performance on the theory of mind measures was concerned, the Look' FB task (Siegal & Beattie, 1991) related to both grasp of references to emotions and to non-evident false beliefs. Responses to the unexpected contents test (UCT) (Perner, Leekam & Wimmer, 1987) employed in

Study 4 were more or less at chance, though – perhaps unsurprisingly, given the age of the sample – and were unrelated to understanding of reference types.

The first point to note, therefore, is that whilst there was a measure of support for some aspects of the main hypotheses, the sequence of emergence of understanding of the four reference types, as gauged from the relative performance of the different age groups, was at odds with that proposed by Ninio & Snow (1996). In particular, in their sequence, past events are referred to and discussed *prior* to conversations concerning non-evident emotions and beliefs, not after. However, Ninio & Snow's sequencing is based on qualitative research, employing observations of mother-child pairs in their natural environment. The observed differences are therefore not necessarily fatal to social constructivist accounts of the growth of the four competences considered here, superficially at least, since observations of the start point for parental support need not actually map in exact fashion onto the emergence of competence per se. For instance, the incidence of parental support might be skewed to some extent by the relative frequency of relevant events. In particular, the need to refer to past events as the basis for current activity may be greater from a younger age than the need to refer to emotions and beliefs, which are in a sense more specialised sets of 'objects' that impinge less often on current activity. The net result may be that supported reference to past events might appear to become commonplace earlier simply in response to functional need, without actual competence following suit in any strict fashion, especially if the complexity and demand imposed by such references are greater, and full competence takes longer to put in place. Overall, then, the differences in sequence are not so large as to be uninterpretable within Ninio & Snow's framework.

Even putting aside the post hoc nature of such interpretations, however, a more searching consideration of the data from the four studies leaves such attempts at qualification seeming inadequate. At a practical level, there are questions about how far the four studies have produced sufficiently comparable data to make it safe to conclude that the sequence of emergence identified above is in fact a reliable one. As will be seen shortly, the answers to these questions are broadly positive, but in dealing with the issue of comparability, it becomes apparent that there are complex variations between the four reference types in cognitive demand, underlying influences and trajectory of growth that leave any *simple* social constructivist account looking unequal to the task of explaining the data. This should not be seen as a failure: recognition of these complexities is itself a step forward, and one which provides in many respects a more powerful basis for further research. It does mean that the initial start point for the research reported here has to be left behind, though. These issues, and related concerns, are dealt with in more detail in what follows.

6.1.2 Cross-task calibration

The issue of cross-task calibration is of central importance, since it concerns how far it is possible to adjudicate accurately on the relative point of emergence of grasp of the four reference types. In bald terms, the sequence of emergence was determined by comparison of the performance of the three age groups employed in the research, both within task, and more critically between them. So, for instance, understanding of references to non-immediate objects was placed first in the sequence, because the oldest age group performed close to ceiling in Study 1, the middle age group performed moderately well, and even the youngest age group showed some signs of comprehension. Similarly, understanding of references to objects within shared past events was placed last in the sequence, because even the oldest age group performed poorly on this. The youngest age group were not involved in the tasks relating to understanding of references to non-evident emotions and non-evident beliefs, because they were expected to do poorly. However, it was still possible to place grasp of these references as emerging at a point intermediate between non-immediate objects and shared past events, since the oldest age group performed at well above chance levels in Studies 3 and 4, and certainly better than they did in Study 2. However, it must be acknowledged that if the four tasks differed in inherent difficulty, the same basic pattern of performance might be apparent. Age effects are therefore potentially confounded with task effects, making the issue of cross-task comparability a crucial one.

The difficulty is that there are in fact various signs of task effects being present, and points where comparability between tasks was not strictly maintained, despite efforts to keep the structures as closely similar as possible. So, for instance, at first sight at least it appeared possible that the need to embed the emotion request task in a relatively complex narrative structure, and check on comprehension of this, served to actually support children’s understanding, resulting in the absence of differences between responses to evident and non-evident emotions. In fact, the absence of such differences in Study 4 as well, where understanding checks were carefully constructed to preclude such support, perhaps makes the task effect account less likely for emotions; for both emotions and beliefs, the relative automaticity of inferences about internal states and their external causes once reference has been made to them seems a more plausible explanation of the absence of evident/non-evident differences.

Nevertheless, Studies 3 and 4 did differ from Studies 1 and 2 in requiring the use of a narrative structure, however comprehensible to the children being tested, potentially increasing task demands. This difference was hard to avoid, though, since there had to be *some* coherent basis for the emotions and beliefs referred to, even if this was left implicit, otherwise random responding would have been the result. It is possible that any increase in demand might have been compensated for by the fact that Studies 3 and 4 were structured so that there were *two* correct targets for each reference, as opposed to just one in Studies 1 and 2. How far these differences acted to balance each other is a moot point, however.

There was also to some extent a degree of unevenness in the amount of information that required tracking between the past event task relative to the non-immediate. Although children did not view the non-immediate items (Study 1) on entering the experimental room, this study employed requests for single objects that would be physically apparent to children using a visual search strategy. In other words, then, all that was required was for children to hold the object request in mind whilst searching within the experimental room. On the other hand, past event references

(Study 2) were directed at the association of object-pairs that had been established the previous day. Thus this task required children to retain the identifying food item referred to in the request, search their memory for the associated toy, and then hold this in mind whilst locating it. Moreover, it also required a temporary suspension of children's focus on the here and now experience of the present event to make this connection. As discussed in Chapter 1, most adult (and child) references to past events occur during present activities (e.g. reading a bedtime story to a child about a Lion can prompt a reference to a shared past visit to the zoo). This potentially made the lack of obvious connection of the past event references to the present context (except in terms of general form of requesting toys via their associated foods) seem unnatural.

A similar point might also be argued to apply to Studies 3 and 4, in fact, since here too, requests for toys were made via an attribute they possessed rather than by direct reference to the toy themselves. This made it necessary to make a connection between attribute and object, whilst retaining information about first the attribute, and then the object, to allow the process of retrieval to be completed. However, this breakdown actually serves to flag up the importance of not taking this process of task deconstruction too far: it is in fact hard to conceive of a reference to an emotion or a belief in everyday speech where this would *not* be as an attribute of a person or character. Similarly, though perhaps less obviously, references to an object in a shared past event are also highly likely to include an attribute or some other identifier to locate *which* object is being referred to in unambiguous fashion.

In other words, then, the main structural differences between the tasks employed in Studies 1 to 4 can be seen as being for the most part inevitable reflections of the inherent character of the references being made. If the tasks differ in difficulty for these reasons, then, this is arguably because of the way the references differ. There are exceptions here, of course: the use of double targets in Studies 3 and 4, and the references to objects *only* by their associated attribute in Study 2 are not particularly likely to be characteristics of everyday references of these types. However, there is little indication that the use of double targets modified ease of responding in any large degree; if it had, it seems likely that the oldest age group would have been responding at ceiling levels. Similarly, the form of reference used in Study 2 was implicitly supported by the fact that the experimenter and child had only strictly limited shared past experience, a major aspect of which focused on the food-toy pairings. Moreover, the reference did make explicit that it was a toy that was being sought. Putting all these points together, therefore, the main non-necessary elements of task variation seem unlikely to have impacted on performance other than marginally.

Overall, then, there seem to be grounds for holding that the reliability, validity and relative calibration of the four different request tasks was broadly satisfactory. Further support for this conclusion is provided by the apparently consistent covariations between performance on the tasks and other indices. One of the latter, pragmatic sensitivity is discussed in more detail below. In general terms, though, children showed a) most awareness of the need to respond to requests, even when unsure of the correct response, and b) greatest differentiation of form of response according to request form, for the task where performance was best, non-immediate references. Similarly, they showed least awareness for the task where performance was worst, past event references. Since these two tasks used identical participants, it is hard to interpret this apparent covariation as reflecting anything other than the relative demands of the tasks themselves, which for the reasons discussed above, appear to have been broadly consistent with the demands of everyday references of these types. The pattern of results here was unexpected, but internally coherent nevertheless, and indicative therefore of convergent validity. By extension, the moderate pragmatic sensitivity shown in the emotion and belief tasks is then also consistent with the moderate levels of performance shown on these. Similar arguments apply to the correlations of request task performance with age, working memory, language, and, in Studies 3 and 4, performance on the 'Look first' FB task (Siegal & Beattie, 1991). The consistency between performance on the four request tasks and this range of other indices is hard to square with the former being subject to widespread and idiosyncratic artefacts.

6.1.3 Variation in demand characteristics of different reference types

If the request task data from the four studies are taken as reliable and valid, though, this leaves untenable any simple social constructivist account under which grasp of each reference type emerges in turn from the same process of parental introduction and interpretative support. This is because the confirmation of reliability and validity also makes it necessary to take seriously the evidence that decoding the different types of reference involves inherently different demands, and that the ability to deal with each is subject to differing influences, and possibly exhibits varying trajectories of acquisition.

To take variation in demand first of all, the notion of evident vs. non-evident referents is too simplistic. For instance, the meaning of ‘non-evident’ for non-immediate references is not the same as for references to shared past events. References to non-immediate objects and those within shared past events both involve forms of secondary intersubjectivity particular to conversation, which depend (at least in part) on deictic mechanisms other than eye gaze and gesture. As such, the *content* to which these mechanisms direct attention (i.e. the subject of the reference) is unfixed and not subject to any specific set of principles – though they may exhibit certain patterns in terms of conventional practice. The child therefore has to learn what the referential possibilities are, what deictic mechanisms help to decode these, and how to search for applications and potential targets for these.

References to objects within past events (or any remote event for that matter, including distant or future) have notably fewer constraints, however, making children’s task considerably more difficult. The only constraints that operate for references to shared past events are that such a reference should indeed be to something that is common to the experience of both speaker and listener, and that it is likely (though not certain) to be related to the immediately preceding topic of conversation. Beyond this, it is free to be anything that might happen to come to mind for the speaker (though speakers will commonly rely on such connections plausibly coming to mind for the listener too, this is not inevitably the case). This relative lack of constraint suggests that the conceptual element involved in understanding references of this kind is likely to be very general in character, perhaps along the lines of a simple awareness that references to past events occur and are generally only meaningful when these events have been shared in some way, or may be assumed to be part of common experience.

This may then be coupled with an efficient search strategy for the probable referent, based on recency, significance, and personal knowledge of the pre-occupations of the speaker. The efficiency or otherwise of this strategy may explain the working memory influence, but in this account the conceptual element is potentially as broad as something akin to Piagetian decentring (e.g. Piaget & Inhelder, 1956). For this reason, the emergence of understanding of references to past events in general at the end of the developmental sequence is plausible: degree of personal knowledge is perhaps the most critical element in any specific instance, and where this is essentially absent, as in the Study 2 test situation, the child (even when somewhat older) may be left floundering to some extent. This of course may also help explain why Ninio & Snow (1996) put the emergence of a grasp of these references somewhat earlier: they were focused on parent-child interactions, where personal knowledge on the part of the child would be much more extensive. By the same token, children's difficulties with the Study 2 task perhaps provide some evidence in favour of Ninio & Snow's (1996) claims for a distinct reliance upon historical support from parents as a crucial conversational practice for past event references to children.

It is not clear that the evident/non-evident distinction is meaningful at all for emotions and beliefs, since understanding of *any* reference to these requires some process of inference of a non-ostensible connection. Though there was also some evidence of differences in influence on grasp of evident vs. non-evident references, these are accountable for in terms of differences in explicitness/embedding (sad in a happy context), and variation in complexity of inference (true vs. false belief) – as discussed in Chapters 4 and 5 – rather than the intended dimension. Moreover, although emotions and beliefs are not physically evident in the same way as objects, whether non-immediate or remote (though emotions often have expressive accompaniment and beliefs have behavioural sequels, these are far from necessary complements), their location is constrained, and they *are* subject to certain principled, generalisable relationships between event and subjective outcome. Thus reference does not need to be subject to the same kinds of conversational deixis as non-immediate or past objects, *once* the child has become inducted into the nature of those principles – which they are through experience and conversation from around 2 years (Butterworth & Jarrett, 1991). Indeed, children appear to become rapidly become adept at searching for sources (reasons) for identified emotions and beliefs, working back from end state to plausible source (see e.g., Harris, 1989) – hence the argument in Chapter 5 regarding the lack of reference type effect in Study 4, and possibly Study 3 as well. The deictic demands of non-immediate and shared past event references are arguably actually greater than those of emotion and belief references; therefore, once the inferential principles underlying the latter have been grasped – although it is plain that each type of reference carries its own distinctive demands and processes over and above these more general points of similarity and difference.

6.1.4 Variation in sources of influence on understanding of different reference types

With regard to the different influences feeding in to the growth of understanding of the four types of reference, the role of parental support was not tested directly here, except in so far as the predicted sequence of emergence of competences was based on the apparent appearance of parental support according to Ninio & Snow (1996). However, some signs of how it might operate were perhaps evident in Study 3, in the possible effects of the understanding check questions discussed then. Given the points about deictic demand, though, there are perhaps grounds for thinking that parental support is most necessary for decoding past event referents, and rather less so for non-immediate, emotion and belief references. In this context, it is perhaps useful to remember that, as previously acknowledged, guidance within the zone of proximal development (ZPD) (Vygotsky, 1978) is by definition focused on the *performance* of an activity, not on abstract understanding, although it may implicitly contain pointers to the nature of underlying principles (see e.g., Philips & Tolmie, 2007). Where no clear abstract understanding or principle is available to guide performance, however, and yet the activity is complex, external support for it is likely to be especially critical and persistent over time.

Data from these studies confirm the operation of a range of *other* influences beyond parental support, though, and it is this that provides the greatest challenge to any simple social constructivist account of the emergence of pragmatic ability. These influences include working memory (though the extent of its impact may vary somewhat across contexts); language (though again this is variable in impact, and it is not clear why the relationship of this to performance is linear, or what exactly it indexes); and an age-related conceptual dimension(s) (perhaps the same one for emotion and beliefs, but different for past event references).

To take the first of these, across three of the four studies, memory tests demonstrated that a basic level of understanding of the task and the references made within it required efficient working memory, in line with research by Gathercole et al., (1994; 1999), which examined children's ability in respect of a series of conventional tests. The age-related working memory finding may also be consistent with Case's (1992) account of children's 'developmental increases in proficiency' across a wide range of areas. Case's theory argues that greater understanding leads to greater working memory efficiency and therefore greater scope for dealing with the detail of the task in hand.

Whilst not tested for reference to emotions, a sizeable working memory-related component was apparent for references to non-immediate objects, non-evident beliefs and past events. It should be noted that working memory was not tested using standard instruments, since the latter are not on the whole well-adapted to the context of examining language competencies, and the preference was thus for use of memory measures that seemed a more natural part of test activity, lending them greater ecological validity. Although fairly constant in terms of its presence, however, the influence of these more contextualised measures was not consistent in character. For non-immediate references (and immediate too, for that matter), the ability to keep track of information and age-related change in that ability was the only detectable influence. This is not to rule out a conceptual element being involved in understanding such references, but the lack of differentiation between age groups in its presence indicates that any such element is understood at an early age i.e. before 18 months. Support for this may be provided by Butterworth & Jarrett's (1991) work on the understanding of pointing as an attention-direction device, which indicates something of what this conceptual element might be (grasp of deixis) and when it emerges (c.12 to 14 months).

In contrast, it was found that true/evident and false/non-evident belief inferences were constrained by different factors, and making one inference was no firm guarantee of being able to make the other. Age-related working memory changes seemed to be the sole influence on understanding of reference to true or evident belief states, consistent with evidence on the relatively early emergence of grasp of the implications for knowledge of line of sight information, (Baldwin & Moses, 1994; Mutter, Alcorn & Welsh, 2006) and the ability to keep track of this. Understanding of references to false or non-evident beliefs appeared instead to stem from a conceptual element relating to inference and representation of alternative belief states to the child's own: what might be regarded as the core aspect of a representational theory of mind (cf. Flavell, 1988). This influenced performance both on formal tasks and conversational reference, and was age-related, so more likely to be present in older children. Given the common relationship to performance on the Look First FB test (Siegal & Beattie, 1991), this same component also appeared to influence performance on the emotion reference task. However, here, there appeared in addition to be a language-related influence at work. Note that this pattern of relatedness and yet distinction between cognition of emotion and belief is consistent with the divergence reported by e.g. Dunn (1995). An age-related conceptual component was also evident for understanding of reference to shared past events, though it was less clear what the character of this conceptual element might be, and it could plausibly be distinct from that operating for emotion and belief. Adjudication on this awaits evidence on the relationship between understanding of past event references and those to emotions and false beliefs, however, which is not as yet available.

The broad point to be noted here, though, is that there are implications of these multiple influences for the theory of the conversational origin of pragmatics and theory of mind: on the evidence presented here, it remains plausible that external support is important, perhaps even necessary for full grasp of the four reference types (cf. work on deaf children by e.g., Peterson & Siegal, 1995, which indicates that the absence of exposure to conversation impairs the emergence of a representational theory of mind), but it certainly does not seem to be sufficient, at least not without its effects being mediated by other internal changes.

6.1.5 Trajectory of growth across different reference types

If the data from these four studies provide clarification of the *type* of developmental change underpinning grasp of different forms of reference, they also provide some indication of the characteristic dynamic of these changes, and this in turn points at the kind of role that parental support (amongst others perhaps) might play, and why it might be important. Specific trajectories are in fact harder to discern for emotion and belief references, because of the complexities of the data (especially the lack of effect of evident vs. non-evident references), the potential masking influences of task structure, and the fact that only two age groups were employed. However, obvious differences are apparent between the growth of understanding of references to non-immediate objects and that relating to references to past events, despite the overlap in influences, with change being more gradual for latter.

Regardless of these variations in rate, though, the data are consistently more indicative of gradual progression as the general dynamic of change rather than sudden insight or conceptual shift, despite the apparent role of various conceptual elements. Stated in broad terms, the overwhelming weight of evidence was in favour of gradual, incremental progress rather than sudden conceptual insight. In both the non-immediate and past event reference tasks, for example, differences between test and baseline (i.e. immediate reference) items manifested as main effects, with no tendency towards the interaction with age that would suggest a rapid spurt in understanding. The lack of context effects in the emotion and belief reference tasks makes the use of this yardstick less helpful in these areas, but even here there was little indication of a rapid shift from chance levels of performance to ceiling. The increase was somewhat more rapid in the emotion reference task, but even here ceiling did not appear to be attained over the course of a year's worth of developmental change, and the linear associations with language and age tend to bespeak incremental rather than sudden growth.

This point is an important one when it comes to consideration of the influences on change. The cross-sectional testing of competence employed in the present studies necessarily entails a lack of direct evidence on the point, but the incremental nature of the observed changes in performance are suggestive of a situation where 'practice makes perfect'. Since for children of this age the opportunity for such practice is largely in the hands of parents (and to an extent older siblings), they must have *at least* this degree of influence on progress. Given evidence from other research that parents are in fact quite variable in the opportunities that they provide (e.g. Meins & Fernyhough, 1999, on mind-mindedness; Brown & Dunn, 1991, on variation in parental use of mentalistic language; and most relevantly, dissertation research by MacColl, 2006 showing that mothers varied substantially in the strategies they employed to help young children unpack references to past events), the potential scale of influence is in fact large, and this perhaps helps explain some of the relatively sizeable variance in the performance on the tasks employed here amongst children in the same age group.

It is therefore at least plausible that Ninio & Snow (1996) have accurately assessed the influence of the forms of parental input they report, and that if there is a limitation in their account, it is in painting this influence as consistent in strength, and operating to the exclusion of or in isolation from other constraints and influences.

There is in addition, other evidence to support the contention that underpins their case, that parents must possess (perhaps in variable degree) implicit sensitivity to the demands made by different types of reference, and to the capability of their children to cope with these demands at different ages. Brown & Dunn (1991) for example, report shifts in the type of emotional reference made by mothers from simple causal linkage to intervening interpretation, contingent upon changes in apparent level of grasp on the part of their children. The potentially widespread nature of such sensitivity is indicated by research reported by Philips & Tolmie (2007) who found that parents systematically adjusted the forms of support they used with their children on problem-solving tasks according to the level of understanding shown, or at least their expectations about it.

6.1.6 Sensitivity to competing pragmatic demands

Whilst the foregoing discussion has attempted a broad commentary on what can be discerned from the present research with regard to the emergence of children's understanding of different types of non-immediate reference, and the factors that influence this, there remain some subsidiary issues to consider, particularly with regard to the wider pragmatic sensitivity shown by children, and the measurement of false belief in relation to understanding of reference. To take the former first, the methodology employed by the studies reported here was designed to examine the impact of variation in pragmatic demands by varying the grammatical form of the requests in which references were either embedded in direct demands or in indirect polite forms (cf. Babelot & Marcos, 1999). Differences were apparent here with regards to children's age, as evidenced by older children's modification of responses, according to the referential context and grammatical form of the request and in recognising that a response was required even when the answer was uncertain.

In fact, even younger children in the present research showed emerging signs of developing pragmatic sensitivity on the Study 1 task in respect of both referential context and form. This finding supported the research of Babelot & Marcos (1999) who found that children aged 1 year and 7 months would give a correct response to a request in a relevant context and children aged 2 years and 5 months could relate the context of a request to its form.

In the event, though, there was little sign that this further layer of pragmatic demand beyond that imposed by the reference itself had *much* bearing on performance. Instead, sensitivity to these demands appeared broadly to follow understanding rather than to constrain its application, and provided a useful extra index of the degree of children's grasp. Thus, for instance, pragmatic sensitivity was not entirely stable across the four studies and the different tasks they employed, i.e. it varied as a function of the level of understanding of the task, and the referential context and form being tested. Where reference was better understood, pragmatic sensitivity was greater.

However, in respect of Studies 1, 2 and 4 at least children generally failed to achieve the third posited level of pragmatic sensitivity i.e. the ability to decode both reference *and* implicature and provide an action response, regardless of whether the request was direct or indirect. The only exception to this pattern was in Study 3, where there were signs that children were operating at this third level, since a majority of children gave action responses to indirect requests for the target toys.

The general lack of pragmatic sensitivity at the third level may be explained by the fact that the precise implicature carried by indirect requests is only decidable via knowledge of conventions which is not typically possessed by children of this age. In effect, then, the third level of pragmatic sensitivity may ultimately require an understanding of Generalised Conversational Implicatures (GCIs) (Grice, 1975; Levinson, 2000; Carston, 2004) i.e. inferences which are derivable without the help of context and the knowledge, aims and intentions of the speaker. To understand an indirect request and respond accordingly, would require knowledge based on violations of the maxim of manner (Grice, 1975) i.e. an awareness that the aims and intentions of a speaker’s request can sometimes be unclear and ambiguous, and therefore that an ostensible request for information can have an alternative meaning or interpretation – and more particularly that requests of this form are commonly used as polite methods of requesting an action. The lack of exposure to GCIs and the conventions surrounding violations the maxim of manner may have meant that children were confined to the obvious literalist interpretation of the request. As previously outlined, parents are least likely to violate the maxim of manner with children at this age, and as a result it is likely that children have not yet learned that something can be implied or meant beyond the utterance itself in these instances.

This fails to explain, however, why many children in Study 3 did apparently attain the third level of pragmatic sensitivity. The reasons for this are unclear. It seems unlikely that this can be attributed to the putative influence of task artefacts from the use of understanding checks (discussed in Chapter 4), since these were employed only to help children follow the story, that is, they were not orientated to specifically aid the decoding of a precise form of request. Moreover, following Study 4, it was unclear how far such artefacts really were in operation, since the lack of differentiation between reference types was apparent there too, without understanding checks of a potentially supportive form being used. However, it is possible that the data here reflect a different kind of artefact, a general bias towards action responses, accentuated in the case of Study 3, possibly by the rapport created by the lengthy introductory procedure, but counteracted in Study 4 by a perceived need for greater politeness, as argued in Chapter 5. Certainly, there are signs of a bias towards action responses in Study 1 as well: despite the tendency towards differentiation of response according to request form, action responses were generally more common. Plainly, though, the whole issue of the point at which children begin to understand these more conventionalised forms of implicature, and under what circumstances, merits further detailed research.

The broader point, as argued previously, is that it would appear that pragmatic sensitivity is a viable experimental measure and an index of competence in given area of conversational ability, allowing understanding of references across all the four studies to be placed in a developmental sequence. It was acknowledged earlier, for instance, that as far as actual sequencing is concerned the placing of the grasp of emotion and belief reference before past event references is plausible given the sizeable differences in pragmatic sensitivity exhibited by children with respect to these different types of reference.

One further implication of this is that demands of this kind only affect performance during a narrow window when the task in hand is beginning to be understood, but competence is patchy in character, and the child has to work harder to keep track of things (hence the working memory effects). Prior to this point, lack of understanding makes the additional pragmatic difficulties irrelevant; once understanding is reasonably well-established, the child has ample mental ‘space’ to cope with this extra element – at least in terms of determining literal conversational requirement; as noted above, understanding of general conventions may emerge later.

The variable nature of pragmatic sensitivity across tasks points up an additional issue: references are never contextually ‘clean’ i.e. stripped down to point where comprehension and decoding is the sole issue determining performance. Instead, they are embedded in conversational settings where much else has to be dealt with, bringing competing demands into play, impacting on performance. It is likely that aspects of this were at work outside of the task manipulation itself too – hence the need for a warm-up period to ensure children could get used to the experimenter and materials, and did not therefore spend time during the task proper wondering about the nature of the event they were engaged in.

6.1.7 Issues regarding the measurement of false belief understanding

A further subsidiary objective of the present research was to evaluate Siegal's (1999) and Siegal & Petersen's (1994) assertions regarding children's conversational abilities. To recap, Siegal makes three claims (1) 3 year olds have the ability to pass standard false belief tasks, but are disadvantaged by the conversational characteristics of the test situation; (2) the reason they are disadvantaged is that they are unused to violations of conversational rules of the kind employed in false belief tests; and (3) when such violations are removed, their performance shows substantial improvement. The general thrust of Siegal's first and third claims is supported by the present research. Study 4 employed standard vs. the 'Look first' FB test (Siegal & Beattie, 1991) as 'follow up' tests in addition to the reference task. Differences were found between performance on the standard false belief test i.e. the UCT (Perner, Leekam & Wimmer, 1987) – argued to rest on Generalised Conversational Implicatures (GCIs) – and the 'Look first' FB test – held to involve Particularised Conversational Implicatures (PCIs). The more or less random responding on the former, and the contrasting systematic shift with age on the latter are certainly consistent with Siegal's claims, and with the findings of previous research cited earlier in the thesis. Such research established that the 'Look first’ question has facilitating effects that enables even a majority of 3 year old children to pass the test (i.e. Surian & Leslie, 1999; Carlson, Moses & Breton, 2002; Leslie, German & Pollizi, 2006 and Yadzi et al., 2006). In contrast, the results of the ‘Look first’ test in Studies 3 and 4 run contrary to Clements & Perner’s (1994) classic study and Wellman, Cross & Watson’s (2002) extensive meta-analysis of 178 false belief tests as both failed to replicate the findings of this effect. The pattern of responding reported here also substantiates Borg’s (2005) argument and Noveck & Posada’s (2003) finding concerning the difficulty that children encounter when decoding GCIs, (discussed above) i.e. that GCIs require general knowledge of the aims and intentions of the speaker: GCIs do not provide the context specific cues to aid children’s interpretation of references that PCIs allow (Grice, 1975; Levinson, 1983; Carston, 2004)

In this respect, these data provide further support for the potential importance of the kind of conversational support discussed by Ninio & Snow (1996) for children’s early conversational performance. More importantly, though, the points made earlier about the limited time window during which pragmatic demands impact on performance also help explain the otherwise potentially puzzling rapidity of the shift from children needing the simplification of the ‘Look first’ FB test (Siegal & Beattie, 1991) to perform well, to them being able to cope with the original versions of such tasks. This pattern is simply a transitional effect of almost classic Piagetian type, and once false belief understanding is secure, the support offered by PCIs is no longer necessary: it is evident that the experimenter’s question relates to the incorrect belief that results from absence of knowledge, and the behavioural consequences of this. Crucially, though, the rapidity of this shift may help to account for the somewhat contradictory findings on the 'Look first' test, including its failure to survive meta-analytic research: unless the precise time window of transition is targeted, the effect will appear to be absent, and its absence in a number of studies for this reason would tend in turn to lead to negative meta-analytic results, due to averaging effects.

However, the overall results of the present research undermine Siegal's (1999) claim that children are unused to violations of the conversational maxims, at least in general terms. The findings of the present studies, showing that three year olds are typically able to deal with at least three forms of what are effectively violations of relevance are consistent with research which has countered Siegal's view, such as that regarding violations of the maxims of quality and quantity by Astington (1999), Dunn (1988), and Lillard (1999). Violations of maxims are evident between caregivers and children during family discourse (e.g. dinner table conversations) and those that serve social functions (Rundquist, 1992; Brumark, 2006). The latter findings also indicate that in dinner table conversations caregivers and children make references to past and future matters i.e. those that lie outside of the present context. Adult initiated violations of conversational maxims appear to be a central conversational practice.

However, Siegal (1999) focuses more specifically on violations of the maxim of manner (to be clear and unambiguous) as the reason that three year old children fail on standard false belief tests – this being the rule that adults are least likely to violate in conversation with children at this age. He may therefore be correct in asserting that children approach the standard false belief test with limited knowledge that people can sometimes be vague and ambiguous in conversation and have ulterior motives for asking questions. The appropriate conclusion is therefore perhaps that the picture is more complex than Siegal paints it, and that it is necessary to consider children's grasp of violations of general conversational principles in terms of each specific type of violation, rather than treating them as somehow uniform in character. The fact that the present research showed that understanding of two forms of conversational reference (and ostensibly violations of the maxim of relevance) were associated with passing the 'Look first' FB test (Siegal & Beattie, 1991) does suggest, however, that the broad thrust of Siegal's argument about the linkage of pragmatic understanding to performance on false belief tasks is correct, even if the specific nature of that linkage is not as he originally conceived it.

6.1.8 Directions for future research

Plainly, there is a need for more systematic and detailed work on variations in parental scaffolding of children's grasp of conversational reference, paying attention to the forms used with different types of reference, and also to the sources of these variations. These include: differences in mind-mindedness; variation in personal conversational skills, which might be assessed using tasks of the form employed by Deleau (1999) and Deleau, Le Sourn & Guehenneuc (2000); the number of children in the family, given that for parents as well as their children practice might make perfect; and the potentially lasting influences of having had older siblings themselves when they were growing up. A central aspect of such work would be to examine in direct fashion the impact of variation in support on children's grasp, both in terms of immediate performance, and longer term competence. MacColl (2006) for instance looked at the relationship between support strategy for decoding references to past events and the turns taken for children to unambiguously identify the referent, and this approach might be extended to include measurement of wider competences.

This kind of research might perhaps entail some refinement of the tasks used in the present thesis, perhaps focused upon the giving of greater historical support in such situations i.e. a protracted conversation between parent and the child concerning the past event during the present activity e.g. 'Let's stop what we are doing for a minute and try to remember where we last saw a Lion like the one in the story'. This combination of more naturalistic observation and competence testing – especially if the second were lagged, and included measures of other cognitive competences, including working memory and basic forms of false belief understanding – would help to establish more precisely the extent of the influence from social mechanisms vs. cognitive mechanisms on the development of pragmatic abilities. A further important aspect that relates to the present research and future enquiry would be to examine the relationship between language and understanding of different forms of reference. It is acknowledged that theory of mind is related to language ability, as typically measured by receptive vocabulary tests. This has been variously explained in terms of language measures indexing children's understanding of syntax (syntactic complements in particular – see e.g. Lohmann & Tomasello, 2003); their semantic grasp (understanding of mental state vocabulary in particular – see e.g. Brown & Dunn, 1991); and their pragmatic grasp (see e.g. Deleau, 1999). The relationship of the aspects of pragmatic grasp investigated here to language might seem to confirm the last of these, except that this was found with measures of productive language, not receptive. The relationship to receptive language therefore needs to be investigated. More importantly, though, more detailed examination is needed of the precise trajectories of these relationships. Their apparently linear character is puzzling, since it is hard to see what incremental features of syntax, semantics or pragmatics general language measures could be tapping into to create linear relationships with an ability that seems to grow by practice. Some more refined conceptualisation of what general language measures index is needed, in order to move on from simplistic treatment of it as a covariate. This work would also help establish the nature of the mechanisms of parental/interpersonal influence on children's conversational skills in terms of conversational, historical and psychological support, making it possible to move towards the development of forms of intervention that might be deployed both indirectly (e.g. via parenting classes) and directly in remedial work with children.

A further area of interest for future research could also focus on children's choice of response mode, given the finding in Study 4 that children primarily used polite verbal responses, when in Study 3, the same choice led predominantly to action responses. To some extent the gap between the two response modes was closing for the oldest children in Study 4, but by and large verbal answers were the response type of choice for all children. Given that children do not typically have choices of this kind in standard false belief tasks (e.g. the requirement is to *point to* or to *state* that it is the 'blue cupboard that Maxi thinks the chocolate is in' or *say* that the next child to come into the experimental room will 'think the smarties' tube contains smarties'), this raises the question as to the extent to which politeness, the formalised setting, or some other unknown concern perceived by children, acts as a barrier to responding correctly to false belief questions with an action.

One means of exploring this might be via refinements of the false belief task aimed at reducing the formalised setting in which these tasks take place, but this seems less than entirely viable: as Siegal & Peterson (1994), and Hilton (1990) argue, the experimenter's concerns are scientific and children's are pragmatic, and this gap cannot be removed completely. One more feasible refinement may therefore be via increased scaffolding or psychological support being given by the experimenter to the child in the lead up to false belief questions (e.g. more understanding checks and/or as the scenario unfolds the adult might ask, 'What do you think might happen next'), given the indications from Studies 3 and 4 that variation in the form of understanding checks and task introduction serves to establish the 'mood' of the encounter.

Further refinements in false belief tasks should consider the implication that task demands account for the differences in terms of either GCIs or PCIs. That particular kinds of conversational implicatures present different levels of difficulty to children is further evidence of the need to set the pragmatic demands of tasks, including the false belief test, at a level that is appropriate to children's age, pragmatic sensitivity abilities and working memory capacity, and perhaps future research may address this.

At the outset, the present research considered Flavell's (1999) questions concerning 'what constitutes a theory of mind?' This thesis has argued that theory of mind acquisition should not be interpreted as a conceptual achievement in isolation from a developing grasp of the pragmatics of conversational practices, and social constructivist influences on this development. This process is itself, as is now more apparent, a complex one. However, it is still contended there are fundamental grounds for holding that the pragmatic linguistic skills and behaviours that children exhibit with adults, in particular their understanding of references to objects or events outside the immediate frame of reference, should be viewed as crucial evidence regarding how knowledge and understanding of other people's minds is put in place in pre-school children.

REFERENCES

Astington, J.W. (1999). What would a theory of conversational awareness look like? Peer commentary. *Developmental Science* **2** (1) (1999), pp 15-16.

Austin, J.L. (1962). *How to do things with words*. Oxford: Oxford University Press.

Babelot, G., & Marcos, H. (1999). Comprehension of directives in young children: influence of social situation and Linguistic form. *First Language*, **19,** 165 – 186.

Baldwin, D.A. (1993b). Early referential understanding: Infant's ability to recognise referential acts for what they are. *Developmental Psychology*, **29,** 832-843.

Baldwin, D. A., & Moses. L. J. (1994). Early understanding of referential intent and attentional focus: Evidence from Language and emotion In C. Lewis and P. Mitchell (Eds.), *Children's Early Understanding of mind.* London; Lawrence Erlbaum Associates.

Baldwin, J. M. (1884). Mental development of the child and the race. *Methods and Processes.* New York: Mc Millan.

Baldwin, J.M. (1906) Social and ethical interpretations in mental development (4th edition) London: Mc Millan.

Baron-Cohen, S. (1993). From attention goal psychology to belief desire psychology: The development of a theory of mind and its dysfunction. In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), *Understanding other minds:* Perspectives from autism, Oxford: Oxford University Press.

Baron-Cohen, S., & Ring, H. (1994). A model of the mind reading system; Neuropsychological and Neurobiological perspectives in C. Lewis and P. Mitchell (Eds.), *Children's Early Understanding of mind.* London; Lawrence Erlbaum Associates.

Bezuidenhout, A., & Sroda, M.S. (1998). Children's use of contextual cues to resolve referential ambiguity: An application of Relevance Theory. *Pragmatics and Cognition* Vol. **6** (1/2), pp 265-299.

Borke, H. (1971). Interpersonal perception of young children: egocentrism or empathy. *Developmental Psychology*, **5**, 263-9.

Borg. E. (2005). Implicatures and modularity. Paper given at 9th International Pragmatics Conference. Riva Del Gorda, Italy

Bretherton. I. (1991). Intentional communication and the development of an understanding of mind in D. Frye & C. Moore (Eds.), *Children's Theories of mind, mental states and social understanding* (pp. 49-75). Lawrence Erlbaum Associates.

Bretherton. I. (1993). From dialogue to representation: The intergenerational construction of self in relationships. In C.A. Nelson (Ed.), *Minnesota Symposia on Child Psychology: Vol. 26 Memory and effect in development* (pp. 237-263). Hillsdale, NJ: Erlbaum.

Bretherton, I., & Beeghly, M. (1982). Talking about internal states: The acquisition of an explicit theory of mind. *Developmental Psychology*, **18**, 906-911.

British Ability Scales II Naming Vocabulary Test, Core Scale 2, Administration & Scoring Manual, (BAS II NVT) (1996). C.D. Elliot, P. Smith & K. McCullough (EDS.), nferNelson Publishing Company Ltd.

Brown, J.R. & Dunn, J. (1991). 'You can cry, mum': the social and developmental implications of talk about internal states. *British Journal of Developmental Psychology*, **9**, 237-256.

Brumark, A. (2006). Non-observance of the Gricean maxims in family dinner table conversation. *Journal of Pragmatics*. Vol. **38**, Issue 8, 1206 -1238.

Bruner, J.S. (1975). From communication to language: a psychological perspective' *Cognition*, **3** pp. 255-87.

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

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

Campos, J.J., & Sternberg, C.R. (1980). Perception, appraisal and emotion: The onset of social referencing. In M. Lamb & L. Sharrod (Eds.), *Infant social cognition* (pp. 273-311). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Carlson, S.M., Moses, L.J. & Breton, C. (2002). How specific is the relation between executive function and theory of mind. Contributions of inhibitory control and working memory. *Infant and Child Development*, **11**, 73 – 92.

Carston, R. (2004). Truth-conditional content and conversational implicature I C. Bianchi (ed) The semantic/pragmatic distinction 2004, CSLI publications Stanford University pp 65-100.

Case, R. (1992). *The mind's staircase*. Hillsdale, NJ: Erlbaum.

Cassidy, J., & Marvin, R. S., with the MacArthur Working group. (1992). *Attachment organisation in preschool children: procedures and coding manual*. Unpublished Manuscript, University of Virginia.

Cassidy, J., & Shaver P.R. (1999). Handbook of attachment, theory, research and clinical applications. The Guildford Press, New York, London.

CHILDES Language Database (2001). B. MacWhinney (Ed.), Lawrence Erlbaum Associates Publishers.

Clark, H.H. (1977). 'Bridging' in Johnson-Laird, P.N. and Wason, P.C. (Eds.), Thinking: readings in cognitive science, Cambridge, Cambridge University Press.

Cole, P.M. (1986). Children's spontaneous control of facial expression. *Child Development*, **57**, 1309-1321.

Crystal, D. (1987). The Cambridge encyclopaedia of language. Cambridge University Press.

Clements, W.A., & Perner, J. (1994). Implicit understanding of belief. *Cognitive Development* **9,** 377-395.

de Rosnay. M., & Hughes, C. (2006). Conversation and theory of mind: Do children talk their way to socio-cognitive understanding? *British Journal of Developmental Psychology* **24,** 7 -37.

Deleau, M. (1999). Assessing links between conversational awareness and cognitive development. Peer commentary. *Developmental Science* **2** (1) (1999), pp 17 -18.

Deleau, M., Le Sourn, S. & Guehenneuc, K. (2000). Conversations and the representation of mind in preschool children. III Conference for Sociocultural Research, Sao Paulo, Brazil.

Diesendruck, G., Markson, L., Akhtar, N. & Reudor, A. (2004). Two-year-olds' sensitivity to speakers' intent: an alternative account of Samuelson and Smith. *Developmental Science*, **7**, 33-41.

Dunn, J. (1988). The beginnings of social understanding. Cambridge, MA: Harvard University Press.

Dunn, J. (1994). Changing minds and changing relationships. In C. Lewis and P. Mitchell (Eds.), *Children's Early Understanding of mind.* London; Lawrence Erlbaum Associates.

Dunn, J. (1995). Children as psychologists: the latter correlates of individual differences in understanding of emotions and other minds. *Cognition and Emotion*, **9**, 187-201.

Dunn, J. (2006). Commentary. A discussion of the Merrill-Palmer Quarterly issue. *Merill-Palmer Quarterly*, Vol **52**, No.1. pp 151-157.

Dunn, J., Brown, J., Slomkowski, C., Tesla, C. & Youngblade, L. (1991). Young children's understanding of other people's feelings and beliefs: individual differences and their antecedents. *Child Development*, **62**, 1352-1366.

Flavell, J.H. (1988). The development of children's knowledge about the mind: from cognitive connections to mental representations. In J.W. Astington, P.L. Harris and D.R. Olson (Eds.), *Developing theories of mind*. Cambridge: Cambridge University Press.

Flavell. J.H. (1999). Cognitive development: Children's knowledge about the mind. *Annual Review of Psychology*, **50**, 21 –45.

Flavell, J.H., Flavell, E. R., & Green, E. L. (1983). Development of the appearance-reality distinction. *Cognitive Psychology*, **15**, 95-120.

First Words & First Sentence Test & Manual, (FWFS), (1997). B. Gilliam, J. Boyle, & N Smith (Eds.), Hodder & Stoughton.

Fodor, J.A. (1992). A theory of the child's theory of mind. *Cognition*, **44**, 283-296

Garfield, J.L., Peterson, C.C., & Perry, T. (2001). Social Cognition, language acquisition and the development of the Theory of mind. *Mind and Language*, Vol. **16** (5) pp. 494-541.

Gathercole, S.E., Willis, C.S., Baddeley, A.D. & Emslie, H. (1994). The children's test of nonword repetition; a test of phonological working memory. *Memory*, **2,** 103-27.

Gathercole, S.E., Service, E., Hitch, G.J., Adams, A. & Martin, A.J (1999). Phonological short-term memory and vocabulary development: further evidence on the nature of the relationship. *Applied Cognitive Psychology,* **13**, 65-77.

Gopnik, A., & Astington, J.W. (1988). Children's understanding of representational change and its relation to the understanding of changes in their mental states. *Child Development*, **59**, 26-37.

Goffman, E. (1983). Felicity's condition. *American Journal of Sociology*, **89**, 1-53.

Greenfield, P.M., & Smith, J.H. (1976). *The structure of communication in early language development*. New York: Academic Press.

Grice, H.P. (1957). Meaning. *Philosophical Review*, **66** 377-388. Reprinted in Grice, 1989b.

Grice, H.P. (1969). Utterer's meaning and intentions. *Philosophical Review*, **78** 144-177. Reprinted in Grice, 1989b.

Grice, H.P. (1975). Logic and conversation. In P. Cole & J.L. Morgan (Eds.), *Syntax and Semantics,* Vol. 3: Speech acts (pp. 41-58). New York: Academic Press.

Grice, H.P. (1982). Meaning revisited. In N.V. Smith (Ed.), Mutual Knowledge. London: Academic Press. Reprinted in Grice, 1989b.

Grice, H.P. (1989a) Retrospective Epilogue. In Grice, 1989b.

Grice, H.P. (1989b) Studies in the way of words. Cambridge, MA: Harvard University Press.

Harris, P.L. (1989). Children and Emotion. Oxford; Basil Blackwell.

Harris., P.L., Olthof.,T., Terwogt., M.M., & Hardman., C.E. (1987). Children's knowledge of the situations that provoke emotion. *International Journal of Behavioral Development,* **10**, No 3, 319-343.

Haslett, B., & Samter. W. (1997). Children Communicating. The first five years. Lawrence Erlbaum Associates.

Haviland, J.M. and Lelwica, M. (1987). The induced affect response: 10 week-old infants' responses to three emotional expressions. *Developmental Psychology*, **23**, pp. 97-104.

Heyes, C.M. (1993). Anecdotes, training, trapping and triangulating: Do animals attribute mental states? *Animal Behaviour*, **46**, 177 – 188.

Heyes, C. M. (1998) Theory of mind in non-human primates. *Behavioural and Brain Sciences,* **21**, 101- 148.

Hilton, D.J. (1990). Conversational processes and causal explanation. *Psychological Bulletin,* **107**, 65-81.

Izard, C.E., Heubner, R.R., Risser, D., McGinnes, G.C., & Dougherty, L.M. (1980) The young infant's ability to produce discrete emotion expressions. *Developmental Psychology*, **16**, pp. 132-40.

Kellerman, K., Broetzmann, S., Lim, T., & Kitao, K. (1989). The conversation MOP: Scenes in the stream of discourse. *Discourse processes*, **12**, 27-62.

Leslie, A.M. (1987). Pretence 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. J. Cohen (Eds.), *Understanding other minds:* Perspectives from autism, Oxford: Oxford University Press.

Leslie, A.M., German, T.P.,& Pollizi, P. (2005). Belief-desire reasoning as a process of selection. *Cognitive Psychology*, **50**, 45-85.

Levinson, S.E. (1983). Pragmatics, Cambridge University Press.

Levinson, S.E. (2000). Presumptive meaning. The theory of generalised Conversational Implicatures. Cambridge, Mass: MIT Press.

Lillard, A.S. (1999). Siegal on Piaget's legacy: Gricean child meets blundering experimenter. Peer Commentary. *Developmental Science* **2** (1) (1999), pp 18-20.

Lock, A. (1978). On being picked up. In A. Lock, (Ed.), *Action, gesture and symbol*. New York: Academic Press.

Lohmann, H., & Tomasello, M. (2003). The role of language in the development of false belief understanding: A training study. *Child Development* **74**, 1130 -1144.

Lourenco, O., & Machado, A. (1999). Rich conversation but vague theory. Peer commentary. *Developmental Science* **2** (1) (1999), pp 21-23.

MacColl, F. (2006). Children's understanding of a reference to a past event through the aid of conversation. Unpublished Honours dissertation, University of Strathclyde.

Manchester Language Study in CHILDES Language Database (2001). B. MacWhinney (Ed.), Lawrence Erlbaum Associates Publishers.

Marvin, R. S., & Greenberg, M. T. (1982). Preschoolers' changing conception of their mothers: A social –cognitive study of mother-child attachment. In D. Forbes & M.T. Greenberg (Eds.), New directions for child development: No 18 Children's planning strategies (pp. 47-60). San Francisco: Josey-Bass.

Meins, E., & Fernyhough, C. (1999). Linguistic acquisitional style and mentalising development: The role of maternal mind-mindedness. *Cognitive Development*, **14**, 363-380.

Mitchell, P. (1994). Realism and early conception of mind: A synthesis of phylogenetic and ontogenetic issues. In C. Lewis & P. Mitchell (Eds.), Children's early understanding of mind: Origins and development (pp 19-45). Hove: Lawrence Erlbaum Associates Ltd.

Mitchell, P. (1996). Acquiring a conception of mind: *A review of psychological research and theory*. Psychology press publishers.

Mossler, D.G., Marvin, R.S. & Greenberg, M.T. (1976). Conceptual perspective taking in 2-to6 year children. *Developmental Psychology*, **12.** 85-6.

Mutter, B., Alcorn, M.B., & Welsh, M. (2006). Theory of mind and executive function: working memory capacity and inhibitory control as predictors of false belief task performance. *Perception and Motor Skills*. **102** (3) 819-35.

Newton, P., Reddy, V., & Bull, R. (2000). Children's everyday deception and performance on false belief tasks. *British Journal of Developmental Psychology*,**18**, 297 –317.

Ninio, A., & Snow, C. E. (1996). *Pragmatic Development*. Boulder, Colorado: Westview Press.

Nippold, M.A., Leonard, L.B., & Anastopolous, A. (1982). Development in the use of polite forms in children. *Journal Speech and Hearing*. **25**, 193 –202.

Noveck., I. & Posada. A (2003). Characterising the time course of an implicature. An evoked potentials study. *Brain & Language*, **85**, 203 -210.

Perner, J., Leekam, S., & Wimmer, H. (1987). Three year olds difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology,* **5**, 125 –137.

Peterson, C. C., & Siegal, M. (1995). Deafness, conversation and theory of mind. *Journal of Experimental Psychology and Psychiatry*, **36**, 459-474.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioural and Brain Sciences,* **4**, 515-526.

Piaget, J. (1959). The language and thought of the child, 3rd edition: Routledge & Kegan Paul.

Piaget, J. & Inhelder, B. (1956). *The child's conception of space*. London: Routledge & Kegan Paul.

Philips, S. & Tolmie, A. (2007). Children's performance on and understanding of the Balance Scale problem: the effects of parental support. *Infant and Child Development*.

Reddy, V., & Simone, L. (1995). Acting on attention: towards an understanding of knowing in infancy. BPS Developmental Psychology Annual Conference, Glasgow.

Reese, E., & Brown, N. (2000). Reminiscing and recounting in the preschool years. *Applied Cognitive Psychology*, **14**, 1-17.

Robinson, E.J. (1981). The child’s understanding of inadequate messages and communication failure: A problem of ignorance or egocentrism. In P. Dickson (Ed.), *Children’s Oral Communication Skills*. New York; Academic Press, 167-188.

Robinson, E.J., & Mitchell, P. (1995). Masking of children’s early understanding of the representational mind: Backwards explanation versus prediction. *Child Development,* **66**, 1022-1039.

Rommetveit, R. (1971). Words, context and verbal message transmission. In A.E. Carswell & R. Rommetveit (Eds.), Social context of messages (pp. 13-26). New York: Academic Press.

Rommetveit, R. (1972). Language games, syntactic structures and hermeneutics. In J. Israel & H.Tajfel. (Eds.), *The context of social psychology* (pp. 212-267). New York Academic press.

Rommetveit, R. (1979a). On the architecture of intersubjectivity. In R. Rommetveit & R.M. Blakar (Eds.), *Studies of Language, thought and verbal communication* (pp. 93-107). New York Academic press.

Rommetveit, R. (1979b). On negative rationalism in scholarly studies of verbal communication and dynamic residuals in the construction of human intersubjectivity. In R. Rommetveit & R.M. Blakar (Eds.), *Studies of Language, thought and verbal communication* (pp. 147-161). New York Academic press.

Rundquist, S. (1992). Indirectness: a gender study of flouting Grice’s maxims. *Journal of Pragmatics,* **18**, 431 -449.

Sachs, J. (1983). Talking about there and then: The emergence of displaced reference in parent-child discourse. In K.E. Nelson (Ed.), *Children’s language,* Vol**.** *IV* (pp. 1-28). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Saltmarsh, R., Mitchell, P., & Robinson, E.J. (1995). Realism and children’s early grasp of mental representation: Belief-based judgements in the state change task. *Cognition* **57,** 297-325.

Schegloff, E.A., & Sacks, H. (1973). Opening up closings. *Semiotica **8*** pp 289-327.

Searle, J.R. (1969). *Speech acts*, Cambridge: Cambridge University Press.

Siegal, M. (1993). Knowing children’s minds. *Behavioural and Brain Sciences,* **16**, 79-80.

Siegal, M. (1999). Language and thought: the fundamental significance of conversational awareness for cognitive development. Article with peer commentaries and response. *Developmental Science* **2** (1) pp 1-34.

Siegal, M., & Beattie, K. (1991). Where to look first for children's understanding of false beliefs. *Cognition* **38,** 1-12.

Siegal, M., & Peterson, C. (1994). Children's theory of mind and the conversational territory of cognitive development. In C. Lewis and P Mitchell (Eds.), *Children's Early Understanding of mind.* London; Lawrence Erlbaum Associates.

Sinclair, J. McH., & Coulthard, R.M. (1975). Towards an analysis of discourse the English used by teachers and pupils, Oxford University Press, London.

Sorce, J. F., Emde, R, N., Campos, J. J. & Klinnert, M.D. (1985). Maternal emotional signalling: its effects on the visual cliff behaviour of 1-year olds. *Developmental Psychology*, **21**, 195-200.

Stevenson, E. J., (1995). The malleability of the developing representational mind. PhD, University of Wales.

Stevenson, E. J., & Mitchell, P. (1992). *The suggestibility of false belief*. Paper presented at the BPS Developmental Section conference, September Edinburgh.

Stevenson, E. J., & Mitchell, P. (1995b). Reversion to realism in a false belief test following suggestion. Unpublished Manuscript. University of Swansea.

Sperber, D., & Wilson, D. (1986). Relevance: communication and cognition, Oxford: Blackwell.

Sperber, D., & Wilson, D. (2002). *Pragmatics, modularity and mind reading*. http://www.dan.sperber.com/pragmatics-modularity-and-mindreading.htm version (pp 1-19). Also to appear in *Mind & Language*, (2002). **17** pp 3 –23).

Surian, L., & Leslie, A.M. (1999). Competence and performance in false belief understanding: A comparison of autistic and normal 3-year old children. British *Journal of Developmental Psychology*, **17**, 141 -155.

Tardiff Language study in CHILDES Language Database (2001). B. Macwhinney (Ed.). Lawrence Erlbaum Associate Publishers.

Tolmie, A. (1991). Structure and content in children's concepts of emotion. Doctoral thesis, City of London Polytechnic.

Vygotsky, L. (1962). *Thought and language*. Cambridge: MIT Press.

Vygotsky, L. (1978). *Mind in Society: The development of higher psychological processes.* M. Cole. V. Johns-Steiner. S. Cribner and E. Souberman (Eds.) Cambridge: Harvard University press.

Wellman, H.M., Cross, D., & Watson, J. (2001).Meta analysis of theory of mind development: The truth about false belief. *Child Development* **72**, 655-684.

Wells Language Study in CHILDES Language Database (2001). B. MacWhinney (Ed.) Lawrence Erlbaum Associate Publishers.

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.

Woodhead, M., Barnes, P., Miell, D. & Oates, J. (1995). Developmental perspectives on emotion in P. Barnes (Ed.), Personal, social and emotional development of children. (pp 135-1870. Blackwell Publishers.

Yadzi, A. A., German, T. P., Defeyter, M.A., & Siegal, M. (2006). Competence and performance in belief-desire reasoning across the cultures. The truth, the whole truth and nothing but the truth about false belief tasks. *Cognition* **100**, 343-368.

APPENDIX A

RELEVANCY VIOLATIONS & THEORY OF MIND ACQUISITION EXPERIMENTAL DATA RECORDING PRO-FORMA 1

EXPERIMENT NUMBER: 1 NAME: AGE: NURSERY: DATE:

IMMEDIATE – NON-IMMEDIATE CONTEXTS

TRIALS **RESPONSES**

QUESTIONS	REFERENCE FORM	REFERENCE TYPE	CORR ACTION RESP	CORR VERBAL RESP	INCORR ACTION RESP	INCORR UNINFORM RESP	INCOMP RESP	NON RESPONSE
Give me the Dog	DIRECT REQUEST	IMMEDIATE						
Have you got the Bear?	INDIRECT REQUEST	NON-IMMEDIATE						
Have you got the Cat?	INDIRECT REQUEST	IMMEDIATE						
Give me the Monkey	DIRECT REQUEST	IMMEDIATE						
Give me the Frog	DIRECT REQUEST	NON-IMMEDIATE						
Have you got the Rabbit?	INDIRECT REQUEST	IMMEDIATE						
Give me the Lion	DIRECT REQUEST	NON-IMMEDIATE						
Have you got the Penguin?	INDIRECT REQUEST	NON-IMMEDIATE						

RESPONSE CONVERSION KEY -
CORR ACTION RESP = **CORRECT ACTION RESPONSE**; (E.G. GIVES MENTIONED OBJECT).
CORR VERBAL RESP= **CORRECT VERBAL/GESTURAL RESPONSE;** (E.G. "IT'S OVER THERE" OR POINTS TO OBJECT)
INCORR ACTION RESP = **INCORRECT ACTION RESPONSE**; (E.G. GIVES DIFFERENT OBJECT).
INCORR UNINFORM RESP = **INCORRECT/ UNINFORMATIVE VERBAL RESPONSE**; (E.G. "YES" OR "NO", BUT NO UNDERSTANDING)
INCOMP RESP = **INCOMPREHENSION RESPONSE**; (E.G. "DON'T KNOW"; "WHAT", PUZZLED)
NON-RESPONSE **= NON-RSEPONSE**; (E.G. NO REACTION FROM CHILD)

APPENDIX B

RELEVANCY VIOLATIONS & THEORY OF MIND ACQUISITION EXPERIMENTAL DATA RECORDING PRO-FORMA 2

EXPERIMENT NUMBER: 1 NAME: AGE: NURSERY: DATE:

IMMEDIATE – NON-IMMEDIATE CONTEXTS

TRIALS **RESPONSES**

QUESTIONS	REFERENCE FORM	REFERENCE TYPE	CORR ACTION RESP	CORR VERBAL RESP	INCORR ACTION RESP	INCORR UNINFORM RESP	INCOMP RESP	NON RESPONSE
Have you got the Frog?	INDIRECT REQUEST	IMMEDIATE						
Give me the Rabbit	DIRECT REQUEST	NON-IMMEDIATE						
Give me the Penguin	DIRECT REQUEST	IMMEDIATE						
Have you got the Lion?	INDIRECT REQUEST	IMMEDIATE						
Have you got the Dog?	INDIRECT REQUEST	NON-IMMEDIATE						
Give me the Bear	DIRECT REQUEST	IMMEDIATE						
Have you got the Monkey?	INDIRECT REQUEST	NON-IMMEDIATE						
Give me the Cat	DIRECT REQUEST	NON-IMMEDIATE						

RESPONSE CONVERSION KEY-
CORR ACTION RESP = **CORRECT ACTION RESPONSE**; (E.G. GIVES MENTIONED OBJECT).
CORR VERBAL RESP = **CORRECT VERBAL/GESTURAL RESPONSE;** (E.G. “IT’S OVER THERE” OR POINTS TO OBJECT)
INCORR ACTION RESP = **INCORRECT ACTION RESPONSE**; (E.G. GIVES DIFFERENT OBJECT).
INCORR UNINFORM RESP = **INCORRECT/ UNINFORMATIVE VERBAL RESPONSE**; (E.G. “YES” OR “NO”, BUT NO UNDERSTANDING)
INCOMP RESP = **INCOMPREHENSION RESPONSE**; (E.G. “DON’T KNOW”; “WHAT”, PUZZLED)
NON-RESP = **NON-RSEPONSE**; (E.G. NO REACTION FROM CHILD)

APPENDIX C

RELEVANCY VIOLATIONS & THEORY OF MIND ACQUISITION EXPERIMENTAL DATA RECORDING PRO-FORMA 3

EXPERIMENT NUMBER: 2 NAME: AGE: NURSERY: DATE:

PAST AND PRESENT CONTEXTS

TRIALS **RESPONSES**

QUESTIONS	REFERENCE FORM	REFERENCE TYPE	CORR ACTION RESP	CORR VERBAL RESP	INCORR ACTION RESP (PA)	INCORR ACTION RESP (PR)	INCORR ACTION RESP (SC)	INCORR UNINFORM RESP	INCOMP RESP	NON RESPONSE
Give me the animal that likes cake (Dog)	DIRECT REQUEST	PAST EVENT								
Have you got the animal that likes tomato sauce? (Lion)	INDIRECT REQUEST	PRESENT EVENT								
Have you got the animal that likes chocolate? (Cat)	INDIRECT REQUEST	PAST EVENT								
Give me the animal that likes burger. (Frog)	DIRECT REQUEST	PRESENT EVENT								
Give me the animal that likes chips (Monkey)	DIRECT REQUEST	PAST EVENT								
Have you got the animal that likes sausages? (Rabbit)	INDIRECT REQUEST	PAST EVENT								
Give me the animal that likes Oranges (Penguin)	DIRECT REQUEST	PRESENT EVENT								
Have you got the animal that likes egg? (Bear)	INDIRECT REQUEST	PRESENT EVENT								

RESPONSE CONVERSION KEY-

CORR ACTION RESP = **CORRECT ACTION RESPONSE** ; (E.G. GIVES MENTIONED OBJECT)
CORR VERBAL RESP = **CORRECT VERBAL/GESTURAL RESPONSE;** (E.G. "IT'S OVER THERE" OR POINTS TO OBJECT
INCORR ACTION RESP (PA) **= INCORRECT ACTION RESPONSE**; (E.G. GIVES DIFFERENT PAST OBJECT) .
INCORR ACTION RESP (PR) = **INCORRECT ACTION RESPONSE**; (E.G. GIVES DIFFERENT PRESENT OBJECT)
INCORR ACTION RESP (SC) = **INCORRECT ACTION RESPONSE**; (E.G. GIVES DIFFERENT SUPPLEMENTAL/CONTROL OBJECT)
INCORR UNINFORM RESP = **INCORRECT/UNINFORMATIVE VERBAL RESPONSE**; (E.G**.** "YES" OR "NO", BUT NO UNDERSTANDING)
INCOMP RESPONSE = **INCOMPREHENSION RESPONSE**; (E.G. "DON'T KNOW"; "WHAT", PUZZLED)
NON-RESPONSE = **NON-RESPONSE**; (E.G. NO REACTION FROM CHILD)

APPENDIX D

RELEVANCY VIOLATIONS & THEORY OF MIND ACQUISITION EXPERIMENTAL DATA RECORDING PRO-FORMA 4

EXPERIMENT NUMBER: 2 NAME: AGE: NURSERY: DATE:

PAST AND PRESENT CONTEXTS

TRIALS **RESPONSES**

QUESTIONS	REFERENCE FORM	REFERENCE TYPE	CORR ACTION RESP	CORR VERBAL RESP	INCORR ACTION RESP (PA)	INCORR ACTION RESP (PR)	INCORR ACTION RESP (SC)	INCORR UNINFORM RESP	INCOMP RESP	NON RESPONSE
Have you got the animal that likes burger? (Frog)	INDIRECT REQUEST	PAST EVENT								
Give me the animal that likes egg. (Bear)	DIRECT REQUEST	PAST EVENT								
Give me the animal that likes chocolate? (Cat)	DIRECT REQUEST	PRESENT EVENT								
Have you got the animal that likes chips? (Monkey)	INDIRECT REQUEST	PRESENT EVENT								
Have you got the animal that likes oranges? (Penguin)	INDIRECT REQUEST	PAST EVENT								
Give me the animal that likes sausages? (Rabbit)	DIRECT REQUEST	PRESENT EVENT								
Have you got the animal that likes cake? Dog)	INDIRECT REQUEST	PRESENT EVENT								
Give me the animal that likes tomato sauce. (Lion)	DIRECT REQUEST	PAST EVENT								

RESPONSE CONVERSION KEY-

CORR ACTION RESP = **CORRECT ACTION RESPONSE**; (E.G. GIVES MENTIONED OBJECT)
CORR VERBAL RESP = **CORRECT VERBAL/GESTURAL RESPONSE**; (E.G. “IT’S OVER THERE” OR POINTS TO OBJECT)
INCORR ACTION RESP (PA) = **INCORRECT ACTION RESPONSE**; (E.G. GIVES DIFFERENT PAST OBJECT) .
INCORR ACTION RESP (PR)= **INCORRECT ACTION RESPONSE**; (E.G. GIVES DIFFERENT PRESENT OBJECT)
INCORR ACTION RESP (SC) = **INCORRECT ACTION RESPONSE**; (E.G. GIVES DIFFERENT SUPPLEMENTAL OBJECT)
INCORR UNINFORM RESP= **INCORRECT/ UNINFORMATIVE VERBAL RESPONSE**; (E.G**.** “YES” OR “NO”, BUT NO UNDERSTANDING)
INCOMP RESPONSE = **INCOMPREHENSION RESPONSE**; (E.G. “DON’T KNOW”; “WHAT”, PUZZLED)
NON-RESPONSE = **NON-RESPONSE**; (E.G. NO REACTION FROM CHILD)

APPENDIX E

RELEVANCY VIOLATIONS & THEORY OF MIND ACQUISITION EXPERIMENTAL DATA RECORDING PRO-FORMA 5

EXPERIMENT NUMBER: 3 NAME: AGE: NURSERY: DATE:

EVIDENT (HAPPY) AND NON-EVIDENT (SAD) CONTEXTS

TRIALS **RESPONSES**

QUESTIONS	REFERENCE FORM	REFERENCE TYPE	CORR ACTION RESP	CORR VERBAL RESP	INCORR ACTION RESP (GDEO)	INCORR UNINFORM RESP	INCOMP RESP	NON RESP
Give me an animal that is happy (Elephant/Frog)	DIRECT REQUEST	EVIDENT EMOTION						
Have you got an animal that is sad? (Bear/Rabbit)	INDIRECT REQUEST	NON-EVIDENT EMOTION						
Have you got an animal that is happy? (Elephant/Frog)	INDIRECT REQUEST	EVIDENT EMOTION						
Give me an animal that is happy (Elephant/Frog)	DIRECT REQUEST	EVIDENT EMOTION						
Give me an animal that is sad (Bear/Rabbit)	DIRECT REQUEST	NON-EVIDENT EMOTION						
Have you got an animal that is happy? (Elephant/Frog)	INDIRECT REQUEST	EVIDENT EMOTION						
Give me the animal that is sad (Bear/Rabbit)	DIRECT REQUEST	NON-EVIDENT EMOTION						
Have you got the animal that is sad? (Bear/Rabbit)	INDIRECT REQUEST	NON-EVIDENT EMOTION						

RESPONSE CONVERSION KEY-
CORR ACTION RESP = **CORRECT ACTION RESPONSE**; (E.G. GIVES MENTIONED OBJECT)
CORR VERBAL RESP= **CORRECT VERBAL/GESTURAL RESPONSE**; (E.G. "IT'S OVER THERE" OR POINTS TO OBJECT)
INCORR ACTION RESP (GDEO) = **INCORRECT ACTION RESPONSE**; (E.G. GIVES DIFFERENT EMOTION OBJECT)
INCORR UNINFORM RESP' = **INCORRECT/ UNINFORMATIVE VERBAL RESPONSE**; (E.G. "YES" OR "NO", BUT NO UNDERSTANDING)
INCOMP RESPONSE = **INCOMPREHENSION RESPONSE**; (E.G. "DON'T KNOW"; "WHAT", PUZZLED)
NON-RESPONSE **= NON-RESPONSE**; (E.G. NO REACTION FROM CHILD)

APPENDIX F

<u>RELEVANCY VIOLATIONS & THEORY OF MIND ACQUISITION EXPERIMENTAL DATA RECORDING PRO-FORMA 6</u>

EXPERIMENT NUMBER: 3 NAME: AGE: NURSERY: DATE:

EVIDENT (HAPPY) AND NON-EVIDENT (SAD) CONTEXTS

<u>TRIALS</u> **<u>RESPONSES</u>**

QUESTIONS	REFERENCE FORM	REFERENCE TYPE	CORR ACTION RESP	CORR VERBAL RESP	INCORR ACTION RESP (GDEO)	INCORR UNINFORM RESP	INCOMP RESP	NON RESPONSE
Have you got an animal that is sad? (Elephant/Frog)	INDIRECT REQUEST	NON-EVIDENT EMOTION						
Give me an animal that is happy (Bear/Rabbit)	DIRECT REQUEST	EVIDENT EMOTION						
Give me an animal that is sad? (Elephant/Frog)	DIRECT REQUEST	NON-EVIDENT EMOTION						
Have you got an animal that is happy (Bear/Rabbit)	INDIRECT REQUEST	EVIDENT EMOTION						
Have you got an animal that happy? (Bear/Rabbit)	INDIRECT REQUEST	EVIDENT EMOTION						
Give me an animal that is happy? Bear/Rabbit)	DIRECT REQUEST	EVIDENT EMOTION						
Have you got an animal that is sad? (Elephant/Frog)	INDIRECT REQUEST	NON-EVIDENT EMOTION						
Give me an animal that is sad (Elephant/Frog)	DIRECT REQUEST	NON-EVIDENT EMOTION						

<u>RESPONSE CONVERSION KEY-</u>
CORR ACTION RESP = **CORRECT ACTION RESPONSE**; (E.G. GIVES MENTIONED OBJECT)
CORR VERBAL RESP = **CORRECT VERBAL/GESTURAL RESPONSE**; (E.G. "IT'S OVER THERE" OR POINTS TO OBJECT)
INCORR ACTION REP (GDEO) = **INCORRECT ACTION RESPONSE**; (E.G. GIVES DIFFERENT EMOTION OBJECT)
INCORR UNINFORM RESP = **INCORRECT/ UNIFORMATIVE VERBAL RESPONSE**; (E.G. "YES" OR "NO", BUT NO UNDERSTANDING)
INCOMP RESPONSE = **INCOMPREHENSION RESPONSE**; (E.G. "DON'T KNOW"; "WHAT", PUZZLED)
NON-RESPONSE = **NON-RESPONSE**; (E.G. NO REACTION FROM CHILD)

APPENDIX G

RELEVANCY VIOLATIONS & THEORY OF MIND ACQUISITION EXPERIMENTAL DATA RECORDING PRO-FORMA 7

EXPERIMENT NUMBER: 4 NAME: AGE: NURSERY: DATE:

EVIDENT (TRUE) AND NON-EVIDENT (FALSE) BELIEF CONTEXTS

TRIALS **RESPONSES**

QUESTIONS	REFERENCE FORM	REFERENCE TYPE	CORR ACTION RESP	CORR VERBAL RESP	INCORR ACTION RESP (GDBO)	INORR UNINFORM RESP	INCOMP RESP	NON RESPONSE
Give me an animal that thinks it got Crisps (Elephant or Frog)	DIRECT REQUEST	TRUE BELIEF- SWAP SEEN						
Have you got an animal that thinks it got a Pencil (Bear or Rabbit)	INDIRECT REQUEST	FALSE BELIEF- SWAP NOT SEEN						
Have you got an animal that thinks it got Crisps (Elephant or Frog)	INDIRECT REQUEST	TRUE BELIEF- SWAP SEEN						
Give me an animal that thinks it got Crisps? (Elephant or Frog)	DIRECT REQUEST	TRUE BELIEF- SWAP SEEN						
Give me an animal that thinks it got a Pencil (Bear or Rabbit)	DIRECT REQUEST	FALSE BELIEF- SWAP NOT SEEN						
Have you got an animal that thinks it got Crisps (Elephant or Frog)	INDIRECT REQUEST	TRUE BELIEF- SWAP SEEN						
Give me the animal that thinks it got a Pencil (Bear or Rabbit)	DIRECT REQUEST	FALSE BELIEF- SWAP NOT SEEN						
Have you got an animal that thinks it got a pencil (Bear or Rabbit)	INDIRECT REQUEST	FALSE BELIEF- SWAP NOT SEEN						

RESPONSE CONVERSION KEY-
CORR ACTION RESP = **CORRECT ACTION RESPONSE**; (E.G. GIVES MENTIONED OBJECT)
CORR VERBAL RESP = **CORRECT VERBAL/GESTURAL RESPONSE**; (E.G. "IT'S OVER THERE" OR POINTS TO OBJECT)
INCORR ACTION REP (GDBO) **= INCORRECT ACTION RESPONSE**; (E.G. GIVES DIFFERENT BELIEF OBJECT)
INCORR UNINFORM RESP **= INCORRECT/ UNIFORMATIVE VERBAL RESPONSE**; (E.G**.** "YES" OR "NO", BUT NO UNDERSTANDING)
INCOMP RESPONSE = **INCOMPREHENSION RESPONSE**; (E.G. "DON'T KNOW"; "WHAT", PUZZLED)
NON-RESPONSE **= NON-RESPONSE**; (E.G. NO REACTION FROM CHILD)

APPENDIX H

RELEVANCY VIOLATIONS & THEORY OF MIND ACQUISITION EXPERIMENTAL DATA RECORDING PRO-FORMA 8

EXPERIMENT NUMBER: 4 NAME: AGE: NURSERY: DATE:

EVIDENT (TRUE) AND NON-EVIDENT (FALSE) BELIEF CONTEXTS

TRIALS **RESPONSES**

QUESTIONS	REFERENCE FORM	REFERENCE TYPE	CORR ACTION RESP	CORR VERBAL RESP	INCORR ACTION RESP (GDBO)	INCORR UNINFORM RESP	INCOMP RESP	NON RESPONSE
Have you got an animal that thinks it got play doh? (Elephant or Frog)	INDIRECT REQUEST	FALSE BELIEF-SWAP NOT SEEN						
Give me an animal that thinks it got a mini eggs (Bear or Rabbit)	DIRECT REQUEST	TRUE BELIEF-SWAP SEEN						
Give me an animal that thinks it got play doh (Elephant & Frog)	DIRECT REQUEST	FALSE BELIEF-SWAP NOT SEEN						
Have you got an animal that thinks it got mini eggs? (Bear or Rabbit)	INDIRECT REQUEST	TRUE BELIEF-SWAP SEEN						
Have you got an animal that thinks it got mini eggs? (Rabbit or Bear)	INDIRECT REQUEST	TRUE BELIEF-SWAP SEEN						
Give me an animal that thinks it got a mini eggs (Bear or Rabbit)	DIRECT REQUEST	TRUE BELIEF-SWAP SEEN						
Have you got an animal that thinks it got play doh? (Elephant & Frog)	INDIRECT REQUEST	FALSE BELIEF-SWAP NOT SEEN						
Give me an animal that thinks it got play doh (Elephant or Frog)	DIRECT REQUEST	FALSE BELIEF-SWAP NOT SEEN						

RESPONSE CONVERSION KEY
CORR ACTION RESP = **CORRECT ACTION RESPONSE**; (E.G. GIVES MENTIONED OBJECT)
CORR VERBAL RESP = **CORRECT VERBAL/GESTURAL RESPONSE**; (E.G. "IT'S OVER THERE" OR POINTS TO OBJECT)
INCORR ACTION REP (GDBO) **= INCORRECT ACTION RESPONSE**; (E.G. GIVES DIFFERENT BELIEF OBJECT)
INCORR UNINFORM RESP **= INCORRECT/ UNIFORMATIVE VERBAL RESPONSE**; (E.G**.** "YES" OR "NO", BUT NO UNDERSTANDING)
INCOMP RESPONSE = **INCOMPREHENSION RESPONSE**; (E.G. "DON'T KNOW"; "WHAT", PUZZLED)
NON-RESPONSE **= NON-RESPONSE**; (E.G. NO REACTION FROM CHILD)

APPENDIX J

Table 2.8: F-values and significance levels for Study 1 responses by reference type, reference form and age

	Ref type		*Ref form*		*Age*		*Ref type x Age*	
	(F (1,76)	*p*	*(F (1,76)*	*p*	*(F (1,76)*	*p*	*(F (2,76)*	*p*
Responses								
Corr Action R	45.90	< .001	18.73	< .001	5.88	< .05	3.28	< .05
Corr Verbal R	3.90	.052	15.78	< .001	2.84	ns	2.33	ns
Ovr Correct R	30.59	< .001	-	-	10.68	< .001	0.89	ns
Incorr Action R	7.18	< .05	0.16	ns	0.60	ns	2.05	ns
Incorr Uninf R	0.01	ns	2.59	ns	0.11	ns	1.55	ns
Non R	17.28	< .001	0.32	ns	20.02	< .001	1.44	ns

	Ref form x Age		*Ref type x Ref form*		*Ref type x Ref form x Age*	
	(F (2,76)	*p*	*(F (2,76)*	*p*	*(F (2,76)*	*p*
Responses						
Corr Action R	1.32	ns	6.10	< .05	0.34	ns
Corr Verbal R	5.38	< .05	3.94	.051	1.60	ns
Ovr Correct R	-	-	-	-	-	-
Incorr Action R	0.28	ns	2.35	ns	0.93	ns
Incorr Uninf R	1.14	ns	1.46	ns	0.24	ns
Non R	0.10	ns	0.49	ns	1.11	ns

Response Conversion Key

Ref type = Reference type; Ref form = Reference form; ns = not significant

Corr Action R = **Correct action responses**; Correct Verbal R = **Correct verbal/ gestural responses**; Ovr Correct R = **Overall correct responses**; Incorr Action R = **Incorrect Action Response, gives different object**; Incorr Uninf = **Incorrect/uninformative verbal response;** Non R = **Non-Response**.

APPENDIX K

Table 2.9: F-values and significance levels for active vs. inactive (Study 1) by reference type, incorrect and age

	Ref type		*Incorrect*		*Age*		*Ref type x Age*	
	(F (1,76)	*p*	*(F (1,76)*	*p*	*(F (1,76)*	*p*	*(F (2,76)*	*p*
Active vs. Inactive	21.11	< .001	1.22	ns	13.41	< .001	1.09	ns

	Incorrect x Age		*Ref type x Age*		*Ref type x Incorrect x Age*	
	(F (2,76)	*p*	*(F (2,76)*	*p*	*(F (2,76)*	*p*
Active vs. Inactive	9.03	< .001	0.67	ns	1.02	ns

Response Conversion Key

Ref type = Reference type: ns = not significant

Active = **Incorrect action responses (gives different objects) plus Incorrect/uninformative verbal responses vs.** Inactive = **Non-Responses**.

APPENDIX L

Table 3.8: ***F-values and significance levels for Study 2 responses by reference type, reference form and age***

	Ref type		*Ref form*		*Age*		*Ref type x Age*	
	(F (1,72)	*p*	*(F (1,72)*	*p*	*(F (2,72)*	*p*	*F2,72)*	*p*
Responses								
Corr Action R	18.11	< .001	0.00	ns	16.03	< .001	3.67	< .05
Corr Verbal R	0.00	ns	0.55	ns	2.12	ns	4.19	< .05
Ovr Correct R	18.85	< .001	-		23.56	< .001	0.60	ns
Incorr Action R	1.80	ns	1.14	ns	0.08	ns	2.98	.057
Incorr Uninf R	0.42	ns	0.01	ns	0.39	ns	0.11	ns
Non R	12.96	< .001	0.31	ns	25.50	< .001	1.06	ns

	Ref form x Age		*Ref type x Ref form*		*Ref type x Ref form x Age*	
	(F (2,72)	*p*	*(F (2,72)*	*p*	*(F (2,72)*	*p*
Responses						
Corr Action R	0.48	ns	0.32	ns	0.49	ns
Corr Verbal R	1.21	ns	1.71	ns	0.23	ns
Ovr Correct R	-		-		-	
Incorr Action R	1.19	ns	2.68	ns	0.85	ns
Incorr Uninf R	0.75	ns	0.41	ns	0.43	ns
Non R	0.87	ns	1.39	ns	1.20	ns

Response Conversion Key

Ref type = Reference type; Ref form = Reference form; ns = not significant

Corr Action R = **Correct action responses**; Correct Verbal R = **Correct verbal/ gestural responses**; Ovr Correct R = **Overall correct responses**; Incorr Action R = **Incorrect Action Response, gives different present object**; Incorr Uninf = **Incorrect/uninformative verbal response;** Non R = **Non-Response**.

APPENDIX M

Table 3.9: F-values and significance levels for active vs. inactive (Study 2) by reference type, incorrect and age

	Ref type		*Incorrect*		*Age*		*Ref type x Age*	
	(F (1,72)	*p*	*(F (1,72)*	*p*	*(F (1,72)*	*p*	*(F (2,72)*	*p*
Active vs. Inactive	14.59	< .001	23.74	< .001	24.35	< .001	1.15	ns

	Incorrect x Age		*Ref type x Incorrect*		*Ref type x Incorrect x Age*	
	(F (2,72)	*p*	*(F (2,72)*	*p*	*(F (2,72)*	*p*
Active vs. Inactive	12.78	< .001	6.87	< .05	1.12	ns

Response Conversion Key

Ref type = Reference type:

Active = **Incorrect action responses (gives different objects) plus Incorrect/uninformative verbal responses vs.** Inactive = **Non-Responses**.

APPENDIX N

Table 4.9: ***F-values and significance levels for Study 3 responses by reference type, reference form and age***

	Ref type		*Ref form*		*Age*		*Ref x Age*	
	(F (1,58)	*p*	*(F (1,58)*	*p*	*(F (1,58)*	*p*	*(F(1,58)*	*p*
Responses								
Corr Action R	1.44	ns	1.54	ns	6.80	< .05	4.31	< .05
Corr Verbal R	0.00	ns	6.75	< .05	0.05	ns	2.29	n.s.
Ovr Correct R	1.60	ns	-		7.73	< .05	1.60	n.s.
Incorr Action R	4.91	< .05	4.36	ns.	0.38	ns	0.01	n.s
Incorr Uninf R	3.81	.056	1.17	ns	4.64	< .05	1.69	n.s.
Non R	0.04	ns	0.14	ns	3.68	.060	3.61	.062

	Ref form x Age		*Ref type x Ref form*		*Ref type x Ref form x Age*	
	(F (1,58)	*p*	*(F (1,58)*	*p*	*(F (1,58)*	*p*
Responses						
Corr Action R	0.82	ns	1.15	ns	3.20	ns
Corr Verbal R	0.00	ns	0.00	ns	1.58	ns
Ovr Correct R	-		-		-	
Incorr Action R	0.70	ns	4.36	.053	0.20	ns
Incorr Uninf R	2.65	ns	0.00	ns	0.27	ns
Non R	1.27	ns	0.83	ns	2.31	ns

Response Conversion Key

Ref type = Reference type; Ref form = Reference form; ns = not significant

Corr Action R = **Correct action responses**; Correct Verbal R = **Correct verbal/ gestural responses**; Ovr Correct R = **Overall correct responses**; Incorr Action R = **Incorrect Action Response, gives different emotion object**; Incorr Uninf = **Incorrect/uninformative verbal response;** Non R = **Non-Response**.

APPENDIX P

Table 4.10: ***F-values and significance levels for active vs. inactive (Study 3) by reference type, incorrect and age***

	Ref type		*Incorrect*		*Age*		*Ref type x Age*	
	(F (1,58)	*p*	*(F (1,58)*	*p*	*(F (1,58)*	*p*	*(F (,58)*	*p*
Active vs. Inactive	1.90	ns	10.92	< .05	8.05	< .05	0.47	ns

	Incorrect x Age		*Ref type x Incorrect*		*Ref type x Incorrect x Age*	
	(F (1,58)	*p*	*(F (1,58)*	*p*	*(F (1,58)*	*p*
Active vs. Inactive	0.03	ns	1.41	ns	1.03	ns

Response Conversion Key

Ref type = Reference type: ns = not significant

Active = **Incorrect action responses (gives different emotion objects) plus Incorrect/uninformative verbal responses vs.** Inactive = **Non-Responses**.

APPENDIX Q

Table 5.11: ***F-values and significance levels for Study 4 responses by reference type, reference form and age***

	Ref type		Ref form		Age		Ref type x Age	
	(F (1,58)	*p*	*(F (1,58)*	*p*	*(F (1,58)*	*p*	*(F(1,58)*	*p*
Responses								
Corr Action R	0.66	ns	1.01	n.s	10.07	< .05	0.37	ns
Corr Verbal R	0.24	ns	1.64	n.s	0.07	ns	0.59	ns
Ovr Correct R	0.00	ns	0.25	n.s	12.86	< .001	1.60	ns
Incorr Action R	0.00	ns	0.15	n.s.	1.32	ns	0.00	ns
Incorr Uninf R	0.10	ns	0.19	n.s.	2.46	ns	0.10	ns
Non R	1.09	ns	0.32	n.s	5.89	< .05	3.61	ns

	Ref form x Age		Ref type x Ref form		Ref type x Ref form x Age	
	(F (1,58)	*p*	*(F (1,58)*	*p*	*(F (1,58)*	*p*
Responses						
Corr Action R	1.62	ns	0.16	ns	2.66	ns
Corr Verbal R	1.64	ns	0.13	ns	0.38	ns
Ovr Correct R	0 06	ns	0.55	ns	0.99	ns
Incorr Action R	0.01	ns	0.48	ns	0.95	ns
Incorr Uninf R	0.19	ns	1.00	ns	1.00	ns
Non R	0.32	ns	0.32	ns	0.32	ns

Response Conversion Key

Ref type = Reference type; Ref form = Reference form; ns = not significant

Corr Action R = **Correct action responses**; Correct Verbal R = **Correct verbal/ gestural responses**; Ovr Correct R = **Overall correct responses**; Incorr Action R = **Incorrect Action Response, gives different belief object**; Incorr Uninf = **Incorrect/uninformative verbal response;** Non R = **Non-Response**.

APPENDIX R

Table 5.12: ***F-values and significance levels for active vs. inactive (Study 4) by reference type, incorrect and age***

	Ref type		*Incorrect*		*Age*		*Ref type x Age*	
	(F (1,58)	*p*	*(F (1,58)*	*p*	*(F (1,58)*	*p*	*(F (,58)*	*p*
Active vs. Inactive	0.05	ns	15.91	< .001	13.57	< .001	0.15	ns

	Incorrect x Age		*Ref type x Incorrect*		*Ref type x Incorrect x Age*	
	(F (1,58)	*p*	*(F (1,58)*	*p*	*F (1,58)*	*p*
Active vs. Inactive	0.14	ns	0.35	ns	0.18	ns

Response Conversion Key

Ref type = Reference type: ns = not significant

Active = **Incorrect action responses (gives different belief objects) plus Incorrect/uninformative verbal responses vs.** Inactive = **Non-Responses**