On the development of the concept of desire
in the theory-theory of mind framework

by

Leonardo Ferres

A thesis submitted to
the Faculty of Graduate Studies and Research
in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

Institute of Cognitive Science

Carleton University

Ottawa, Ontario, Canada

December, 2004

© Leonardo Ferres, 2004

Library and
Archives Canada

Bibliothèque et
Archives Canada

Published Heritage
Branch

Direction du
Patrimoine de l'édition

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

ISBN: 0-494-00797-4

Our file Notre référence

ISBN: 0-494-00797-4

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

ABSTRACT

This dissertation has as its main objective the analysis of the hypothesis that ascribing mental states of desire is present at early stages of child development.

In order to investigate this issue, several analyses in the NOS (Naturally Occurring Speech) paradigm were carried out on a novel database of child language, replicating and extending those analyses pioneered by Bartsch and Wellman (1995). The analyses supported the hypothesis that children are able to ascribe desires by the youngest age analyzed (18 months) and by the lowest mean length of utterance analyzed ($MLU < 1.75$) as well. This points to an early grasp of the concept of desire, definitely earlier than the concept of belief, as reported in the developmental literature (Wellman, Cross & Watson, 2001), though some reservations about this issue are put forward as well in the last chapter. Several tests applied on the data (all with the NOS paradigm) yielded further support to the hypothesis.

This dissertation concerns the following topic areas: developmental psychology, cognitive science, concepts, language acquisition, philosophy of mind, and Theory Theory of Mind.

To Loreto.

TABLE OF CONTENTS

1	Introduction	1
1.1	Introducing H1	1
1.2	Introducing the project: Testing H1	4
1.3	The road map	8
1.4	Overview of the studies	12
2	Theory of mind Research	15
2.1	Charting the conceptual map: Situating desires in the general Theory-Theory of Mind framework	15
2.1.1	Folk psychology (FP)	18
2.1.2	Theory of mind (ToM)	21
2.1.3	Theory-theory of mind (TToM)	23
2.1.4	Simulation theory of mind (SToM)	26
2.1.5	Mentalizing and folk psychological practice	28
2.1.6	Summary	29
3	Language and TToM: The NOS paradigm and children's refer- ence to [desire]_p	31

3.1	Introduction	31
3.2	The adequacy of the NOS approach	32
3.3	The data so far	33
3.4	The philosophy of statistical tests in the NOS paradigm	46
3.5	Limitations of the paradigm	51
4	The corpus	56
4.1	The Data I: The corpus from non-autistic children	56
4.1.1	The corpus by age	56
4.1.2	The corpus by Mean Length of Utterance (MLU)	60
4.2	The Data II: The corpus from autistic children	63
4.2.1	The corpus by Mean Length of Utterance (MLU)	64
4.3	Identifying ‘want’ utterances	66
5	Desire in normal monolingual English children	73
5.1	Identifying genuine reference to $[\text{desire}]_p$	73
5.1.1	Bartsch and Wellman’s categorization scheme: some ex- amples, some problems, and a rationale for its continued use	77
5.2	Study 1: Replicating Bartsch & Wellman’s three hypotheses	85

5.2.1	Age as the independent variable	86
5.2.2	Self versus other by age	94
5.3	Study 2: Extending Bartsch & Wellman	99
5.3.1	The issue with MLU (mean length of utterance)	99
5.3.2	MLU as the independent variable	101
5.3.3	Self versus other by MLU	104
5.3.4	Communicative uses of ‘want’ as a function of age and mean length of utterance	107
5.3.5	Desires And Their Intentional Objects as a function of age and mean length of utterance	113
5.4	General Conclusions	130
6	Desire in autistic children	134
6.1	Examining Bartsch & Wellman’s three hypotheses with autistic children	135
6.2	Study 1: GPRDs in autistic children	136
6.2.1	Coding GPRDs: MLU as the independent variable	137
6.2.2	Self versus other	140
6.2.3	Desire and communicative uses	142
6.2.4	Desire and their intentional objects	144

6.3	General conclusions	147
7	Comparing normally developing and autistic children	149
7.1	A comparison of children's reference to [desire] _p by means of 'want'	149
7.2	Comparison 1: GPRDs	150
7.3	Comparison 2: Self versus other	152
7.4	Comparison 3: GPRDs versus communicative uses of 'want'	155
7.5	Comparison 4: the objects of GPRDs	159
7.6	Discussion and conclusions	162
8	Conclusions, speculations and future work	167
8.1	H1: Making sense of all the results	167
8.2	What is the representational status of GPRDs?	170
8.2.1	Problems, future work and last thoughts	179
	References	184

LIST OF FIGURES

1.1	Engaging in folk psychological practice	1
1.2	The different studies carried out to test the validity of H1 using the NOS paradigm	13
2.1	Conceptual map of ToM research	19
3.1	Genuine psychological references to desire by clinical condition (Autistic Vs. Down's Syndrome) as frequency of occurrence ev- ery 1000 utterances, From Tager-Flusberg (1992)	44
4.1	A comparison between the frequency of occurrence of the term 'want' in the non-autistic versus autistic populations by MLU (er- ror bars are Standard Deviations	72
5.1	GPRDs by age as a percentage of the total number of utterances for that child at that age. (Error bars are standard errors.)	90
5.2	Individual differences of GPRD as a percentage of total utterances for that child at that age	92
5.3	Self versus other as a percentage of GPRD utterances. (Error bars are standard errors.)	96

5.4	Mean GPRDs as a percentage of total utterances by child by mean length of utterance (MLU). (Error bars are standard errors.) . . .	103
5.5	Self versus other as a percentage of GPRD utterances by MLU stage. (Error bars are standard errors.)	105
5.6	Frequency of the communicative uses of 'want' as a percentage of total utterance and as a function of age. (Error bars are standard errors.)	109
5.7	Frequency of the different communicative uses of 'want' as a percentage of total utterance and as a function of MLU compared to GPRDs. (Error bars are standard errors.)	110
5.8	The different kind of desire objects as a percentage of the total number of GPRD utterances for that child at that age (Error bars have been omitted for readability)	126
5.9	Physical objects (PHY) compared to states of affairs plus actions (SOA+ACT) as a percentage of the total number of utterances for each child at each particular age.	128
6.1	GPRDs by MLU as a percentage of the total number of utterances for that autistic child at that MLU stage	140
6.2	Frequency of GPRDs to self versus other in children with autism as a function of MLU	142

6.3	Average frequency of the different communicative uses as a percentage of total utterances as a function of MLU compared to age for autistic children (Error bars are standard errors.)	143
6.4	Average frequency of the different objects as a percentage of total utterances as a function of MLU for autistic children. (Error bars omitted for readability.)	146
7.1	Frequency of talk about genuine desires for both autistic and normally developing children as a percentage of total utterances and as a function of MLU	151
7.2	Frequency of talk about genuine desires for self versus other for both autistic and normally developing children as a percentage of total utterances and as a function of MLU. Figure A shows a comparison of self between autistic and normally developing children and Figure B shows a comparison of desire terms about others. . .	154
7.3	Frequency of talk of the different communicative uses of 'want' for both autistic and normally developing children as a function of MLU. (Error bars are standard errors)	156
7.4	Frequency of talk of the different objects of desire for both autistic and normally developing children as a function of MLU	160

LIST OF TABLES

4.1	Information on the samples.	58
4.2	Total utterances available from each child by age in months. . . .	59
4.3	Total number of utterances available from each autistic child at each MLU stage.	62
4.4	Information on the samples from the autistic children.	65
4.5	Total number of utterances available from each autistic child at each MLU stage.	66
4.6	Total 'want' utterances available from each of the non-autistic group by AGE stage.	68
4.7	Total 'want' utterances available from each of the non-autistic group by MLU stage.	70
4.8	Total 'want' utterances available from each of the autistic group by MLU stage.	71
5.1	Total GPRD utterances available from each child by AGE.	89
5.2	Ages at first available transcript, first (and third) use of desire terms and first (and third) genuine reference to desire.	93
5.3	Total number of utterances and age at first (and third) genuine reference to the desires of self versus other	97

5.4	Total GPRD utterances available from each child by MLU.	102
5.5	Number of desire utterances for physical objects (PHYs), actions (ACTs), states of affairs (SOAs) and GPRDs with no object in them (NOBs) by age	123
5.6	Mean and standard deviations of the frequency of physical objects (PHYs), actions (ACTs) and states of affairs (SOAs)	124
6.1	Number of GPRD utterances available from each autistic child by MLU stage.	139

ACKNOWLEDGMENTS

I have many people to thank. I really wouldn't know how to start or who to start with. I guess that, if I look back on my graduate student years, four people have been instrumental in my formation. In no particular order, I would like to thank Prof. Andrew Brook, who was always, but really always there when I needed either (or sometimes both) academic or/and personal advice and help. He always knew what to tell me so I would not suffer the bureaucracy and other nuances of academia and its different peoples. One sticks out, particularly, and although the context need not concern us here, I would like to put it down in writing, for me to always remember when the time comes: "But Andy, what do I do? He'll be so angry at me!", Leo said, extremely worried. Andy may have thought for a second (though Leo doesn't remember) and replied: "He'll get over it." And magically, I guess, Leo wasn't worried anymore.

I would also like to thank Prof. John Logan. I would not have been able to make it without him. His constant support and also always being there for me helped me carry on. There was not even once that he was not available for me, even when he was insanely busy having just taken office as the chair of the Psychology Dept. at Carleton. It really was amazing to see how he always went the extra mile to help me. He never (not once, not a little bit) failed to provide me with incredibly sound academic advice on my dissertation and was never tired to go over a section or a chapter for the third time if it need be. He also used

to say things that helped me lots. The phrase I remember from him (again, the context should not concern us here) is that I always have to “keep a stiff upper lip.”. I try John, I swear I always try.

I would like to thank Bruno Emond, with whom I’ve had hundreds of hours of very interesting academic discussions and with whom I share many interests. He has been the catalyst of various loose-end ideas and his support has always been unfailing. I owe him a great deal.

I would like to thank Karen Bartsch. Her book and research have been a real inspiration for me. I know this is a minute contribution, compared to what she did in her investigations, but I hope my own investigations, inspired by hers, serves both to the progress of science and as a recognition of her work.

My thanks go also to Colleen Fulton and Lianne Dubreuil. Both have been the kindest at times of desperation. I have to thank Colleen particularly. I don’t know how many hugs I’ve received from her and how many treats... I still eat all of her candy when I’m around.

Thanks go also to my fellows in the Cognitive Science programme, with whom I’ve spent many, many hours. Particularly, and in no specific order, my thanks go to Matthew Rutledge-Taylor, Sam Scott, Terry Stewart, Jerzy Jarmasz, Kamilla Johannsdottir and Diana Destefano.

Although only for the last year and a half (out of six!), I’ve had the privilege of spending... somebody might say *too* many hours with my fellow Argentinean

friends, Juan Carlos Afonso and Ezequiel Glinsky. Realmente, no se que hubiera hecho sin ustedes, chochamus.

I would like to thank Fundacion YPF and Fundacion Antorchas, who helped me for many years. Part of this dissertation has been funded by them.

Finally, I'd like to thank my family, Enrique, Velia, Martin y Martha for, well, all. Everything. Los quiero muchisimo a todos. And last but not least, I'd like to thank my wife, Loreto, who's giving me everything I'd ever wanted. This thesis is, of course, dedicated to her. Gracias, peces.

CHAPTER 1

Introduction

1.1 Introducing H1

Here's a comic strip from the Overboard series by Chip Dunham,

Figure 1.1: Engaging in folk psychological practice

But it was not really that Louie the dog did not want a treat. In fact, knowing dogs, “he”[sic] probably did. However, for some reason, the character offering the treat (let us call him Ezekiel) decides to explain Louie’s behavior (not coming to get the treat) as a matter of Louie’s (nonexistent) desire to eat this treat (“He probably just didn’t want a treat”). Moreover, for some reason, Ezekiel opposes Nate’s offering a “knowledge” explanation of the dog’s behavior (“Oh, he didn’t

know, Nate. Don't give him so much credit"), seemingly assuming it is too much for the dog to "know" about their scheme, when in fact the dog did *seem* to know (hence the punch-line of the comic, should anyone have failed to get it).

Despite the potentially never-ending discussion of what the dog in the comic strip (or dogs at large) may or may not know or want, what is interesting is Ezekiel's use of terms such as "know" and 'want' to explain Louie's behavior; namely, his leaving the scene in the presence of an otherwise perfectly delicious treat.

Being able to provide these "folk psychological" explanations is a benchmark of human cognition, and the general topic of this investigation. In particular, the work reported here will focus on the development of one of the terms used in these explanations; namely, the term 'want', *qua* a proxy for the *concept* of desire, *qua*, in turn, a concept referring to the *mental state* of desire.

Technical details aside, the studies reported here build on a puzzling hypothesis put forward by many different researchers in many different ways: the hypothesis (henceforth, H1) that children make genuine psychological references to desire at around the second year of life, at least a year before belief has made its way into the child's conception of mind (see Tan & Harris, 1991; Wellman & Bartsch, 1994; Bartsch & Wellman, 1995; Harris, 1996; Astington, 2001; Ferres, 2003; among many others).

If H1 is true, as evidence coming from the developmental area suggests, then

it suggests several important implications. The most intriguing one is that desires should then differ from beliefs in their representational status, a claim many developmental psychologists have been happy to embrace (see Wellman, 1990; Perner, 1991a; Perner, 1991b) but that many philosophers, in turn, mostly reject. In modern philosophy of mind, propositional attitudes, beliefs and desires are usually grouped together, at least when it comes to their basic representational status. Although philosophers come in different flavors and it would be a real crime to generalize¹, it is relatively uncontroversial that beliefs and desires are both “representational states”, a relation between people and internal representations (Fodor, 1975, 1987, 1990, 1991, 1992).

Very little is clear in the developmental literature about this issue: are desires representational at younger ages (Wellman, 1990; Bartsch & Wellman, 1995) or something else (Perner, 1991a), are they metarepresentational as belief is assumed to be (Baron-Cohen, 1995) or still something else at older ages? The discussion in fact is very much alive (see Astington, 2001, for example). But, is H1 really true? For example, could it be the case that results obtained in desire research is an artifact of the different research paradigms, or, at least, of the one assumed in these pages? The purpose of this work is to explore this and other relevant questions derived from this issue.

¹cf. for example Fodor (1987) and Searle (1983) for a subtle difference about the representational status of propositional attitudes.

1.2 Introducing the project: Testing H1

This investigation has, as its main objective, to empirically test H1 against several different potential arguments against it. A note of caution is due: the studies reported here are not supposed to prove or disprove H1 beyond any shred of doubt, even if the results show exactly what is expected. Instead, the studies reported are meant to clear the way for other, less naturalistic, investigations of the child's use of the concept of desire.

Logically, the first stage of this project will be to somehow test H1 against empirical data. If H1 ends up being false, then there is, in fact, no project at all and the discussion about it in both the developmental and the philosophical literature has been, at best, misguided or, at worst, in vain.

A brief digression. Given its characteristics, H1 is difficult to study without resorting to methodologies using language as the main dependent measure. Actually, some experiments have been devised to behaviorally test this hypothesis (Astington & Gopnik, 1991) and although they so far yield results consistent with H1, they fall prey to a Dennettian criticism (Dennett, 1978a) in which the philosopher argues that, in this field of research at least, the only actions that meet experimental conditions naturally are linguistic acts (rather than pointing or other sorts of non-linguistic behavior). Given the characteristics of language, naturally occurring speech (as opposed to elicitation or act-out in laboratory settings) would be the best possible source of knowledge about this area. Of course,

this is not without perils.

Even if H1, studied by means of a given linguistic measure, appears to be true, then it may very well be that the phenomenon could be given a linguistic explanation, rather than a developmental one. In other words, if a given measure of linguistic complexity explains the staging of the development of the concept of desire better than age, then it may be the case that children are simply being constrained by their linguistic prowess rather than their conceptual knowledge. If this is the case, then H1 is an artifact of the instrument used to tap into children's knowledge about a given concept. If this is the case, then once again there is no project at all, with any investigation of H1 by means of language for nought.

Now, what if H1 cannot be given a mere linguistic explanation? It may be the case that at early stages of linguistic development (which will be highly correlated with age anyway), children already have the linguistic tools for expressing desires reliably and they do not... or it could be that at early stages they do not have the linguistic complexity and still are able to reliably talk about the mental state of desires. Should we now conclude that H1 should be true and worry about its most pressing implications? The issue is still far from clear.

Of course, for this research to be informative researchers should be able to single out when children are using a desire word without having the concept of desire and when they are using a desire word to refer to the concept of desire. Given this, categories have been devised to differentiate between (and group

together) those instances when children use desire terms (mere words) in a way that makes it evident that they have a real grasp of [desire]_p and those instances which do not evidence this same grasp (Shatz, Wellman, & Silber, 1983; Wellman & Bartsch, 1994; Bartsch & Wellman, 1995). This “grasping” of the concept will be indicated by the ability to talk about one’s own and others’ desires, express finely differentiated kinds of desires clearly and so on. This is, of course, not always easy to do, as will be discussed below in page 7.

Also, putting together and then analyzing only the category that groups real reference to the concept of desire is not enough in itself. Suppose that a given grouping technique yields four mutually exclusive categories such that one and only one groups utterances that use desire terms standing proxy for the concept of desire *A* while the other three categories group other different uses of desire terms, say *B*, *C*, and *D*. Then it could be the case that at least one of the four categories that are not really desires, say *C*, follows the same developmental trajectory as the allegedly real desire category *A*. If so, then, it would be hard to conclude that what is being grouped as genuine desires are, in fact, genuine desires and nothing else. Thus, if there are other seemingly unrelated categories that develop in the same way as the allegedly genuine desire category, then one may conclude that the instrument is not sensitive enough to individuate desires as opposed to utterances fulfilling other functions in the corpus of data, such as mere use of desire words in a formulaic way or a repetition of an adult utterance.

If so, again, we have no reason to suppose H1 is true. The instrument is not really up to the challenge of indicating the difference between terms referring to desire states and the same terms being used for something else.

But now suppose the instrument for studying H1 passes this test of grouping genuine desire talk (talk that makes evident that the child has grasped the concept of desire) and nothing else. Suppose further, and quite optimistically, that plotting the development of the use of desire terms to refer to desires yields a very different pattern of development than the other non-desire categories. Is H1 now free from methodological issues? The short answer is no, but taking together all the measures discussed so far, it would be hard indeed to prove the opposite; namely, that H1 is false.

Two more problems remain: first, there is still the possibility that the scheme cross-codes utterances that indicate a perfectly good grasp of the concept of desire. In other words, mistakenly assigning utterances that show the child is using the word 'want' as a proxy for the concept of desire to the wrong category. This could be the case because in Bartsch and Wellman (1995), the study that most closely resembles the present investigation, they use a scheme in which the only test of having grasped the concept is an ability to *refer to* the desire via the use of the word 'want' and some other words that stand for the same concept. Second, there may be other ways to indicate grasp of the concept of desire than linguistically sophisticated utterances. Specifically, there is the problem that

many utterances indicating grasp of the concept of desire may not use the word 'desire' or cognates at all. These issues are important and will be taken up in Chapter 5.

Given all this, the first goal of this investigation is to go through all the requirements discussed above and analyze whether H1 should be taken seriously, at least when it comes to the methodology assumed in these pages; namely, one in which language is the main dependent measure. Since, as will be shown, it is highly likely that H1 is still true even after all the reservations imposed on it, the second part of this work will be devoted to exploring the theoretical implications of H1 being true.

1.3 The road map

As said in the previous section, the first step of this research is to replicate one of the studies that most clearly support H1. In this case -and for many reasons- efforts will be made to replicate part of the work reported in Bartsch and Wellman (1995); focusing, of course, on the development of the concept of desire.

Chapter 5 will carry out the replication itself, together with relevant studies. However, to achieve this purpose, many intermediate steps are needed. The following is the road map for the analyses described in this research. Chapter 2 will deal with definitional matters; in other words, with proposing a general, articulated framework to deal with folk psychology, some of the assumptions

made and related issues.

The next chapter, Chapter 3, will critically assess the advantages and disadvantages of what will be termed the Naturally Occurring Speech (henceforth, NOS) paradigm as the tool for providing support or falsifying H1. Basically, this paradigm supposes that words stand as proxy for concepts children have been able to acquire at a given point in time. In other words, if a child uses a certain word in a certain context to refer to desire itself, then one might be warranted in believing that at that point they know the concept, simply because they are referring to. A description of all the studies using this paradigm will be analyzed, together with a commentary on the difficulty of carrying out inferential statistics on the results and a possible solution to the problem.

Chapter 4, in turn, will present a detailed revision of the characteristics of the corpus of data that have been analyzed in this investigation. It will present the data in various perspectives to preview the different sorts of comparisons that will be used later in the dissertation. Also, some of the mechanics of identifying utterances and coding them will be made explicit in the spirit of making the investigation and its results amenable to replication by other researchers.

Foreshadowing the results, H1 is given further empirical support based on analyses from a different database of children's language from the ones used up to now in the literature and coded by two different individuals. But as discussed above, several tests are needed to confirm the reliability of the result obtained

in the replication. This task is also taken up in chapter 5 where H1 is put to the tests described above in section 1.2; namely, (a) whether H1 is an artefact of using language as the main instrument (analyses using MLU –mean length of utterance– are carried out to evaluate this issue); (b) whether utterances coded as genuine desires show no distinctive pattern of development compared to other uses of the term ‘want’, and whether (c) taking a more conservative stance about what is considered to be a genuine desire shows a pattern of development similar to the more lax version. One example of this would be plotting only children’s use of ‘want’ with desire objects such as actions and states of affairs (i.e. desires for things that are not directly there and have to be represented somehow), instead of physical objects (which may be visually accessible to the child and, according to certain researchers, not necessarily mentally represented (Bartsch & Wellman, 1995)).

All these tests on H1 yield roughly the picture expected if H1 were true: MLU studies show that even at the lowest MLU, children find a way to refer to their own and someone else’s desires. Plotting reference to desire (as understood in these pages) across time shows a different developmental trend compared to mere communicative uses (such as behavioral requests, direct repetitions and other non-desire utterances): while the concept of desire seems to undergo some development across time, other uses of the term ‘want’ stay relatively unchanged. Lastly, being more strict about what is taken to be a desire (considering only non-

physical propositional objects such as actions and states of affairs) yields, in fact, the same developmental pattern as the laxer coding scheme.

Chapter 6 replicates and extends studies by Tager-Flusberg (1993) on a sample of autistic children. These studies were selected for replication because this population is said to have some metarepresentational deficits, at least when dealing with the concept of belief (Baron-Cohen, 1995, 1993; Bryson, Clark, & Smith, 1988). Without going into detail, this meta-representation deficit can be defined as these children's inability to entertain a proposition (about someone else) that stands for a counterfactual state of affairs (Davies & Stone, 1995c). The aim of these analyses in the present investigation is to see whether there is a developmental trend of the concept of desire in children with autism and also to lay the foundations for the next chapter, which introduces a comparison between a sample of normally developing children and a sample of children with autism.

Chapter 7 presents an in-depth comparison between a sample of children with autism and a sample of normally developing children that, to the best of the author's knowledge, has not been investigated elsewhere. Analyses of the data suggest no differences between normally developing and autistic children's abilities.

The last chapter of this work is devoted to providing a summary of the empirical results and their possible implications. This last chapter will also discuss the strengths and weaknesses of this approach to TToM research and consider

the work that remains to be done.

1.4 Overview of the studies

Table 1.2 shows the different analyses carried out in the present investigation and how they fit into the general framework of the area. The analyses can be divided into two categories. The first involves normally developing children (NDC) and the second involves autistic children. Basically, the same battery of studies were carried out on both populations, but they constitute different contributions to the field. Some of them are replications, others are extensions of previous research while still others are completely new analyses.

Each of the analyses and its hypothesis will be discussed in much more detail later on, but I will outline them here. In all, there were four main analyses carried out on each of the two populations. The first involved coding utterances as genuine desires, as Bartsch *et al.* did in their seminal studies. A second step was to compare what were identified as genuine desire references with other utterances to see whether they could be characterized as something different or whether in fact they underwent an similar development. The third analysis was aimed at being much more conservative in what was taken to be a genuine reference to desire. This was done by considering only those desire utterances that had, as their object of desire, something that could not be seen or touched directly, such as actions and states of affairs. Lastly, but not least, an analysis

Part	Analysis	Version	Replication	Extension	New Analysis
Part I - NDC	Coding	AGE	X		
		MLU		X	
	Comm Uses	AGE		X	
		MLU			X
	Objects	AGE			X
		MLU			X
	Self vs. Other	AGE		X	
		MLU			X
Part II - Autistic children	Coding	AGE	N/A		
		MLU	X		
	Comm Uses	AGE	N/A		
		MLU		X	
	Objects	AGE	N/A		
		MLU			X
	Self vs. Other	AGE	N/A		
		MLU			X

Figure 1.2: The different studies carried out to test the validity of H1 using the NOS paradigm

was carried out to see whether there was also a developmental trend in children's reference to their own versus other people's desires. In particular, the Self versus Other study is, I believe, a unique contribution to the study of how the concept of desire develops under many possible contexts. This has, in my view, strong implications for theoretical issues raised by the empirical literature, such as representational/metarepresentational issues.

All these analyses were carried out on the normally developing population in terms of both chronological age and mean length of utterance (MLU) to try

to answer some of the questions about the linguistic side of H1. On the other hand, only MLU made sense for the autistic population given this population's individual differences when it comes to language acquisition.

The Replication/Extension/New Analysis categories are intended to refer to the two base analysis addressed in this dissertation; namely, Bartsch and Wellman (1995) and Tager-Flusberg (1993). By replication I mean that the studies were carried out in one of these two papers (or indeed in other papers dealing with this issue and using this methodology) but using different populations. By extension I mean looking at the data in several ways that were not present in the original studies. An extension would be to look at Bartsch and Wellman's developmental data plotted as a function of MLU, for example. By a new analysis, I mean a novel way to look at how certain aspects of the child's use of 'want' develops; such as, for instance, new ways to divide the categories proposed by Bartsch and Wellman into other different, hopefully more telling, categories.

CHAPTER 2

Theory of mind Research

2.1 Charting the conceptual map: Situating desires in the general Theory-Theory of Mind framework

The main goals of this section are twofold: (a) to try to define the different foundational concepts that I will be working with and (b) to situate the concept of desire in the conceptual map resulting from the definitions in (a). Unfortunately, there is no comprehensive, consistent theoretical framework to use in order to define any of the concepts satisfactorily. Most researchers assume one definition or another of these concepts when carrying out their agendas, but they rarely make them explicit in their writings and this has led to a chaotic use of terms across researchers. This is, however, not the place to go into an in-depth *a priori* consideration of the definitions of the concepts used by researchers. Therefore, the task at hand is to define, as carefully as possible, all the theoretical concepts to be used in these pages, while considering the general ideas that can be decanted from the literature to-date. This is not an easy task given the different shades of

basic concepts such as belief, false-belief, desire, theory, theory-theory of mind, etc. that researchers use. So a few preliminary definitions are due.

First, the reader should be warned that there are two kinds of concepts that will appear in the course of this work. On the one hand, there are formal scientific concepts that psychologists and philosophers use when trying to explain the human ability to describe, explain, and predict/expect behavior: these scientific concepts (henceforth, $[\text{Concept}]_s$, the subindex s standing for a scientific concept, discovered by a formal scientific discipline) include those identified by the words ‘theory’ and ‘folk psychology’, and even ‘desire’ and ‘beliefs’ for example. These concepts are better thought of as the researcher’s toolbox that he or she uses to systematically and scientifically reflect on the human ability to calculate their own and other people’s behaviors. On the other hand, there is a second kind of concepts: pre-scientific concepts (henceforth, $[\text{Concept}]_p$, the subindex p standing for pre-scientific, a folk concept) that humans know (perhaps tacitly) and make use of when describing, explaining and predicting behavior “in the wild”. These concepts include those of $[\text{belief}]_p$ and $[\text{desire}]_p$ and the rest of the so-called propositional attitudes (see, for example, Davies & Stone, 1995c) and are better thought of as the lay person’s toolbox that he or she uses to calculate their own and other people’s behaviors.

For example, the word ‘desire’ in these pages could be used to refer to three different things: (1) the mental state per se; (2) to $[\text{desire}]_p$ the concept that the

child uses to express and refer to (1) and what the scientist is trying to give an account of; and (3) [desire]_s, the scientist's concept for the child's concept (i.e. 2). Every use of these kinds of words will be marked to make its meaning clear, unless they manifest in a context which makes its use unambiguous.

[Desire]_s, as taken in these pages, is just one of the theoretical entities postulated by one of the hypotheses of theory of mind research; namely, the theory-theory of mind hypothesis. It is important to know what all these terms mean before going into the characteristics of desires per se.

Figure 2.1 shows a general conceptual (Cs) map of the area. At the top lies Folk Psychology (henceforth, FP), a natural phenomenon that is the object of study of Theory of Mind (henceforth, ToM), which is a research field with the specific agenda of providing a scientific account of FP. There are two scientific accounts of FP that have been proposed by ToM so far: the theory-theory of mind hypothesis (henceforth, TToM) and the simulation-theory of mind hypothesis (henceforth, SToM).

On the one hand, TToM postulates the existence of two components: (a) a conceptual repertoire that includes [beliefs, desires]_p and the rest of the propositional attitudes and (b) some computational mechanisms (like the practical syllogism, see Fodor, 1992 or Fodor, 1987). These computational mechanisms have the function of putting the aforementioned conceptual repertoire to work. The computational processes that make use of beliefs, desires and the rest of

the propositional attitudes yield generalizations about behavior which, taken together, crystallize in a naïve theory about the psychology of people, hence the name *theory*-theory of mind. One can think of this naïve theory (making use of naïve concepts) as *knowledge* about the psychology of people in this case.

On the other hand, for Simulation Theory of Mind (henceforth, SToM) to work, it only needs some computational mechanisms that would help an agent *A* to imaginatively identify with the organism it is trying to explain the behavior of. There is no need for the foundational concepts that are required for the TToM. For SToM, it is enough to entertain thoughts about self and then project those thoughts to the mind of other people. The resulting product of this mechanisms is a change in *skill*, rather than in knowledge.

2.1.1 Folk psychology (FP)

Rudder-Baker (1999) writes:

[...] there are two different things that "folk psychology" has come to mean, and they are not always distinguished: ([FP]1) commonsense psychology that explains human behavior in terms of beliefs, desires, intentions, expectations, preferences, hopes, fears, and so on; ([FP]2) an interpretation of such everyday explanations as part of a folk theory, comprising a network of generalizations employing concepts like belief, desire, and so on. The second definition [...] is a philosophical

Figure 2.1: Conceptual map of ToM research

account of the first. (Rudder-Baker, 1999)

Humans provide descriptions like “Mika Pooch is running to get her tennis ball”, explanations like “the character on the left is cheating because he wants to earn easy money” and even predictions “Mika Pooch will cry tonight because I lost her tennis ball”. This natural ability is what will be called folk psychology in this investigation. Exactly how humans do it or how researchers might end up explaining this phenomenon is irrelevant in defining folk psychology. In fact, for all it matters, humans may do it by either ascribing functional units like [beliefs]_p and [desires]_p or by analyzing strange covert behavioral cues like a blink occurring on Fridays in February, or the tone of voice as Louie the dog does in the comic in Chapter 1. In turn, researchers might end up explaining it by postulating particular brain states, such as neuron number 236 firing at a particular rate and having an influence on nuclei *Y*, or by a set of enzymes acting together. Whatever is the case, what is relevant is that humans do it, one way or another.

For the purposes of this research, folk psychology will mean something roughly similar to Rudder-Baker (1999)’s first definition of folk psychology, with some provisos: i) FP will be understood simply as the human ability to describe, explain and predict (or expect) behavior and ii) also, these explanations, descriptions and predictions could be applied not only to human behavior (as Rudder-Baker, 1999 seems to imply), but also to the behavior of members of other species (like primates, cats, insects, etc.) and sometimes even things that perform some sort

of function, like computers or thermostats (see Dennett, 1978b, for an excellent discussion of some of these issues).

To emphasize even more: folk psychology will not be taken to mean how researchers, qua developmental psychologists or philosophers, explain the human ability to provide folk psychological descriptions, explanations and predictions of behavior, but merely the human (non-scientific) ability to provide them. This natural phenomenon can be taken as an object of study, of course, and the scientific endeavor that aims at giving a scientific account of folk psychology will be called Theory of Mind (research).

2.1.2 Theory of mind (ToM)

A brief digression. The word theory will appear in at least two contexts in these pages: on the one hand, in the context of a scientific endeavor that is concerned with explaining the human ability to ascribe psychological states to ourselves and others (FP), “theory of mind”. This use of the word theory implies a scientific account and it should be therefore characterized as [theory]_s. On the other hand, the word theory will be used in the context of a hypothesis in theory of mind research that characterizes human folk psychology as some sort of implicit pre-scientific (naïve) theory, which will be called the theory-theory of mind hypothesis (see section 2.1.3 below) as opposed to, for example, the simulation-theory of mind hypothesis. This use of the word ‘theory’ implies a non-scientific account

and it should be therefore characterized as [theory]_p. One of the general purposes of this research is to investigate one of the concepts that is present in this (naïve) theory, so a (scientific) [theory]_s of [desire]_p will be put forward.

It seems that in the developmental literature, theory of mind is just another name for Rudder-Baker's (1999) FP2 (see for example, Baron-Cohen, 1995; Wellman, 1990). Some researchers may be happy running these two concepts (FP and TToM) together under the phrase "folk psychology" (or indeed ToM sometimes!), as Rudder-Baker implies they do in the literature. However, this is quite confusing and they should be teased apart.

For developmental psychologists, accepting this "combo" definition of theory of mind "as is" would mean accepting that "an interpretation of such everyday explanations [is] part of a folk theory, comprising a network of generalizations employing concepts like beliefs, desires and so on" (Rudder-Baker, 1999). But this commits researchers to much more than some are prepared to commit to by this term if in search of a well-articulated scientific conceptual toolbox. Specifically, it commits investigators to an understanding of folk psychology as a theory (in the sense of Morton, 1980; Premack & Woodruff, 1978; Fodor, 1987, 1992). Rudder-Baker's view that theory of mind is a philosophical, scientific account of folk psychology (as defined above in section 2.1.1) seems right. But that account need not necessarily assume that folk psychology is "theoretical" in nature; that is, taking the form of a theory (see section 2.1.3 below).

The argument for this is as follows: suppose that we do take theory of mind to be FP2 in Rudder-Baker's (1999) sense. Then what happens to the alternative scientific account of folk psychology; namely, simulation theory of mind? Is it not a theory of mind itself? Carruthers and Smith (1996) seem to agree on this point. They write: "the term 'theory of mind' has come to designate a particular research domain, whose general goal is to provide an explanation of the ability [...] to explain and predict the behavior of both oneself and of other intelligent agents". So, obviously, we need to call FP2 something else, and I will call it the theory-theory of mind (hypothesis), to contrast it with the simulation-theory of mind (hypothesis). However, there is no doubt that they are both theories of mind. The concept [theory of mind]_p, therefore, will be taken to mean how we (scientifically) characterize folk psychology (as defined in section 2.1.1).

The reader should think of theory of mind, then, as the general scientific framework that aims at explaining FP or how minds understand and explain their own and other people's behavior, without committing to one particular view of how this understanding/explanation happens; that is, whether by theory (TToM) or off-line simulation (SToM).

2.1.3 Theory-theory of mind (TToM)

Basically, then, in the context of the present work, theory-theory of mind is more or less Rudder-Baker's FP2; namely, the (scientific) hypothesis that our

folk psychology is a theory, not unlike professional scientific theories (Wellman, 1990; Bartsch & Wellman, 1995; Perner, 1991b). It's theory-like in the sense that it postulates theoretical entities (like [belief]_p, [desire]_p and the rest of the propositional attitudes), laws which give rise to generalizations, and which, in turn, relate these mental states causally, either to one another, to external stimuli on the input side (the observation of a particular behavior) or to actions on the output side (the performance of a particular behavior) (Davies & Stone, 1995c). Also, for folk psychology to be a theory means more than simply being able to give a theoretical account of it (Blackburn, 1995). It also means that we have stored psychological knowledge in the form of a (naïve) theory about the psychology of people, as discussed above.

Here is an example: i) we see behaviors out there in the world; ii) on the basis of our folk psychology, we analyze those behaviors, and iii) we provide an explanation and probably predict/expect a future behavior. It is (ii) that takes the form of a theory, if we assume that folk psychology is a theory inside our heads. Take formal scientific theories (henceforth, FSTs) as an analogy. They seem to have, at least, the following characteristics: FST1) in explaining a phenomenon FSTs invoke abstract entities (quarks, phlogiston, etc.); FST2) FSTs invoke laws (causal laws) that govern the interaction of the entities invoked in (FST1), and lastly, FST3) on the basis of (FST1) and (FST2), FSTs provide explanations for the phenomenon in question and also build generalizations (forming a theory) that

make them amenable of being extrapolated to new evidence. Taking (FST1)-(FST3) into account, theory-theorists would like to tell the following story about folk psychological explanations (henceforth, FPEs): consider that, like FSTs, FPE1) in explaining behavior (the phenomenon) FPEs seem to invoke abstract, unobservable entities like [beliefs]_p, [desires]_p and the rest of the propositional attitudes; FPE2) FPEs seem to invoke causal laws that govern the interaction of the entities in (FPE1), for example: IF X THEN P ; and lastly FPE3) on the basis of (FPE1) and (FPE2), FPEs provides explanations of the phenomenon at stake (behavior) and build generalizations that taken together would form a theory of the phenomenon in question (behavior in this case) and that could in principle explain new evidence. For example, consider the well-known ‘practical syllogism’ (see Fodor, 1987; Wellman, 1990; Wellman & Gopnik, 1995), as a particular case of FPE generalization: IF $\{x \text{ wants that } P\}$ AND $\{x \text{ believes that } \sim P \text{ UNLESS } Q\}$ THEN $\{x \text{ tries to bring it about that } Q\}$. Let’s analyze a ‘real-life’ example: the phenomenon (behavior): Behavior1) Loreto drinks hot water (not tea or coffee) during class; Behavior2) Loreto told me: “I like hot drinks but I want to sleep at night”; Therefore: If Loreto wants a hot drink during class and she believes that tea or coffee would prevent her from having a good night’s sleep, then she will avoid tea or coffee and drink hot water instead. This a plausible explanation, reached by invoking (FPE1)–(FPE3).

Basically, then, the main idea behind the theory-theory of mind hypothesis

is that folk psychology could be explained by the existence of a (naïve) theory about psychology, a theory that resides (at least, as far as we know) inside the human mind. In similar words, it means that the way we explain and predict (expect) behavior is by having/using a theory, not unlike a formal scientific theory, which is found right inside our heads. To use the fashionable expression in developmental psychology, we are untrained psychologists (as we are untrained physicists, physicians, biologists, etc.).

2.1.4 Simulation theory of mind (SToM)

Simulation-theory of mind, in turn, postulates that there is not the need for an agent A to entertain thoughts that include the concepts that are inherent in TToM (such as those of $[\text{belief}]_p$ and $[\text{desire}]_p$) while trying to explain, describe or predict behavior. In a simulation-theorist's view, the explaining agent needs to "identify" with the organism whose behavior A is trying to explain and imagine a counterfactual situation (Goldman, 1989, 1993, 1995). In other words, according to the SToM, there is no need for an agent to possess the full-blown concept of $[\text{belief}]_p$ or mutatis mutandis, that of $[\text{desire}]_p$, but merely to have beliefs and desires (Davies & Stone, 1995a, 1995b; Gordon, 1995a, 1995b). Thus, the ability ascribe mental states is not conclusive in SToM as it is in TToM.

Let me exemplify once again. As in TToM, i) we see behaviors out there in the world; ii) on the basis of our folk psychology, we analyze those behaviors, and

iii) we provide an explanation and probably predict/expect a future behavior. It is (ii) that takes the form of some ability, if we assume that folk psychology is a skill inside our heads. Here, the agent trying to explain a behavior is required to merely have thoughts of the following kind:

[1] P

while P may stand for whatever proposition *I believe that there's a dog in my backyard* or *I want to have a grey Nissan Pathfinder*. The difference between SToM and TToM is that to ascribe a thought to someone else is to imaginatively think from the other organism's point of view. The agent has to actually put itself in the place of the organism that it is trying to explain the behavior of, thus resulting, loosely, in a thought of the following kind:

[2] I [Sally] believe that there is a dog in my [her] backyard.

The SToM view has never been satisfactorily worked out to be as comprehensive as the TToM hypothesis in its psychological grounds (Harris, 1995a, 1995b). The philosophical literature provides it with relatively strong foundations (Stitch & Nichols, 1995b, 1995a; Perner & Howes, 1995), but very few of its implications have been found to be coherent with the evidence so far (Astington & Gopnik, 1991; Astington, Harris, & Olson, 1988; Astington, 2001; Fodor, 1990). Thus, the present investigation raises and discusses a problem in the TToM framework.

The same problem could arise in the SToM view, but this will be left for future work. Giving an account of SToM and analyzing the data in the light of this hypothesis makes enough work for another investigation and the issue will not be taken up again.

2.1.5 Mentalizing and folk psychological practice

Finally, it seems intuitive that there is a difference between possessing the (naïve) theory that gives rise to our ability to explain behavior and actually using this body of knowledge. The actual use of folk psychology (in the sense discussed in section 2.1.1 above) in a particular task and a particular situation, that is, actually engaging in the task of describing, explaining and predicting (expecting) behavior, I will call, following Davies and Stone, 1995c, folk psychological practice.

The developmental literature refers to this practice with other terms. For example, some researchers call it *mentalizing* (Frith, 1989), Baron-Cohen (1995 and elsewhere) calls it *mindreading*. I prefer Davies and Stone's definition mainly because the concept (in the [Concept]_s sense) of mentalizing is, at times, either vague or ambiguous. We can actually mentalize about many things, not only about the psychology of other people. To be sure, this is no more than a definitional matter and will not be explored further.

It is important to realize that the present investigation is indeed about FP

and TToM (the human competence to provide explanations), but that the way we have access to that competence is by children's use of FP and TToM, their performance, to borrow further on the linguistic analogy (Chomsky, 1965). Folk psychological practice is exactly this performance.

2.1.6 Summary

This chapter has, I hope, clarified and articulated the foundational concepts that will be used in this dissertation. Here is a brief terminological summary:

- a) **folk psychology** refers to the human ability to describe, explain and predict behavior and action;
- b) **theory of mind** refers to a research agenda that aims at giving a scientific account of (a);
- c) **theory-theory of mind** refers to a hypothesis of (b) to explain (a), which states that (a) is a naïve theory about the psychology of people;
- d) **simulation-theory of mind** refers to a hypothesis of (b) to explain (a), which states that (a) is a not necessarily a naïve theory about the psychology of people, but an ability that allows agents to ascribe mental states by imaginatively identifying with them;
- e) **folk psychological practice** refers to the act of putting (a) to actual use in a given situation.

In TToM, then, $[\text{desires}]_p$ are used in a $[\text{theory}]_p$ as one of the ubiquitous concepts humans make use of when engaging in folk psychological practice. $[\text{desire}]_s$, in turn, is one of the concepts of what researchers call the TToM hypothesis of ToM research. Zooming in on the “conceptual repertoire” branch of the tree, it can be seen that this node is further subdivided into the concepts of $[\text{belief}]_s$, $[\text{desire}]_s$ and the rest of the propositional attitudes $[\text{yearn, longing, want}]_s$, etc. (Fodor, 1987; Davies & Stone, 1995c; Stich, 1985). Within the ToM framework, then, and the TToM hypothesis of ToM in particular, this investigation focuses on the problem of the development of one of the concepts that pertains to TToM; namely, the development of $[\text{desire}]_p$. More specifically, whether the analyses carried out by others so far portray the correct picture of the development of desire talk, given several different empirical tests.

CHAPTER 3

Language and TToM: The NOS paradigm and children's reference to [desire]_p

3.1 Introduction

The goal of the present project is to investigate children's capacity to explicitly impute and express their own mental states and/or those of others; particularly reference to and expression of the child's own and others' desires. In order to answer important questions in this general framework, I have chosen to use what I call a natural-occurring-speech paradigm (or NOS paradigm, for short). Basically, this methodology involves building an as-large-as-possible database of child utterances at different points in their development and analyzing them in the light of hypotheses about mental state talk. The purpose of this section is, then, to discuss this paradigm, its results to-date, and to address its potential problems.

3.2 The adequacy of the NOS approach

The vast majority of research in TToM has been carried out using language in one way or another as the main instrument to tap into the child's (or indeed the adult's) conception and knowledge of the mind (Bartsch & Wellman, 1995; Wellman et al., 2001). It is difficult to imagine how else this kind of research could be potentially carried out. Some attempts have been made to use some other communicative means, such as pointing and the like (Astington & Gopnik, 1991), but they are subject to the same criticisms raised by Dennett (1978a) against the cognitive ethology literature; namely, that without real language and unambiguous intentional communicative acts, it is extremely hard to know whether a given communicative act is or is not such.

More constrained experimental techniques also involve many other restricting factors such as instructing participants, the presentation of experimental material, instructions to carry out the task, etc. (Astington & Gopnik, 1991). These and many other limitations render the environment rather artificial, to put it mildly. Thus, given how this agenda relies on language as its main instrument, it stands to reason that one of the best ways to carry on with this is to use language in its rawest form; that is, as it occurs naturally; or as I have referred to it earlier, in the wild.

Basically, the NOS paradigm has characteristics opposite to laboratory experimentation: it does not involve training of any kind (the sampling occurs without

the child even realizing that he or she is being recorded, such as in Wells (1981), for example. There is no need for instructing the participant about what the responses should be or when to speak and no materials need to be presented. The NOS paradigm systematically collects data at periodic intervals (say once every hour) for a set amount of time (10 minutes). Thus, there is no right or wrong thing to say, and no particular set of materials to present.

It is also known that taking children out of their familiar context and asking them to perform alien tasks degrades what they are able to show as their competencies. In the NOS paradigm, the child's context is her usual, familiar context, with not even an unfamiliar person (the experimenter) present in the room. In certain cases, like the samples chosen for this investigation, the microphone recording the child's speech is set to start recording automatically, without the presence of the investigator (Shatz et al., 1983; Wellman & Bartsch, 1994; Baron-Cohen, 1995). The assumption here is that since the child is absolutely unconstrained, the procedure is sensitive in that it can capture early locutions that children might be shy about using in the laboratory, thus providing a window into the child's early understanding of the topic at hand.

3.3 The data so far

As far as I know, the first two studies to use a naturally-occurring-speech paradigm to address issues in ToM were carried out by Bretherton, McNew, and Beeghly-

Smith (1981) and Bretherton and Beeghly (1982).

Bretherton and Beeghly (1982) explored the hypothesis that the capacity to explicitly impute internal states to self and other emerges late in the 2nd year, at the onset of communicative intentions. Their methodology consisted of i) asking mothers to listen to their children's speech for a week and write down phrases containing any one of 73 internal-state words and ii) a follow-up interview in which mothers were also asked to provide the context in which the target word was used, and, particularly important for our purposes here, they were asked to report on whether the child was also using the target word to refer to self or other. Their results showed that by 28 months, 93% of children have already acquired the word 'want' and that 77% used this target word for both self and others. These results stand in sharp contrast to the word 'think', for example, used only 33% of the 28 month-old children in the sample. Also noteworthy is that just 17% of that 33% used the word 'think' to refer to both self and others.

By the same token, these results are not necessarily due to the syntactic complexity of the kind of complement clauses that the target words take, for although the word 'need' can take the same complement clauses as the word 'want' ("I need an apple", "I need it red" and "I need to be six feet tall"), it was used by only 67% of 28-month-olds, 26% less than 'want'; in the same vein, the word 'know' (which takes the same complement clauses as the word 'think') was used by 66% of 28-month-old children, 33% more than 'think'. This means that, at least in general

terms, there are words in the child's vocabulary that are acquired later, probably due to conceptual difficulty and or pragmatic issues rather than syntactic difficulty. Similarly, Shatz et al. (1983) have also paid close attention to the kinds of complementation internal-state words take. They analyzed transcripts of one child (Abe) from ages 2;4 to 4;0 (years; months). They found that early in this period, mental verbs of belief and knowledge were more likely to be "complex", which they defined as containing more than one verb phrase rather than "simple", containing just one verb phrase. They also found that Abe's earliest utterances (early in that they occurred before the appearance of mental state terms of belief and knowledge) already contained instances of different kinds of complementation: they found 22 instances of utterances with wh-complementation ("Show me how to work this"), 15 instances of utterances with that-complementation ("I guess (that) I'll have a bobber") and 32 with to-complementation ("I tried to eat it all"). All examples come from Shatz et al. (1983). At the very least, these results show us that it is not necessarily grammatical complexity that is at stake here. If this were the case, then both 'need' and 'want' and 'think' and 'know' would have been used by roughly the same percentage of children at a very early age (see Wellman & Estes, 1987).

Of course, one could argue that these results may only be an artifact of word frequency; that is, words that are more frequent should be/are used first by most children (at any age). That does not seem to be the case in the Bretherton

and Beeghly (1982) study. Consider again the words ‘want’, ‘need’, ‘think’ and ‘know’ and the percentage of 28-month-olds that use them. Consider also that the frequency list of the British National Corpus (BNC) reports an occurrence of 94,293 for the word ‘want’, 62,201 for the verb ‘need’, 153,881 for the word ‘think’ and 185,534 for the word ‘know’, all values over a one-hundred-million-word (100,000,000) corpus. Given then that i) ‘want’ has a higher frequency of occurrence than ‘need’ and that ii) ‘know’ has a higher frequency of occurrence than ‘think’, it is surprising that while words that make reference to beliefs and knowledge are far more frequent than words that refer to desires, it is the latter kind of words that get used by the highest proportion of 28 month old children. If the results reported above were an artifact of word frequency, then one would expect that ‘think’ and ‘know’ would be far more frequent than ‘want’ and ‘need’. In analyses carried out by Bretherton and Beeghly (1982), results indicate this is not what is going on: ‘want’ is used by more children than ‘know’ or ‘think’, and ‘need’ is used more than ‘think’, counter to what would be expected from frequency counts. It is important to emphasize that the word frequencies I have reported here correspond to frequencies of both a written English¹ corpus and a spoken and written English² corpus. It seems rather obvious then that the method used in Bretherton and Beeghly (1982) is tapping much more than simply the acquisition of certain lexical items in terms of frequency effects or even the child’s

¹See <ftp://ftp.itri.bton.ac.uk/bnc/lemma.al>, for the BNC lemmatized frequency list.

²See http://www.comp.lancs.ac.uk/ucrel/bncfreq/lists/2.1_spokenvwritten.alpha.txt.

syntactic competence.

However, a real problem with Bretherton and Beeghly's (1982) analyses is that to study children's capacity to explicitly impute mental states by means of mental (state) words, it is not enough to merely count the occurrence of words like 'know', 'think' or 'want' in children's language or, similarly, to report how many children actually make use of those words at a particular age in a particular sample. Take for instance the fossilized expression "you know" in "I went to the movies yesterday, you know, and guess who I ran into?". The word 'know' in this example is not even referring to an internal state (let alone a mental state) but is just serving a communicative purpose of some other kind (a pause filler in this example). Although this is a clear-cut example of an internal-state word not being used to impute an internal state, one can think of several less clear examples in which mothers in the Bretherton and Beeghly (1982) study may have confused the mere uttering of a particular word for a word actually being used to impute an internal state. To be fair, Bretherton and Beeghly (1982) did ask for some contextual information in order to disambiguate the target words. The problem is that they did this *post facto*; that is, they asked mothers about their children's utterances a week after they had written down the utterances. This may have biased the data in several ways. Retrospective categorization has the potential for either over-reporting or under-reporting of usage in appropriate contexts. In either case, it would have added to the "noise" in the data.

A solution to this problem is to create a large database of fully-transcribed samples of children's spontaneous speech production and a set of functional categories that could help distinguish the various uses of internal-state words. The Shatz et al. (1983) study described above analyzed 30,160 utterances of one child's (Abe) spontaneous speech, grouped by age in 5-month intervals. Utterances containing one or more mental (state) terms were categorized into one of seven categories: i) mental state; ii) modulation of assertion; iii) directing the interaction; iv) clarification; v) expression of desire; vi) action-memory and vii) "I don't know" expressions. Categories (ii-iv) and (vi-vii) seem to have been proposed in order to avoid the problem mentioned above; namely, confusing conversational uses of mental state terms with the real imputing of mental states by means of these words. The authors found that although mental state terms of belief and knowledge were already present by 2;4, these words were used most often for conversational function rather than the imputing of a mental state of belief or knowledge. Interestingly, frequency³ and variety of functions seemed to have increased over time, with mental state terms used to refer explicitly and unambiguously to mental states appearing more or less at 28 months of age (see also Wellman & Estes, 1987). Unfortunately, for our purposes there are two problems with Shatz et al. (1983)'s study. First, they carried out a case study of only one child. While case studies are useful for evaluating the feasibility of hypotheses, they are fraught with limitations on generalizability. Second, they

³Whenever frequency is discussed, it is always relative frequency unless otherwise specified.

only reported/analyzed the use of mental state terms of belief and knowledge and not of desire. As noted above, they do have an “expression-of-desire” category (see (v) above), but this category applies only to mental state terms of belief and knowledge that are used to impute or refer to desire (“I hope we have popcorn”). Shatz *et al.* make it clear that they did not use “want” as a mental state term here: 1) they did not search the corpus for the term ‘want’ and 2) they took “mental terms” such as ‘hope’ or ‘wish’ as being paraphrasable by ‘want’ statements and thus coded separately.

The most comprehensive study of mental (state) terms using the naturally-occurring-speech paradigm is reported in Wellman and Bartsch (1994) and Bartsch and Wellman (1995). Especially important in the context of the present investigation is that these two analyses address the issue of teasing apart those utterances in which it is obviously the case that children have grasped the concept of desire from those in which it is not as obvious, or from those showing that they have not done so.

As noted above, previous analyses had either treated the issue rather superficially or completely neglected it. Wellman and Bartsch (1994) and Bartsch and Wellman (1995) analyzed a corpus of 200,000 child utterances from 10 children (whose ages ranged from 18 to around 72 months), searching mostly for internal-state terms of thoughts and beliefs (‘think’, ‘know’, ‘wonder’, ‘believe’, etc.) and internal-state terms of desires (‘want’, ‘wish’, ‘hope’, etc.). They found approx-

imately 12,000 containing these terms - mostly variations of 'think', 'know' and 'want'. Of course, as in Shatz et al. (1983), Bartsch and Wellman (B&W will be used hereafter to refer to Wellman & Bartsch (1994) and Bartsch & Wellman (1995) collectively)⁴ were not merely interested in the use of those terms, but in what children actually meant by using them. They basically reformulated the categories used by Shatz et al. (1983), differentiating mostly among the utterances that were genuinely imputing an internal state, the ones serving communicative purposes, and those which could not be clearly coded.

Interestingly, Bartsch and Wellman also used what they called a "window" of context; namely, while doing the computational analyses of the child speech database, they instructed the software they used to return not only utterances containing internal-state terms but also their immediate "context". The "context" consisted in three utterances before and three utterances after the utterance containing the target internal-state term. This practice allowed Bartsch and Wellman to disambiguate certain unclear occurrences of these terms.

Roughly in line with the results reported in Bretherton and Beeghly (1982), Bartsch and Wellman found that between the ages of 18 and 24 months, children already used internal-state terms of desire to refer to states of desire rather than using them for merely communicative purposes. In contrast, at this age range children did not seem to be using nearly as many internal-state terms to refer

⁴I will however use "Bartsch and Wellman" from time to time as well, which will also mean Wellman & Bartsch (1994) and Bartsch & Wellman (1995) collectively and it will simply be a stylistic variation to avoid repetition.

to beliefs and thoughts. This is, in fact, a striking finding. This result raises interesting and mostly unanswered questions in theory of mind research. Some of these questions include the following: why do children express desires before beliefs? What are the characteristics of desire talk (and, more generally, the concept of [desire]_p) that makes it easier to grasp by younger children? These and related questions about the representational status of desires will be part of the focus of this dissertation. Based on purely qualitative analysis of their data (i.e. without using statistical tools to test for significant differences), Bartsch and Wellman reported another result that merits some attention, specifically, the seeming decline in frequency of reference/expression to the mental state of desire at around the 37th month. Intriguingly, this decline seems to happen at the same time as an increase in expression of/reference to beliefs and thoughts. By 61-72 months, expressions of/reference to beliefs and thoughts on the one hand and about desires on the other seem to stabilize at roughly the same percentage of uses.

Although figure 4.3 on page 74 of Bartsch and Wellman (1995) portrays group data (that is, the data collapsed across all children), the same general developmental pattern is evident in individual children; at least, that is, for children for whom the authors had enough developmental data. Basically, then, the developmental picture of the concept of desire seems to take the following form: 1) use of internal-state terms of desire to refer to an internal state is already present by

around the 18th month of life; 2) a dramatic increase in reference to the state of desire occurs around the 36 months; 3) which use of desire terms falls back and stabilize at around 44 months; probably when children start to use beliefs terms (Bartsch & Wellman, 1995). These three findings motivate the present investigation: how reliable is this result? Can these results be observed in other children with roughly the same characteristics? If we replicate the study and a similar pattern of development is observed, what are the implications for theory or mind research?

An important reliability check of the naturally occurring speech paradigm is that it seems to be reliable in teasing apart differences among populations with different, relevant characteristics, such as autistic and Down's Syndrome children. In other words, it could have been possible that no difference between or among groups (age, sex, clinical condition, etc.) would be observed because the results are, to a certain extent, an artifact of the coding. In yet other words, the NOS paradigm would be useless if it could not give an account of certain known facts about these clinical populations. However, there is some evidence that this is not the case. This allows me to introduce the results of another set of fairly recent studies on a population of autistic and Down's Syndrome children.

Tager-Flusberg et al. (1990) recruited six boys diagnosed with autism (Rutter, 1978) and four boys and two girls diagnosed with Down syndrome. The two groups were well-matched in age and MLU (the number of morphemes divided by

the number of utterances in the samples). The procedure consisted in collecting natural speech protocols for a period of around three years and transcribing these protocols into computer readable files, to be later analyzed using computational tools.

Autistic children are impaired in their ability to impute mental states to themselves and others (Baron-Cohen, Leslie, & Frith, 1985; Frith, 1989; Wellman et al., 2001). If so, then this fact should be reflected in the way autistic children use (or do not use) certain words that stand proxy for certain concepts that refer to mental states of certain kinds (belief and knowledge, for one, and maybe desire if they have the same representational status as the former two). In other words, if the NOS paradigm works, then it should capture some difference between groups; namely, one would expect to see differences between the development of autistic children's talk about certain internal states (belief for sure, and maybe desire) and that of non-autistic children or, indeed, children with other non-theory-of-mind-related conditions such as, for example, Down syndrome (DS). Figure 3.1 below shows the results of Tager-Flusberg (1992) as reported in Tager-Flusberg (1993). For desires, children with autism show the same pattern of development as Down's syndrome children in the first few stages of their MLU development (Brown, 1973). However, by the time DS children start making reference to beliefs and knowledge and decrease their references to desires (as non-autistic children would, see Bartsch & Wellman, (1995)), children with autism are doing

exactly the opposite: they do not start talking about beliefs and desires and their references to desires increase considerably, at least by visual inspection of the data.

Figure 3.1: Genuine psychological references to desire by clinical condition (Autistic Vs. Down's Syndrome) as frequency of occurrence every 1000 utterances, From Tager-Flusberg (1992)

Clearly, this paradigm differentiates between the two clinical groups; namely, autistic children and Down syndrome children. This difference has been replicated using other paradigms (Tan & Harris, 1991). Several other observations from this study also suggest the appropriateness of the NOS paradigm with respect to teasing apart certain populations.

The first observation has to do with syntax and the acquisition of mental verbs. Just to beat a dead horse: autistic children were matched within MLU

stages as well as age groups. A comparison of the DS and autistic groups yielded only one statistically significant difference: IQ, which favored autistic children over DS children. In spite of this fact and in spite of the Tager-Flusberg et al. (1990) finding “that autistic children acquire specific grammatical structures in the same general order as has been found in non-autistic and DS children”, autistic children seem to have a different pattern of development (that the NOS paradigm is capturing) and that is not due to either syntactic complexity or general intelligence, at least of the kind that IQ (using the Leiter International Scale) measures. The second observation has to do with the source of the data: Tager-Flusberg (1992) and Tager-Flusberg (1993) obtained roughly the same results as Wellman and Bartsch – at least when it comes to the gross developmental picture. Although there is more to be said about these results, I’ll delay their treatment until Chapter 6.

However, Tager-Flusberg’s and Bartsch and Wellman’s analyses are not perfectly parallel. For one, both Wellman and Bartsch (1994) and Bartsch and Wellman (1995) used chronological AGE as the grouping factor, while Tager-Flusberg (1992, 1993) and Tager-Flusberg et al. (1990) used the five MLU stages of linguistic (mostly syntactic) competence, proposed by Brown (1973). This leaves some room for questions. What would happen if the groupings in Wellman and Bartsch (1994) were changed and MLU was used as the independent variable? Secondly, it is unfortunate that Tager-Flusberg’s analyses didn’t have a normal

group for control. However, given the considerations below with regard to the methodological characteristics of the NOS paradigm, it is possible to provide this control group *post hoc*, using data from other analyses.

3.4 The philosophy of statistical tests in the NOS paradigm

The objective of this section is to explain the rationale behind the statistical methods used in the analyses that follow.

Historically, the NOS paradigm has made only limited use of inferential statistics as an analytic tool. The primary reason is because the NOS paradigm relies on categorizing large amounts of naturalistic data that are difficult to come by and thus there is, many a times, a shortage of data, thereby limiting statistical analyses. Naturalistic data, as discussed above, has the advantage of being a relatively unconstrained record of child language compared to the kind of data obtained in controlled experiments. However, this feature of naturalistic data is also its Achilles heel, so to speak. Naturalistic data are by definition heterogeneous, meaning that samples can be of varying size and highly dependent on conversational topics. There is, of course, the added problem that the data should be longitudinal. Any samples collected over time from the same individuals are subject to the problem of missing data, especially in the case of data from children. In the type of analyses carried out in the present investigation, a major problem resulting from missing samples is that categories in which the data will

be coded will yield many missing data points.

Of the studies that have used the NOS paradigm in the past, only three have used any kind of inferential statistical method to test the hypothesis or to explore the results. Bretherton and Beeghly (1982) used a repeated measures ANOVA to assess the difference among percentages of the use of perceptual, physiological, volitional, cognitive and moral judgement/obligation terms and among uses of internal state terms for self and other. They used Neuman-Keuls post-hoc tests to assess particular contrasts between different conditions in their study.

Although a repeated measures ANOVA is the right kind of inferential statistical procedure for use with data that have these characteristics, its use is not always warranted. Given the small sample of children in Bretherton and Beeghly (1982), there are no missing data in any of the categories analyzed. However, more typically when working with longitudinal data at difference ages the problem of missing data from successive samples is something that quickly becomes evident. The repeated measures ANOVA procedure does not handle missing data gracefully. Because the underlying computations associated with analyzing longitudinal data rely on calculating difference scores between conditions, any missing values in a data set make the calculation of a difference score impossible. The overall effect of missing data from successive samples is to effectively remove the data from that child, even if only one sample is missing. Thus, if one starts with data from 10 children across five successive sampling periods and data from only

one sample (i.e., any sample) is missing from three of the children, data from those three children can no longer be incorporated in the analysis, resulting in the sample size being reduced to seven from ten children. In short, it is not always possible to carry out repeated measures on the data.

In order to still provide statistical analyses of the NOS paradigm data, other authors (e.g., Hickling & Wellman, 2001; Tager-Flusberg, 1993) have resorted to using an independent groups ANOVA procedure to test the hypotheses. In Tager-Flusberg's analyses, a Multivariate Analysis of Variance (MANOVA) was used to test for group differences in the frequencies of lexical terms used in a set of categories on samples coming from autistic and Down's syndrome children. Carrying out an independent groups ANOVA on the sample has the advantage that missing points in the data can be dealt with successfully. In order to test each category independently, several MANOVAs were carried out for *each* category including other variables. Univariate analyses were used to test for some of the inter-group differences.

Being a more restrictive calculation, the results obtained by this method may be in fact prone to Type II errors (i.e., deciding that no difference exists between conditions when, in fact, a difference does exist in the population). This is indeed a problem and its presence needs to be kept in mind when working with this statistical method. On the other hand, if statistically significant results are indeed found, then it could be assumed that those results will be replicated,

perhaps with somewhat more power, using a repeated measures ANOVA in a sample with no missing data. In other words, the only results that should be interpreted with some degree of suspicion are those approaching significance, rather than those which are in fact statistically significant or those which are far from reaching significance.

Hickling and Wellman (2001) are the only ones who have tried to give a rationale behind their use of statistics while working in the NOS paradigm. They have argued, following Bakeman and Gottman (1997) that although the NOS paradigm utilizes longitudinal data coming from the same sample of children and are thus repeated measures of the same participants, "these data may be considered independent in the sense required for valid inferences of statistical significance". In order for successive events to be regarded as independent for the purposes of carrying out inferential statistics, they have to comply with the following two conditions: the first is that there are several (at least two) different raters (presumably independent) that code the data separately; the second is that the coding system should consist of several different, mutually exclusive categories (Hickling & Wellman, 2001; Bakeman & Gottman, 1997; Bartsch & Wellman, 1995; Wellman, Harris, Banarjee, & Sinclair, 1995). In their study, Hickling and Wellman (2001) used z tests, which, they argue, are appropriate for pooled data in the sense they have discussed.

Given that there is no consensus as to the nature of the statistical methods

that should be used in this paradigm to test for significant results, the investigations reported here will adopt the following assumptions: a) since the correct statistical method to apply is without doubt a repeated measures ANOVA, this is the method to be applied whenever there are enough data to support it. Post-hoc and pairwise comparisons will be carried out to test for significant differences in the different variables. These post-hoc tests will all be adjusted to the Bonferroni criterion.

In accordance with the discussion above, when there are missing data, as in the Tager-Flusberg (1993) analyses, an independent groups ANOVA procedure will be applied on the data. This rationale seems to be warranted on two grounds: a) analyses that have been published have used this method and b) that other researchers Bakeman and Gottman (1997) have argued that this method of evaluating the data to be independent even if coming from the same individuals. The data reported here also adhere to the two conditions proposed by Bakeman and Gottman (1997): a) they have been rated by two independent coders, one of them with no previous experience with the topic at hand and no theoretical preconceptions, and b) all the categories in which the utterances are coded are mutually exclusive. Here also, pairwise comparisons and post-hoc tests will be carried out using a Bonferroni criterion for adjustment.

One more note regarding the use of independent groups ANOVA in the NOS paradigm. Since, as discussed above, using independent groups ANOVAs in-

creases the chance of making Type II errors, some results will be interpreted even if they are only approaching significance. No special emphasis will be put on those effects that are close to reaching significance, in the fashion of Tager-Flusberg (1993), but they will be mentioned in view of future availability of more data. This “approaching significance” approach will at least suggest a road to take for future investigations when there are enough data to test the issues further.

These provisos notwithstanding, I believe that the assumptions for choosing and carrying out the analyses mentioned above are strong enough to give us some insight about what may be happening developmentally for the samples of children when it comes to their use of the term of desire.

3.5 Limitations of the paradigm

The NOS paradigm is not without constraints. One of the most important problems is representativeness. Naturalistic language samples usually sample language from only a few children. This has serious consequences for the generalization of the results to other children or even across children within the sample. In order to deal with these constraints, researchers have leaned towards creating large databases of child language, so one child may have many thousands of utterances associated to different (systematic) stages in her development. Fortunately, several researchers have made their data available to other researchers,

thus yielding samples of language from several hundred children. The analysis reported here incorporate data from 12 children, 6 boys and 6 girls, at different stages in their development. This sample size is similar to other analyses in the literature (Bartsch & Wellman, 1995; Tager-Flusberg, 1993).

Two other related problems have to do with amounts of raw data available. First, given that each child's file may contain many thousands of utterances, there are problems with searching and individualizing all possible interactions that bear on the topic at hand. That is why researchers have to choose certain words that would stand proxy for the different concepts they are investigating. It would be prohibitively time-consuming to go over every single child utterance, one by one, and code them in a certain manner. Thus, some conversations about desire may be overlooked. Second, even if the assumption that in relative terms children will use certain words in the same way that they become aware of the concept these words stand for, it is extremely hard to gauge the accuracy of the codings. There may be the case of categories cross-contaminating other categories. This issue is taken up again in more detail in Chapter 5.

The problems just listed are serious, but not more serious than the the problems of other approaches. The NOS paradigm has been devised to complement experimental techniques, not to replace them. This paradigm would be of no use, however, if the results it yielded were not consistent with various known phenomena. In other words, it could be the case that the NOS paradigm is insensitive

to well-established facts in the experimental literature. However, this does not seem to be the case. The next section presents the results of the studies that have used the NOS methodology. As will become clear, these analyses suggests that the NOS paradigm yields results that are both internally consistent (separating populations of participants in certain conditions) and externally consistent as well (by confirming some of the results obtained in experimental studies).

Another big limitation draws on an analogy to linguistics. In one of his seminal works, Noam Chomsky drew a distinction between competence and performance (Chomsky, 1965). For him, competence was the underlying linguistic knowledge that an individual has (about the syntax of language in his case), while performance was putting that knowledge to work (actually producing sentences that reflect knowledge about particular syntactic structures), somewhat akin to our distinctions between folk psychology and folk psychological practice.

Since this is the case, and since the analyses presented in these pages rest heavily on purely linguistic evidence, it may be the case that performance (actually saying the word 'want' or not) may both under- and over-represent the child's knowledge about the concept in question, the concept of $[\text{desire}]_p$. For example, it may very well be the case that the child does not have the linguistic resources to meaningfully say "I want a cookie" and says merely "Cookie!" instead. These expressions with no word 'want' or its other instantiations "wanna, wants, wannu, etc." would be missed outright by the tools used in this disser-

tation. However, these utterances may sometimes stand for real expressions of desires, thus making a child appear that he or she has less folk psychological knowledge than he or she really does. The NOS paradigm cannot currently give an account of this phenomenon because it relies on computational procedures in which software is instructed to locate explicit markers for desire; namely, words normally associated to the expression of desire.

By the same token, there are times when the opposite is true; namely, there are times when a child appears to have the knowledge, producing utterances such as “I want a cookie”, where, contrary to appearances, it may be no more than a parroting of a particular pattern of sounds, not an act of referring to anything at all, least of all to a complex concept such as desire or even self. Fortunately, since these kinds of expressions are picked up by the tools used to identify want utterances, the NOS paradigm can and does give an account of them. It does so by being very conservative with the kinds of expressions that are taken to be expressions or references to desire. As will be shown, there are categories (not without problems themselves, see Chapter 5) that help to tease these utterances apart and categorize them correctly (see section 5.1.1).

Regarding the first kind of utterances, it should be said that they remain a problem for the NOS paradigm. To be sure, no experimental paradigm is without problems and it is all these paradigms taken together that should paint a close picture of reality, more than any one taken individually. Thus, the results of

the NOS paradigm should be backed up by more controlled experiments, laboratory experiments, so to speak. For now, I work under the assumption that the results reported here will paint *a* picture of the development of the concept of desire that should be replicated and extended and researched by means of other methodologies.

CHAPTER 4

The corpus

4.1 The Data I: The corpus from non-autistic children

The corpus from non-autistic children was organized in two ways. First, the corpus was organized taking age as the grouping factor, following Bartsch and Wellman's original age stages. Second, the corpus was also organized into MLU stages, using those stages proposed by Brown (1973).

4.1.1 The corpus by age

A total of 14,896 child utterances were taken from the Wells corpus (Wells, 1981) in the CHILDES database (MacWhinney, 2000). These data come from the longitudinal observation of spontaneous speech production of 12 children (6 boys and 6 girls) whose ages ranged from 18 months at the time of the first observation to 60 months at the time of the last observation and who were acquiring English as their mother tongue. Each child was observed a total of 10 times, for about 40 minutes each, at 3-month intervals. Since the objective of Well's (1981) research

was to obtain spontaneous speech samples, a timing mechanism was devised to set off a tape recorder –hooked up to a wireless microphone in the child’s garment– at different times between 9am and 6pm, to prevent parents from planning activities, for example. Twenty-four 90-second samples were recorded in each observation. These were later transcribed into several files using normal English orthography.

Table 4.1 shows some general information about the samples. First, it is important to notice that the corpus was carefully chosen to avoid missing data at the different ages for the different children, a problem in the Bartsch and Wellman analyses. The data used here seem to be as homogeneous as possible given the characteristics of the paradigm being used. A second important issue that is worth mentioning here is that the MLU values given in table 4.1 were calculated by means of the <mlu> command in the CLAN program (v26Mar2002, see MacWhinney, 2000), which counts the number of words in each child utterance in file *Y* and then divides them over the total number of child utterances in that same file *Y*. Thus, the MLU values given below correspond to the average of all samples for all children at that age range. Since sometimes there is more than one sample associated with a particular child at a particular age, then the value used to compute the general MLU for that age range was the mean MLU of those files. For example, Ellen has three different files at ages 18-24, say *W*, *X* and *Z*; each, not surprisingly, with a different MLU value. To allow for general statistics of the kind given in Table 4.1, the three MLU values corresponding to files *W*, *X*

and Z were averaged and it was this resulting value that was entered to calculate the average of the group.

Ages (mo.)	N	#Samples	#TotUtt	MLU(x')	MLU(SD)
18-24	12	36	3483	1.480	0.252
25-28	12	13	1830	1.697	0.430
29-32	12	17	2694	2.246	0.566
33-36	12	18	2972	2.709	0.446
37-40	12	12	2041	2.981	0.390
41-44	12	12	1876	3.202	0.425

Table 4.1: Information on the samples.

In Table 4.1, Ages (mo.) = ages in months, N = number of participants in each age group (all twelve participants are the same children at different ages), #Samples = the number of files containing transcriptions for all children at that particular age, #TotUtt = the total number of utterances in the samples, MLU(X') = average mean length of utterance for that age group, and MLU(SD) = the standard deviation of the mean for the MLU values for that age group.

Table 4.2 shows the total number of utterances available from each child by age in months. Given that this chapter will be, in part, a replication of those analyses by Bartsch and Wellman, the age groups correspond to their age groups, with the difference that they also analyzed transcripts from much older children. The omission of data from older children in the present study is appropriate for several reasons. First of all, Bartsch and Wellman were concerned not only with the acquisition of mental state terms of desire, but also with mental state

Child(*)	Age Groups (in months)						Total
	18-24	25-28	29-32	33-36	37-40	41-44	
Abigail(f)	216	146	127	184	139	121	933
Benjamin(m)	427	131	169	282	170	80	1259
Darren(m)	204	88	319	155	164	164	1094
Debbie(f)	395	264	176	136	134	125	1230
Ellen(f)	497	177	504	251	239	282	1950
Elspeth(f)	208	117	294	91	110	98	918
Frances(f)	306	68	203	175	201	170	1123
Gary(m)	341	159	189	425	124	270	1508
Geoffrey(m)	133	138	170	196	227	133	997
Harriet(f)	290	174	120	307	156	128	1175
Jack(m)	334	273	296	506	253	149	1811
Jason(m)	132	95	127	264	124	156	898
Sum	3483	1830	2694	2972	2041	1876	14896

Table 4.2: Total utterances available from each child by age in months.

terms of belief and thought, the latter allegedly appearing at around forty-eight months. In the present work, however, I am only concerned with the acquisition of one internal state term of desire; namely, ‘want’, already present, allegedly, at around the twenty-fourth month of age. Second, if the developmental trend for the concept of desire found by Bartsch and Wellman is a solid result, then these six age groups should be enough to observe the trend. Also, although my sample has fewer utterances than B&W, this should not affect the overall result, given that most of the time I will be talking about percentages, rather than absolute frequency values.

4.1.2 The corpus by Mean Length of Utterance (MLU)

MLU is a widely used index of grammatical ability that is calculated by taking all utterances in a sample of child language, counting either the number of words or the number of morphemes in each utterance and then dividing the resulting number by the total number of utterances in that sample. Let me explain this calculation by using an example: Suppose we have the following sample corpus of child language:

*JON: I've been a good boy.
*MOT: Have you?
*JON: Yes. *JON: Well give us some sweets.
*MOT: Pardon?
*JON: Please #3.
*MOT: I feel tired #2.
*JON: Have some sweets Mummy [=! vocative] please?

This sample corpus has 5 utterances belonging to the child Jon. These utterances contain the following number of words: utterance1 = 5 words (whether you count "I've" as one or two words is not that important if you are systematic; my justification, however, for counting "I've" as just one word is that these usually function as single morphemes for children, like 'gouna' and 'wanna', see Brown, 1973); utterance2 = 1 word; utterance3 = 5 words; utterance4 = 1 word; and lastly, utterance5 = 5 words. Thus the total number of child words for this corpus is 17. When this total is divided by 5 utterances, the resulting MLU (mean

length of utterance (in words)) is 3.4. However, there is another way to calculate MLU that has been favored in the linguistic literature: instead of counting the number of *words* in a sample, it is morphemes that are counted. Therefore, for instance, the sample corpus would be “parsed” in the following way: utterance1 = 5 morphemes (my rationale for counting “I’ve” as just one morpheme is the same as above); utterance2 = 1 morpheme (Brown’s, 1973, rules for calculating MLU allow one to count ‘yes’, ‘yeah’, ‘no’ as morphemes; utterance3 = 6 morphemes (the word ‘sweets’ counts as two morphemes: ‘sweet’, the base morpheme and ‘-s’, one of the allomorphs of the plural morpheme in English); utterance4 = 1 morpheme, and utterance5 = 6 morphemes (‘sweets’ again), yielding a total of 19 morphemes that is divided by the 5 utterances, resulting in an MLU (in morphemes) of 3.8. I will refer to MLU calculated in terms of number of words per sentences as MLU_w and to MLU calculated in terms of number of *morphemes* as MLU_m .

Following Brown (1973) and later work in the NOS paradigm by Tager-Flusberg *et al.* (1990) and Tager-Flusberg (1992, 1993), I will divide the developmental period covered by the samples into Brown’s five equidistant “Stages”. There are, in total, five stages (named in Roman numbers, beginning at MLU I = 1.75 and ending at MLU V = 4.00, and with increments of 0.50 in most of the cases (see, however, Stages III to IV in the following description)). STAGE I begins at MLU=0 and continues up to and including MLU = 1.75 (so any sample

of an MLU between 0 and 1.75 (included) would belong to this stage), STAGE II comprises the range between $MLU = 1.76$ to $MLU = 2.25$, STAGE III ranges from $MLU = 2.26$ to $MLU = 2.75$, STAGE IV ranges from $MLU = 2.76$ to $MLU = 3.5$, and lastly, STAGE V ranges from $MLU = 4.00$ onwards. This division is preliminary and only for the purposes of presenting the data in a fashion consistent with what is the general practice in the field.

MLU Groups						
Child	MLU I	MLU II	MLU III	MLU IV	MLU V	Total
Abigail(f)	121	95	397	139	351	1103
Benjamin(m)	318	–	579	362	143	1402
Darren(m)	457	–	154	896	–	1507
Debbie(f)	656	–	215	259	179	1309
Ellen(f)	381	534	514	905	–	2334
Elsbeth(f)	325	–	–	521	316	1162
Frances(f)	374	98	308	546	–	1326
Gary(m)	341	–	348	1084	–	1773
Geoffrey(m)	271	–	536	360	85	1252
Harriet(f)	319	449	–	407	150	1325
Jack(m)	816	294	402	299	194	2005
Jason(m)	618	–	280	306	–	1204
Sum	4997	1470	3733	6084	1418	17702

Table 4.3: Total number of utterances available from each autistic child at each MLU stage.

The same database of utterances from Wells (1981) was then rearranged according to the MLU_w values of the samples they occurred in. MLU_w was chosen rather than MLU_m because the counting of utterances is less prone to mistakes compared to morpheme categorization. Besides, MLU_w is calculated semi-

automatically by the CLAN program. CLAN stands for Computerized Language Analysis and it is a tool designed specifically to analyze transcriptions in the format of the Child Language Data Exchange System (CHILDES) database and it can be applied regardless of the researcher and language of the transcripts (MacWhinney, 2000). Table 4.3 shows a summary of the number of available utterances by MLU_w stage and by child. The discrepancy with respect to the total number of utterances (14896 versus 17702) is due to the fact that some of the children's transcripts belonged to an age group not analyzed in this investigation (older than 44 months of age). Thus, some children's transcripts belonged to a possible age group of more than 44 months of age, but the same transcript belonged to an MLU of say MLU IV, so these utterances had to be considered in the MLU count. There are, in fact, 2806 more utterances in the MLU analyses than in the AGE analyses.

Unfortunately, unlike the previous arrangement of utterances by age, here there are some empty cells of data. This will be dealt with by working with proportions and indeed rearranging the data when necessary. However, I would like the reader to keep this in mind during the discussions of the analyses below.

4.2 The Data II: The corpus from autistic children

The corpus of utterances coming from children with autism is special in at least one respect: there is not the possibility of grouping the children producing those

utterances by age. Usually, autistic children have great difficulty in speaking fluently (Frith, 1989; Baron-Cohen, 1995), although their difficulty lies more in the pragmatics of language than its phonology, morphology or syntax (Tager-Flusberg et al., 1990; Tager-Flusberg, 1993, 1992, 1988). In the Tager-Flusberg analyses (which are henceforth taken to be the ones in the previous reference, taken collectively), children differed in chronological ages but were matched on MLU and other linguistic scales like the Leiter scale (Leiter, 1974) and IQ/Mental Age. Thus, the corpus coming from autistic children will be grouped only according to MLU stages, rather than age.

4.2.1 The corpus by Mean Length of Utterance (MLU)

A total of 63 samples totaling 19,413 child utterances from the Tager-Flusberg corpus (Tager-Flusberg et al., 1990) in the CHILDES database (MacWhinney, 2000) were analyzed for mental state language. These data come from the longitudinal observation of spontaneous speech production of 6 autistic children (6 boys) diagnosed by Rutter's criteria (Rutter, 1978) and in conformity with the DSM-III-R criteria (American Psychiatric Association, 1987). That the samples come from boys rather than from both boys and girls has to do with the fact that the incidence of autism in girls is very low compared to boys, usually around 2.5 boys to 1 girl (Lotter, 1966; Bryson et al., 1988; Frith, 1989).

Each child was observed for a total of 8 to 13 times and their speech recorded

for about 40 to 70 minutes each, at 2-month intervals. Two researchers did the visits. One of them was responsible for recording the child while the other took relevant notes about the child's activities and behavior (Tager-Flusberg, 1992). Unlike the study with non-autistic children, where children were recorded unexpectedly, mothers in the Tager-Flusberg et al. (1990) study were asked to prepare activities and select toys that would elicit the child's language.

MLU Stage	N	#Samples	#TotUtt	MLU($\bar{x}$)	MLU(SD)
MLU I	3	23	6616	1.441	0.151
MLU II	4	13	4160	2.011	0.126
MLU III	4	17	5025	2.461	0.147
MLU IV	3	10	3612	2.990	0.185

Table 4.4: Information on the samples from the autistic children.

In Table 4.4, N = number of participants in each MLU group (all six participants are the same children at different MLU stages)¹, #Samples = the number of files containing transcriptions for all children at that particular MLU stage, #TotUtt = the total number of utterances in the samples, MLU($\bar{X}$) = average mean length of utterance for that group, and MLU(SD) = the standard deviation of the mean for the MLU values for that group.

¹the Roman numeral names of MLU stages do not necessarily correspond to the MLU values themselves. They are completely arbitrary.

Child(*)	MLU Groups				Total
	MLU I	MLU II	MLU III	MLU IV	
Brett(m)	–	–	472	3069	3541
Jack(m)	–	2003	1229	–	3232
Mark(m)	4803	538	–	–	5341
Rick(m)	492	1229	1084	191	2996
Roger(m)	–	390	2240	352	2982
Stuart(m)	1321	–	–	–	1321
Sum	6616	4160	5025	3612	19413

Table 4.5: Total number of utterances available from each autistic child at each MLU stage.

4.3 Identifying ‘want’ utterances

The first step was to identify all and only the instances of the internal state term ‘want’ in all and only the target child’s utterances. To do this, the <kwal> command in the CLAN program (MacWhinney, 2000) was used. It is not enough to ask the program to identify only the lexical item ‘want’ in the child corpus. Due to the fact that inflected forms of this term (e.g. ‘wanting’, ‘wants’, etc.) or the original researcher’s transcription idiosyncrasies (e.g. ‘wanna’, ‘wanning’, etc.) would be overlooked, the program had to be instructed to look for all lexical forms that started with the root ‘wan-’. This practice posed a problem of its own; namely, the inclusion of false positives, like the lexeme ‘wander’, produced by one of the children in our sample (Elspeth, at 30 months). However, this side-effect was preferred to having false negatives and it was easily worked around by manually going through all instances of ‘want’ identified by the computer and again manually deleting all false positives. The exact command used to identify

instances of 'want' was <kwal +t*CHI +s 'wan* '>, where <kwal> instructs the computer to search for certain key word in a corpus and output that word embedded in the context utterance it occurred in. The <+t*CHI> segment instructs the computer to include only utterances belonging to the target child, instead of all utterances in the sample, like those of parents and/or siblings. The <+s 'wan* '> segment instructs the machine to look only for words that start with 'wan' irrespective of what it ends with (e.g. 'want', 'wanna' and 'wander' would be picked up by the program, while 'water' would not).

Given that the outputs of the CLAN program take the form of a plain structured ASCII file, the second step was to order the results in such a way as to make querying these results much faster. I developed a small application in Visual Basic to extract information from the files and then put the extracted information in a Microsoft Access database. There were five pieces of information to be extracted from the ASCII files: the word recognized (e.g. 'wanting'), the line in the original transcript where the instance of the word was found, the name of the file being analyzed and its path (e.g. "c:\temp\wells\ellen\ellen07.cha"), the word embedded in the context utterance (e.g. "I want to say good bye") and the name of the speaker who produced the utterance (e.g. "Anne"). Once all the information was appended to the different fields of the database table, this table was related by a many-to-one relationship to another table in the same database that contained all the developmental information taken from each sample in the

corpus.

For example, for the file called "c:\temp\wells\ellen\ellen07.cha", there was another table with a record that had the same file name to which it would be related containing the following information: the total number of utterances in that file, the general MLU value for that file, the standard deviation of the MLU value, the age of the child in CLAN format (e.g. "1;11.18" (years;months.days)) and finally, a calculation of the age in months only (e.g., for our example, 1;11.18=23 months). Building this relational database was necessary because querying a database with all the information associated to the children and the utterances in the corpus is much simpler than visually inspecting the CLAN output and, of course, much less prone to identification and counting errors.

Child(*)	Age Groups (in months)						Total
	18-24	25-28	29-32	33-36	37-40	41-44	
Abigail	3	3	6	5	3	3	23
Benjamin	15	4	1	24	7	8	59
Darren	1	0	14	12	2	9	38
Debbie	5	0	4	11	12	4	36
Ellen	11	1	21	21	19	10	84
Elspeth	1	3	22	5	6	5	44
Frances	1	1	13	6	12	6	39
Gary	6	5	7	25	10	17	74
Geoffrey	0	14	28	17	7	11	68
Harriet	13	6	4	40	1	11	77
Jack	1	0	1	34	5	7	48
Jason	0	0	1	4	2	5	12
Total	57	37	120	206	86	96	602

Table 4.6: Total 'want' utterances available from each of the non-autistic group by AGE stage.

Unlike B&W's analyses, I concentrated on the term 'want' rather than other possible terms that could also potentially hint at the child's grasp of the concept of desire. This was done because of two reasons: first, it is well known that 'want' is the most frequent desire word in the child's language. Other fixed phrases ("I'd like", "I wish I had") are much less common and should not in principle have a large effect in the relative frequency at different stages, see Bartsch and Wellman (1995). Second, the term 'want' is the canonical term by which children refer to desire. Other phrases might actually involve other linguistic constructions that are more difficult either syntactically (conditionals, etc.) and/or semantically for the child at the earlier ages. Adding more potential desire terms should not materially alter the overall developmental pattern.

The final step was actually counting all the occurrences of the word 'want' and assigning frequency counts to each child at a particular age. To do this, the database was queried using standard SQL syntax, where, basically, the computer is asked to retrieve all and only the 'want' utterances that occur between, for example, the ages of 41 and 44 months and only for the child Ellen. This provided us with the number of records in the table that satisfied those conditions. The important information to be retrieved was the word, the name of the child and the age of the child. It was important to retrieve the word because there were cases in which the word 'want' (in any possible form, that is, gerund, past, present simple, etc.) appeared more than once in the same utterance and these had to

be added up in the final count.

Table 4.6 shows the total number of 'want' utterances for each individual child as a function of age. Given that part of the aim of this dissertation is to provide a replication to the analyses reported by Bartsch and Wellman, grouping non-autistic children as a function of age was mandatory. Most of their analyses in fact take age as an independent variable.

Given that the autistic population could not be matched to the non-autistic population by chronological age, the same analyses also had to be carried out as a function of MLU. Tables 4.7 and 4.8 show the raw number of 'want' utterances for each child in the non-autistic and autistic groups as a function of MLU, respectively.

Child	MLU Groups					Total
	MLU I	MLU II	MLU III	MLU IV	MLU V	
Abigail	1	2	14	3	4	24
Benjamin	2	0	25	30	0	57
Darren	2	0	11	29	0	42
Debbie	5	0	15	20	0	40
Ellen	8	7	40	35	0	90
Elsbeth	4	0	0	31	5	40
Frances	2	6	11	24	0	43
Gary	6	0	11	59	0	76
Geoffrey	14	0	49	17	0	80
Harriet	8	23	0	46	0	77
Jack	2	25	11	16	0	54
Jason	5	0	7	12	0	24
Total	59	63	194	322	9	647

Table 4.7: Total 'want' utterances available from each of the non-autistic group by MLU stage.

Child(*)	MLU Groups				Total
	MLU I	MLU II	MLU III	MLU IV	
Brett(m)	–	–	27	292	319
Jack(m)	–	60	24	–	84
Mark(m)	27	2	–	–	29
Rick(m)	38	11	69	17	135
Roger(m)	–	19	180	26	225
Stuart(m)	139	–	–	–	139
Sum	204	92	300	335	931

Table 4.8: Total ‘want’ utterances available from each of the autistic group by MLU stage.

Figure 4.1, in turn, shows a comparison between the non-autistic and autistic groups when it comes to the mere frequency of ‘want’ utterances. As we can see, non-autistic children display a small increase of the frequency of ‘want’ utterances from MLU I to MLU II. At MLU V, children seem to slightly decrease the frequency of ‘want’ utterances. Children with autism, on the other hand, do not show the same pattern. In fact, children with the syndrome start by producing more ‘want’ utterances than non-autistic children, but there is a small, unexpected, decrease in the production of ‘want’ utterances at MLU II compared to MLU I. This decreasing phenomenon is not very common in non-autistic children (Tager-Flusberg et al., 1990; Brown, 1973). After this, autistic children seemingly increase the frequency of ‘want’ utterances, making them seem more similar to non-autistic children. Unfortunately, there are no data for the autistic children at MLU V and thus there is no way of comparing whether the small decrease apparent in non-autistic children is there as well.

Figure 4.1: A comparison between the frequency of occurrence of the term 'want' in the non-autistic versus autistic populations by MLU (error bars are Standard Deviations)

CHAPTER 5

Desire in normal monolingual English children

5.1 Identifying genuine reference to [desire]_p

As we can see from chapter 4, children in the first age group (18-24 months) are already using the verb 'want' in their spontaneous speech. Given this, one might be inclined to claim that, just by using this term, children must have already acquired the concept this term stands for. In other words, one might be inclined to say that these children at these very early ages (in the context of ToM development) are actually able to express their own desires, refer to those of others' and even explain actions by means of this concept.

However, it is not the case that, *by default*, all instances of this term involve the concept of desires, which in turn makes reference to an organism's mental state of desire (Tager-Flusberg, 1992, 1993; Bartsch & Wellman, 1995). Given the functional differences in the use of the term 'want' by children, to study these children's engagement in folk psychological practice, it is not enough to merely count the occurrences of internal state terms of desire such as 'want',

'care (about)', etc. What is needed is a set of functional, mutually exclusive, categories that would help us distinguish the various uses of this term. Below, I will provide a possible list of these categories, first proposed by Wellman and Bartsch (1994).

5.1.0.1 Genuine psychological references to desire (GPRDs)

Genuine psychological references to desire (henceforth, GPRDs) are instances of children's unequivocally referring to themselves or other people as being in a mental state of desire, like in the following exchange from Bartsch and Wellman (1995):

CHILD: Daddy! Fill my squirt gun.

FATHER: Why?

CHILD: Because.

FATHER: Why?

CHILD: Because I want to play with it.

Some of the conditions that utterances should fulfill to be GPRDs are: a) that the objects mentioned in the desire complement clause (or direct object) should not be in plain view and/or they should be unattainable; b) that the same utterance should not have been uttered by the child in the previous and next 10 child utterances; c) that the utterance should not be a very frequent, formulaic expression of the language; d) that the utterance should not be a 'want' word

with nothing after it (unless the context makes it absolutely clear what the child was referring to).

5.1.0.2 Behavioral requests (BRs)

Behavioral requests, in turn, (henceforth, BRs) are ‘unadorned’ instances in which the child uses a desire term to fulfill an immediate goal, like receiving something that is beyond her reach but in plain view. Bartsch and Wellman (1995) take these instances to mean nothing more than ‘give me x ’. For example, in Bartsch and Wellman (1995):

[Father is rubbing some gel on child’s arm]

CHILD: I want that [some gel].

Some of the conditions that utterances should fulfill to be BRs are: a) that the object of the desire complement clause or direct object be in plain view and/or reachable by the child; b) that the same utterance should not have been uttered by the child in the previous and next 10 child utterances; c) that the utterance should not be a very frequent, formulaic expression of the language; d) that the utterance should not be a ‘want’ word with nothing after it.

5.1.0.3 Direct repetitions (DRs)

Direct repetitions (henceforth, DRs) are dialog turns in which the child merely repeats the adult (or his own) utterance, as in Bartsch and Wellman (1995):

CHILD: I want to go.

MOTHER: You want to go?

CHILD: I want to go.

Some of the conditions that utterances should fulfill to be DRs are: a) that the same utterance be uttered by the child in the previous or next 10 child utterances; b) that the utterance should not be a very frequent, formulaic expression of the language; c) that the utterance should not be a 'want' word with nothing after it.

5.1.0.4 Idiomatic expressions (IEs)

Idiomatic expressions (henceforth, IEs) are high-frequency collocation of words in the particular language. This is the case of the English expressions "I don't want to" or "I want more", when they appear without an object, as in the following example in Bartsch and Wellman (1995):

CHILD: I wanna.

Some of the conditions that utterances should fulfill to be IEs are that the utterance be a very frequent, formulaic expression of the language.

5.1.0.5 Uncodable utterances (UUs)

Uncodable utterances (henceforth, UU) are instances of the desire term ‘want’ for which categorization was impossible, due mainly to failure in retrieving contextual information from the dialog. Bartsch and Wellman do not give an example of this category.

5.1.1 Bartsch and Wellman’s categorization scheme: some examples, some problems, and a rationale for its continued use

The categories used in this dissertation are those used by the two most comprehensive studies in the literature, those of Bartsch and Wellman (1995) and Tager-Flusberg (1993). These categories, of course, have advantages and disadvantages.

The most important advantage of using this particular set of categories is that the results can be compared to previous work using these same categories. Having a completely new categorization scheme would have implied that the results obtained would have to wait to be validated by a replication or an extension further down the road. Since not many studies have used the NOS paradigm, and even fewer have used it to research children’s knowledge of folk psychological concepts such as [belief]_p and [desire]_p, it seems useful to validate the results obtained by previous studies rather than to refine the coding schemes.

But what do the categories really pick out and what are the limitations of this

categorization scheme? I have talked about the categories in general terms, but I do need to go into a bit more detail. Below is a discussion and some examples of each of the categories, together with what they pick out as instances of those categories and the problems these instances raise in the overall coding scheme. I will start with what is taken to be the most central category of these studies, genuine psychological references to desires.

The category "Genuine psychological references to desire" (GPRD) is probably the most committal of all the categories and the most difficult to identify. In this dissertation, several different characteristics pointed at a GPRD. The "rules" given in section 5.1 for identifying them are, of necessity, more of a guideline. There is no really infallible algorithm for identifying GPRDs. Below are a few examples of what was taken to be a GPRD in this dissertation and the issues they raise:

The most common instance of GPRDs were those in which the child uses simply the word 'want' and a physical object such as tea in the case of Benjamin (23 months) "I want my tea" or even Abigail's (27 months) "I want Mommy". In both these cases, neither Benjamin's tea nor Abigail's mommy were present in the scene. In fact, it was simply that Benjamin was hungry in the first instance (while he was just playing with some toys) and that it was only Abigail's father who was in the house at that particular time. It is supposed, then, that these desires are not merely formulaic language. In uttering those words, children are

meaningfully using a lexical term that refers to desires, thereby showing a grasp of the concept. These kinds of utterances will be taken to be references to desire in these pages, not necessarily a reference to the desire state itself, but a reference to $[\text{desire}]_p$. It is the word 'want' that is doing the referring and $[\text{desire}]_p$ is the referent.

Allegedly, these are the simplest forms of GPRDs, containing only an object. The problem one runs into when trying to code these utterances is that they are too close to being possible instances of behavioral requests, rather than GPRDs. However, most of the time, the context makes it clear that they are in fact GPRDs because it would not be possible to paraphrase them into "give me X", for example, and because the child does not repeat the same expression over and over in subsequent lines, making the line a piece of formulaic language that the child has acquired. For example, it could be the case that when saying "I want mommy", Abigail is only saying "I'm bored" or "I'm sad", but has strategically learned that "I want mommy" can get some attention to her feeling sad or bored and uses it without really wanting her mom. Fortunately, these formulaic utterances are easy to identify once you have read the whole transcript belonging to the child. In fact, the researcher learns to actually "know" the child enough to tease apart those formulaic utterances from the ones that are not (see the discussion below and also section 3.5).

Other instances of GPRDs are significantly clearer as to their status; such as

Benjamin (23 months) "I want to turn it around" or Gary's (27 months) "want to go greenhouse" or Gary's (27 months) "I want that one there". In these examples there is no doubt that the child is using the concept of desire as an adult would. Gary and Benjamin are both communicating their mental state of desire for a particular state of affairs in the world ("I want that one there") or for a given action that they are not performing at the time ("I want to go to the greenhouse"). As is obvious, these utterances are really about the child's own mental state. He or she is referring to the particular mental state, that of desire in this case, by using the concept of desire in the right way. It would not be possible to confuse these particular utterances with behavioral requests, as is the case with the flavor of GPRDs discussed above. This is partly the rationale for some of the analyses presented in the following chapter.

There are two more cases of GPRDs that do not lend themselves to confusion as to their status. One of them is illustrated by the following: Benjamin (27 months) "You wanted those shoes" or Deborah (23 months) "Maybe he wants this basket to lie on" or even Frances (39 months) "My daddy wants to talk to you". These examples clearly refer to the mental states of desire of other people, not the child's own and as such, are pretty clear as to their status. When children start using 'want' and they are talking about someone else, it is even more clear that they are in fact somehow reasoning by means of those desires.

The second case is quite clear, though little attention has been given to it in

the literature: the case of questions. Take for example the following utterances coded as GPRDs: Geoffrey (29 months) “D(o) you want go up with me?” or Ellen’s (17 months) “want do it?”. Questions, in my view, are a very good reflection of children’s understanding of the concept of desire. When they ask the question (and these questions are not formulaic language), they do so because the piece of information they are looking for is very important to their folk psychological reasoning. It is extremely important to know how people feel about things and whether they want these things or not. Thus, questions are taken to be a strong indication that children know the concept of desire, that definitely other people have those states and that knowing the contents of these desire states helps them navigate socially. When asking questions about the desire of others, children are in fact trying to ascribe desires to other people and to reason by means of them, although they may not know which desires to ascribe yet.

Both direct repetitions and to a lesser extent idiomatic expressions are more straightforward categories than GPRDs, as discussed above. Benjamin (23 months), for example, keeps repeating “I want beer” for about three or four turns in the dialog, even while doing something else or while being alone. Benjamin’s first use of the expression “I want beer” in the transcript may have actually been coded as a real GPRD or maybe a behavioral request (depending on whether the beer was in front of him or not, see the discussion below), but the next few times are merely playful repetitions of the first, meaningful utterance. These kinds of rep-

etitious utterances have not been assigned a GPRD status because the meaning of the desire term is clearly not one that evidences the child grasp of the concept the word stands for.

Idiomatic expressions, in turn, could in fact be somewhat misleading as to their status. It could be the case that by saying "want more", the child is indeed asking for more milk or more sweets or anything. Two things should be said about this: first, the total number of idiomatic expressions in the database is not very large. There are only a total of 20 IE (3% of the total want utterances) and the ages at which the children actually say these utterances range between 27 and 58 months, which should not critically influence the GPRD category results in any significant way, as if, for example, these utterances had appeared in the lowest age range (17-24 months of age). Second, the degree of inter-rater reliability was relatively high. Of the 20 utterances identified as IE, only 4 were rated differently by the two raters, and the 4 were rated as BR by one of the raters and as IE by the other. This suggests that it is not impossible to find agreement as to their status as IE. This agreement notwithstanding, there may have been a few false positives and false negatives in the set of utterances belonging to the IE category.

Behavioral requests are a problematic category in the sense that some utterances coded as BRs could potentially be confused with a GPRD with a physical object as the object of the desire and indeed vice-versa. There could be times in which GPRDs can be confused with BRs, given that the context did not make it

clear which category the utterance belonged to. This notwithstanding, as with idiomatic expressions, most, though not all of the utterances categorized as BRs were quite straightforward. Usually, the context makes it clear that the object the child is referring to in his or her expression or reference to desire is in front of him or her. According to Bartsch and Wellman, then, these utterances, should probably not be categorized as GPRDs because they may be no more than a pragmatic strategy for asking parents, etc. to give them something, without really entertaining the concept of desire at all. This is not to say, of course, that the child does not want what is in front of him or her, but that BRs may just be a simple expression of desire without necessarily using the concept of desire itself.

For several reasons, the false positives and false negatives criticism applies more strongly to BRs and GPRDs. These are, indeed, hard to tease apart. One of the analyses in this investigation directly addresses part of this problem. As we will see, only GPRDs for a physical object could be wrongly categorized as BRs. If we were able to get the same results by only considering other kinds of GPRDs (GPRDs for states of affairs or actions, for instance), then we would at least make sure that the GPRD category is not contaminated by BRs. As we will see, this is the case, and indeed a goal of this analysis.

It is obvious then that the set of categories chosen here (following Bartsch and Wellman) may be deficient in teasing apart utterances as belonging to one

or another category, generating both false positives and false negatives in all categories, but most importantly in the GPRD and the BR categories. In general terms, if the comparison group contains some items that are in fact members of the target group (or vice-versa), the comparison would be contaminated and would misrepresent reality. This is fair criticism. However, certain measures have been taken to keep the results as clean as possible of these contaminations.

Also, a more general point could be made about the nature of the data in these and other analyses. Data can be carved in many different ways, in potentially infinite numbers of ways, in fact. Because of the nature of data, false positives and false negatives will creep into any categorization scheme. In other words, any other categorization scheme will almost surely contain false positives and false negatives and, to be sure, there is no independent method of confirmation down the road to ensure that we “learn” from these categorization errors, as there is in other disciplines. By the same token, if everybody came up with their own set of categories, there would be no common ground for comparison.

In summary, it should be said that although there may be some categorization errors in the studies, there are several more advantages than disadvantages for sticking to the scheme proposed by Bartsch and Wellman: 1) there will be false positives and false negatives in any other categorization scheme, with no independent method of really knowing how many, 2) inter-rater reliability is quite high, even given the nature of the studies, 3) there already are results using these

same categories, which allows for comparing results, 4) the categories themselves have indeed undergone evolution as to what they pick out, going back to the seminal studies of Shatz *et al.* in 1983 and 5) previous studies using these same categories are in line with many of the results coming from the traditional experimental literature, which is a quite strong independent confirmation that the categories could not be *that* wrong. Compare this to the sole advantage of other categorization schemes in *maybe* containing fewer false positives and false negatives, with no way of knowing (even *a posteriori*) that this is true. I think the case is strong enough to keep the categories rather than refactoring them because another scheme *may be* more telling.

5.2 Study 1: Replicating Bartsch & Wellman's three hypotheses

Research on the understanding of desires using the NOS paradigm has yielded three major hypotheses that I will focus on in this replication/extension: Hypothesis One (H1): Children make reliable genuine references to desire at around 2 years of age; Hypothesis Two (H2): psychological talk using desires increases markedly at around the 24th month, reaching a peak at around the 33rd month to stabilize later on, when the concept of belief has found its way into the child's conceptual repertoire (Wellman et al., 1995; Wellman, 2001; Flavell, Flavell, Green, & Moses, 1990). Lastly, Hypothesis Three (H3) predicts that children

begin by expressing their own desires rather than referring to those of other people. All these hypotheses have been given strong empirical support by Bartsch and Wellman and have been discussed also by Harris (1996) and others, both in normal (Perner, 1991a, 1991b; Fodor, 1992; Cassidy, 1998) and abnormal populations (Tager-Flusberg, 1992, 1993; Rieffe & Terwogt, 2000). The objective of this study is to provide evidence for H1, H2 and H3.

Notice also that the analyses carried out in the present study go beyond mere replication. These analyses actually complement the results in Bartsch and Wellman's analyses in two ways: a) they were carried out with a different and, arguably, better controlled population of children (see section 4.1) and b) inferential statistics were used to check for significant differences, something that the first two analyses (Wellman & Bartsch, 1994 and Bartsch & Wellman, 1995) did not do.

5.2.1 Age as the independent variable

5.2.1.1 The data

The data come from the already identified 'want' utterances in the corpus described in Chapter 4. To recapitulate: there are 602 'want' utterances belonging to 12 children at six different age groups (in months): AGE1 (18-24); AGE2 (25-28); AGE3 (29-32); AGE4 (33-36); AGE5 (37-40); AGE6 (41-44). Unlike Bartsch and Wellman (1995), there are no "blank cells" in the data; that is, there are

developmental data for all 12 children at all ages.

5.2.1.2 Procedure

Each occurrence of 'want' in the corpus was coded as belonging to one of the five exhaustive and mutually exclusive categories defined in section 5.1 above; namely, GPRDs (section 5.1.0.1), BRs (section 5.1.0.2), DRs (section 5.1.0.3), IEs (section 5.1.0.4) and UU (section 5.1.0.5). To do this, utterances were not taken in isolation, but embedded in a window of the four previous and the four following utterances of the whole sample. However, sometimes this short context did not help defining which category the utterance belonged to. Thus, the whole transcript had to be analyzed in order to assign a category to the utterance in question. The box below shows an example of this contextual window.

MOT: Do you want to play?
CHI: play those .
CHI: one, two, three, go .
CHI: one, two, go .
CHI: want (a)nother one .
MOT: there isn't another one .
MOT: sorry .
MOT: I used them .
CHI: (a)nother one .

An independent rater, unaware of the hypothesis of the study, rated a subset of the data (10%=60 utterances), Kappa reliability tests yielded a high score, $\kappa=0.87$. Disagreements were resolved by discussion and, in the light of the dis-

cussions, there was a second coding pass to the whole data set.

5.2.1.3 Results

In total, there were a 602 utterances with the verb 'want'. Of these 602 utterances, 347 (58%, against the 74% in B&W) were genuine psychological references to desire, 106 (18%, against the 3% in B&W) were deemed uncodable and finally 149 (24%, against the 22% in B&W) were categorized as a communicative use of the word 'want'. Communicative uses of 'want', again, are those utterances that do not refer to a mental state of desire (either the child's or someone else) per se, but to other uses of the word.

Given that there were many more uncodable utterances than in the B&W's studies, the two coders were, it seems, overly conservative in what they took to be GPRDs. Proportions of the other conversational uses (BRs, DR's and IE's) remained practically the same in both this and B&W's studies.

In general terms, the results I obtained are quite similar to those of Bartsch and Wellman. Table 5.1 below shows the result of the coding efforts. This table only shows the counting of instances of the word 'want' qua genuine psychological reference to desires, I have made no reference yet to the communicative uses of 'want'.

Figure 5.1 shows the average number of GPRD utterances as a percentage of total utterances for each child at the different age groups. It is evident that

Child(*)	Age stages						Total
	AGE1	AGE2	AGE3	AGE4	AGE5	AGE6	
Abigail(f)	1	2	4	2	3	2	14
Benjamin(m)	6	3	1	19	5	7	41
Darren(m)	0	0	2	9	2	7	20
Debbie(f)	4	0	3	7	7	2	23
Ellen(f)	7	0	15	8	9	4	43
Elspeth(f)	0	2	12	5	4	4	27
Frances(f)	1	1	12	6	8	3	31
Gary(m)	3	2	2	19	8	15	49
Geoffrey(m)	0	1	18	10	2	4	35
Harriet(f)	2	3	2	24	0	5	36
Jack(m)	0	0	1	16	0	6	23
Jason(m)	0	1	0	0	0	4	5
Sum	24	15	72	125	48	63	347

Table 5.1: Total GPRD utterances available from each child by AGE.

the developmental picture I have obtained resembles the one in Bartsch and Wellman very closely, even the ranges of the percentages are similar (see Bartsch & Wellman, (1995), p. 73, Figure 4.2B). Reference to the mental state of desire seems to be present at the first age analyzed (AGE1), to increase slowly by AGE2 and then more drastically again at AGE3. The frequency of genuine reference to the mental state of desire seems to peak at around AGE4 and stabilize thereafter.¹

In order to test for significant differences in children's talk about desire states at any given age group, a repeated measures ANOVA with AGE as a 6-level variable was used. There was an overall significant main effect at $F(5, 55)=6.72$, $p<.000$. *post hoc* analyses using the Bonferroni criterion for significance indicated

¹That the results are so similar to those reported by Bartsch and Wellman is quite remarkable given the differences in the two studies: two different coders, a different population and even different techniques for individuating the utterances and coding them. This, I believe, lends a lot of credibility to the NOS paradigm.

Figure 5.1: GPRDs by age as a percentage of the total number of utterances for that child at that age. (Error bars are standard errors.)

that the average frequency of GPRDs for AGE1 ($M=0.51$, $SD=0.55$) and AGE2 ($M=0.96$, $SD=0.80$) were significantly lower than frequency of GPRDs at AGE4 ($M=4.28$, $SD=2.24$) and AGE6 ($M=3.54$, $SD=2.10$).

Of course, the results reported in the last section show the group data. Figure 5.2 on page 92 shows the data for every child in the study. Seven out of 12 children in the study are already making genuine references to desire in the first age group analyzed (age <24 months). By AGE3 (29-32 months), all but one of the children (Jason) are making genuine psychological references to desire. This is indeed quite early when compared to Bartsch and Wellman's data for genuine reference to beliefs. All children are making genuine references to thoughts and

beliefs between the 41-44 months interval; that is, a year after the appearance of genuine psychological references to desire. By the last age range analyzed (AGE6, 41-44 months of age), desire talk seems to have stabilized, even for Jason, who was a bit delayed at the onset.

This is more evident in Table 5.2 on page 93 below. Table 5.2 presents the data showing the age for the first transcript, the first (and third) use of 'want' regardless of its meaning and the first (and third) genuine use of the verb 'want' to talk about the mental state of desire. Half the children start using 'want' by their first available transcript; the other half by or shortly after their second birthday. All but one child (Jason, who seems to lag behind the group average all across the board) have made their third use of the word 'want' before their third year. When it comes to genuine references to desire, most children start talking about desire by their second year of life and they seem to be talking about desires reliably before the third year. It is important to notice that although the samples start at 17 and 18 months, only three out of 12 children have made their first reference to desire at that age. The remaining 9 have to wait until their second birthday to produce their first GPRD. That none of the children have made their third reference before the 24th month seems to confirm that the concept has been successfully established by that age and not before. Therefore, although genuine desire talk begins quite early indeed when compared to reference to beliefs and thoughts, it does not begin until the 24th month.

Figure 5.2: Individual differences of GPRD as a percentage of total utterances for that child at that age

I believe that B&W failed to see this phenomenon simply because there weren't enough data available at the earlier stages; half of their transcripts start after the 24th month. In the case of this study, 7 out of the 12 earliest transcripts start at age 18 months and the rest at age 17 months. Still, children in the sample analyzed in these pages refer to desire relatively reliably only after the 24th month.

Child	Age at First Available Transcript	First Use of Desire Term	First Genuine Reference to Desire
Abigail	1;5	1;5 (2;0)	2;0 (2;3)
Benjamin	1;5	1;5 (1;11)	1;5 (1;11)
Darren	1;6	1;6 (2;8)	2;8 (3;0)
Debbie	1;6	1;6 (2;6)	2;6 (2;6)
Ellen	1;5	1;5 (1;5)	1;5 (1;9)
Elspeth	1;5	2;0 (2;2)	2;2 (3;0)
Frances	1;6	2;0 (2;6)	2;0 (2;6)
Gary	1;6	1;6 (1;6)	1;6 (2;0)
Geoffrey	1;6	2;2 (2;2)	2;2 (2;5)
Harriet	1;6	1;7 (2;0)	2;0 (2;3)
Jack	1;5	2;0 (2;9)	2;5 (2;9)
Jason	1;6	2;3 (3;0)	2;3 (3;5)

Table 5.2: Ages at first available transcript, first (and third) use of desire terms and first (and third) genuine reference to desire.

5.2.2 Self versus other by age

5.2.2.1 The data

The data come from the ‘want’ utterances that were coded as genuine talk about desire states in the corpus described in Chapter 4. To recapitulate: there are 347 ‘want’ utterances that are genuine talk about desires and that belong to 12 children at six different age groups (in months): AGE1 (18-24); AGE2 (25-28); AGE3 (29-32); AGE4 (33-36); AGE5 (37-40); AGE6 (41-44).

5.2.2.2 Procedure

Each occurrence of ‘want’ qua GPRD in the corpus was coded as either belonging to the expression of the child’s own desire (self), or to refer to someone else’s desire (other). Some examples of expression of the child’s own desires and reference to other people’s desires are given below.

Abigail (39 months)

[Self] I want to go on this first

[Other] Perky wants to come in.

An independent rater, unaware of the hypothesis of the study, rated the whole database. Disagreements were resolved by discussion and, in the light of the discussions, there was a second coding pass on the whole data set.

5.2.2.3 Results

Figure 5.3 on page 96 shows the data for GPRDs of self and other as a percentage of total GPRD utterances by child by age. Most talk about desires is an expression of the child's own desire. Interestingly, however, the ratio of children's expression of their own desires does not seem to vary as a function of age. Being a ratio of the total number of identified GPRDs, then, something similar happens with children's reference to the desire of others. Figure 5.3 shows that reference to other people's desire is indeed present at AGE1 and that it does not really change that much from AGE1 to AGE6. Children's use of the term 'want' (in a way that shows they have understood the concept the word stands for) shows that the child's own desire utterances dominate over other utterances produced to talk about the desire of others. Indeed, more data are needed to support this claim.

Although this is a repeated measures study by design (with age and self versus other as the two variables (AGE*SelfOther)), the absence of some data points make the use of a repeated measures statistical analysis impossible. Thus, and in keeping with other studies in the literature (Tager-Flusberg, 1992), an independent-group ANOVA with age and SelfOther as the fixed factors was used to test for significance. A main effect of was found for SelfOther $F(1, 104)=426.412$, $p<.000$ and for the interaction between age and SelfOther references $F(5, 104)=4.755$, $p<.001$. There is no main effect for age.

Figure 5.3: Self versus other as a percentage of GPRD utterances. (Error bars are standard errors.)

In order to test this issue further, Table 5.3 on page 97 shows the data for children's age at first (and third) genuine reference to the desires of self versus other (the "—" symbol means that there was either no reference to desire at all or just that there was not a *third* reference).

As this table shows, expression of the child's own desires predominates: 88.2% versus 11.8% of reference to someone else's desire. In more detail: two out of 10 children (Debbie and Ellen) start referring to someone else's desires before they talk about their own; the remaining 10 children do the opposite. They start expressing their own desires before referring to those of other people. Obviously, then, there's a tendency for children to start talking about their own desires, in

Child(*)		No. of References	Age at First (and Third) References
Abigail	Self	13	2;0 (2;3)
	Other	1	3;1 (-)
Benjamin	Self	35	1;5 (1;11)
	Other	6	2;3 (2;11)
Darren	Self	19	2;8 (3;0)
	Other	1	3;0 (-)
Debbie	Self	18	2;6 (2;6)
	Other	5	1;11 (1;11)
Ellen	Self	38	1;9 (1;11)
	Other	5	1;5 (1;11)
Elspeth	Self	23	2;2 (2;6)
	Other	4	2;6 (3;2)
Frances	Self	24	2;0 (2;6)
	Other	7	2;11 (3;3)
Gary	Self	44	1;6 (2;7)
	Other	5	2;0 (2;9)
Geoffrey	Self	32	2;2 (2;5)
	Other	3	2;5 (3;6)
Harriet	Self	34	2;0 (2;3)
	Other	2	2;0 (-)
Jack	Self	21	2;5 (2;9)
	Other	2	2;9 (-)
Jason	Self	5	2;3 (3;5)
	Other	0	(5;0) (-)
Subtotal			
	Self	306	
	Other	41	
Total utterances		347	

Table 5.3: Total number of utterances and age at first (and third) genuine reference to the desires of self versus other

accordance with B&W's studies. The difference between the onset of talk about desire of self versus other is, on average, roughly 7 months for the first reference and 5 months for the third. In other words, children start talking about their own desires around 7 months before they do so with other people's desires (for their first reference to them) and around 5 months for their third reference. Interestingly, however, this is still earlier than their references to beliefs.

5.2.2.4 Conclusions

I think I have replicated B&W's results quite closely. I have shown that there is talk about the mental state of desire in the 18-24 month period and that there is a significant increase in talk about desires in the 33-36 month period, to stabilize later on. This is all in accordance to B&W's results. Therefore, two of the hypotheses presented in section 5.2; namely, H1 and H2 have been given some more empirical support.

I have shown that, also in accordance with B&W, children prefer expressing their own desires rather than referring to those of someone else. The third hypothesis in section 5.2 has also been given some more empirical support. I have also found that there possibly is an increase in talk about someone else's desire at age AGE5, a few months after the sudden increase in talk about the child's own desire and, allegedly, when talk about belief finds its way into the child's conceptual repertoire. This is, I believe, an interesting finding that will be left

for further study when the data become available. For now, by comparing B&W and possibly all the data in the false belief task (see Wellman et al., 2001) the concept of belief seems to be present at age 41-44; that is, around a year after the concept of desire.

The question that will occupy us further is this: how does [desire]_p develop from AGE1 to AGE4 such that there is such a significant increase in talk about it? I will try to answer this question on a variety of levels: a) a linguistic prowess level; b) by looking at the different objects that children desire; c) by looking at some possible metarepresentational issues that may develop.

5.3 Study 2: Extending Bartsch & Wellman

Children acquire language at very different rates. Thus, chronological age is not usually a good predictor of language competence (see Brown, 1973). Therefore, in this chapter, I have decided to analyze use of the internal state term 'want' as a function of mean length of utterance rather than age.

5.3.1 The issue with MLU (mean length of utterance)

Only two sets of studies using the NOS paradigm have categorized children's language samples as a function of MLU: Bartsch and Wellman (1995), Wellman and Bartsch (1994) and Tager-Flusberg (1990, 1992 and 1993, all of them reports on different aspects of the experiment carried out in her 1990 paper). I review

some of the key points of these studies below, starting with Bartsch and Wellman's work.

In their 1995 book, Bartsch and Wellman report on the relationship between internal (mental) state terms of belief and thought and MLU at 36 months of age. Not surprisingly, the conclusions they reached were that a) there was some evidence for a relationship between MLU and frequency of talk about thoughts and beliefs qua genuine psychological reference to these mental states ($r=.66$, $p<.06$) but also that b) there was no apparent relationship between MLU and first reference to thoughts and beliefs ($r=-.19$, n.s.). In other words, "MLU is related to how much these children talk about thoughts and beliefs but is relatively unrelated to the timing of this important milestone".

There are several problems with Bartsch and Wellman's use of MLU in their studies. First, they never mention explicitly how the MLU values of their samples were calculated; namely, whether MLU was calculated over the number of *words* in each utterance (MLU_w) or over the number of *morphemes* in each utterance (MLU_m). At first sight, it seems they calculated MLU_w rather than MLU_m due to the fact that they say, for example, that "MLU at 36 months ranged from 2.3 to 6.0 words per child". Despite this, I am not sure what method of calculation they used, or what they call "word" for that matter. Given this, I have decided, for practical purposes, to carry out MLU_w analyses of the child samples and carry out statistical tests to find out whether the differences are significant or not.

A second problem with Bartsch and Wellman's (1995) MLU studies is that they never talk about the relationship between MLU (of any kind) and the frequency of occurrence of terms associated with reference to the mental state of desire (terms like 'want'). As I remarked above, they confined their analysis to talk about thoughts and beliefs. Thus, one goal of the MLU study I carried out in this section was to do an MLU analysis that would, I hope, shed some light onto issues of talk using the desire term 'want' qua genuine talk about mental states and its relationship to linguistic ability/competence. A third problem I find in Bartsch and Wellman's (1995) MLU studies is that they calculated MLU (of whatever kind) only for children at 36 months of age. There is, therefore, no "developmental data" of talk about mental state terms (either belief or desire) and MLU as the independent variable.

5.3.2 MLU as the independent variable

5.3.2.1 Procedure

The 347 GPRD utterances already identified in section 5.2 were assigned to the different MLU groups identified in the previous section. Given that the utterances were in database form, they could just be assigned the MLU stage and then queried with respect to this variable.

5.3.2.2 Results

Table 5.4 shows the results of this recoding effort. Figure 5.4, in turn, shows the development of mean GPRDs by MLU stage, as a percentage of the total number of utterances for that child at that MLU stage.

Child	MLU Groups					Total
	MLU I	MLU II	MLU III	MLU IV	MLU V	
Abigail	0	1	8	3	5	17
Benjamin	1	—	14	26	1	42
Darren	0	—	2	24	—	26
Debbie	3	—	10	9	4	26
Ellen	5	4	21	20	—	50
Elspeth	2	—	—	22	12	36
Frances	2	6	9	17	—	34
Gary	3	—	4	43	—	50
Geoffrey	1	—	32	6	4	43
Harriet	3	11	—	22	2	38
Jack	1	9	6	7	5	28
Jason	1	—	4	10	—	15
Sum	22	31	110	209	33	405

Table 5.4: Total GPRD utterances available from each child by MLU.

The developmental picture is similar to that with AGE as the independent variable. An interesting similarity is that even by MLU I (with a total average of less than 1.76 words per transcript), desire talk is already present. After MLU I, specifically at MLU II there is a fairly marked increase in the frequency of talk about desire states. From then on, the frequency of desire talk seems to increase a small amount at MLUs III and IV and then, by MLU V, the frequency stabilizes at more or less the level of MLU II. A one-way independent group ANOVA was carried out and the effect of the independent variable was significant,

$F(4,37)=6.204$, $p < .001$. *Post hoc* analyses using the Bonferroni criterion for significance indicated that the average frequency of GPRDs for MLU I ($M=0.47$, $SD=0.40$) was significantly lower than frequency of GPRDs at MLU III ($M=3.33$, $SD=2.47$) and MLU IV ($M=3.76$, $SD=1.67$). No other significant effects were found.

Figure 5.4: Mean GPRDs as a percentage of total utterances by child by mean length of utterance (MLU). (Error bars are standard errors.)

5.3.2.3 Conclusion

From this study, evidence suggests that somewhere between MLU II and MLU III, children have acquired the linguistic capacities to talk reliably about desire. In fact, from around MLU II onwards, there does not seem to be much change in the frequency of genuine desires at the different MLU stages. This suggests again,

that by MLU II onwards, children can use their present linguistic capacities to talk about the mental state of desire.

5.3.3 Self versus other by MLU

5.3.3.1 The data

The data for this analysis come from the 'want' utterances that were coded as genuine talk about desire states in the corpus described above on page 73. To recapitulate: there are 347 'want' utterances that are genuine talk about desire states and that belong to 12 children at five different MLU groups: MLU I (more than an average of 1.76 words); MLU II (more than an average of 1.75 and less than 2.26 words); MLU III (more than an average of 2.25 and less than 2.76 words); MLU IV (more than an average of 2.75 and less than 3.5); MLU V (more than an average of 3.5 words).

5.3.3.2 Procedure

Each sample of child language was linked to the MLU value associated with it. The database of child utterances was linked to a "developmental data" database, which included the name of the particular file, the child's age (in months), the child's name and the MLU calculation for that file. Thus, for each child utterance which occurred in a particular file, there was a coding for self or other (see page 94) and there was an MLU value associated to it. The database was queried to

return, by name of child, the number of utterances that were present at each MLU stage.

5.3.3.3 Results

Figure 5.5 presents the results of children's mean expression of their own or reference to someone else's desires as a percentage of total number of GPRDs for that child at that particular MLU stage.

Figure 5.5: Self versus other as a percentage of GPRD utterances by MLU stage. (Error bars are standard errors.)

As in the case of age as the independent variable, children do seem to talk a lot more about their own desires, rather than those of somebody else. A 2 (self versus other) by 5 (MLU stages I-V) independent groups ANOVA was carried out on the frequency of talk about the mental state of desire (dependent variable).

The omnibus F test shows a significant main effect of expression versus reference to other people's desire, $F(1,66)=651.17$, $p < .001$. However, the difference in expression of their own and reference to other people's desires by the different MLU stages is non-significant, $p > .05$. There is also an interaction effect, $F(3,66)=6.59$, $p=.001$.

5.3.3.4 Discussion

The results suggest that, at least considering MLU stages as we have, children do not really undergo a development of the concept of desire for self versus other (there seems to be no developmental pattern at all). In other words, they seem to maintain the same proportion of GPRDs for self and other across MLU stages, though they favor talk about their own desire most of the time by far.

A word should be said about the interaction effect. It looks like using desire terms in their first person or third person conjugation changes with MLU. There is a reduced number of 'self' and a greater number of 'other' utterances by MLU V. This interaction is difficult to explain given that, even if it would in principle make sense that children become more adept at using desire terms to talk about other people, the same is not true when age is taken as the independent variable. If we compare Figure 5.5 with Figure 5.3, we can see that the latter shows the opposite interaction effect. In Figure 5.3 the frequency of children's use of desire terms decreases for 'other' as a function of MLU, while it increases for 'self'. Interpreting

this interaction goes beyond the objectives of this dissertation. Suffice it to say that children prefer talking about self when using desire terms and that they do so relatively consistent across both age and MLU. This interaction, however, is a very interesting result and merits further research.

5.3.4 Communicative uses of ‘want’ as a function of age and mean length of utterance

Bartsch and Wellman did not report what they found for codings other than GPRDs. This is unfortunate but understandable, given that such a detailed analysis for all those mental states terms identified would have made the project unattainable. Fortunately, given that my objective is much more modest, I am able to actually explore the relationship between GPRDs and other uses of want, what Bartsch and Wellman call “communicative uses”.

Given that behavioral requests (BRs), direct repetitions (DRs) and idiomatic expressions (IEs) should in principle not contain the difficulties that GPRDs present, we would expect that there should be no developmental change like there is for GPRDs. It is this set of hypotheses that I now turn to explore in more detail.

As in previous studies, it should be noted that there is still the possibility of cross-contamination of the categories (BRs could be wrongly categorized as GPRDs, or vice-versa), so the same provisos discussed in the replication analyses

(those analyses having age as the independent variable) also hold for this analyses.

5.3.4.1 The data

The data come from the already identified ‘want’ utterances in the corpus described in Chapter 4 and described in more detail above in section 5.1 on page 73. In this case, however, I will be paying attention to those utterances coded as BRs, DRs and IEs when compared to GPRDs.

5.3.4.2 Results

Figure 5.6 shows the development of children’s communicative uses of ‘want’ as a function of age. Although direct repetitions are present at the earliest MLU stage (MLU I)², neither idiomatic expressions nor behavioral requests appear in the first age stage sampled (AGE1=18-24). By AGE2, all three categories of communicative uses are present, although not dramatically different from the previous age. The frequency of DRs seems to grow and separate from the main trend at AGE3 and then again at AGE4, while at AGE3 both BRs and IEs are at the same level. Something indeed seems to happen at AGE4, when all three categories seem quite different in their frequencies, with DRs leading the frequency count, followed by BRs and IE in the last place. Both AGE4 and AGE5 seem to show the new stability of these categories. It seems then that after AGE4, all

²It is important to remember that the names of MLU stages are arbitrary. In this case, MLU I means (arbitrarily) less than 1.75 words per utterance.

three kinds of communicative uses of 'want' stabilize. However, these observations are somewhat mitigated by the statistical analyses described below.

Figure 5.6: Frequency of the communicative uses of 'want' as a percentage of total utterance and as a function of age. (Error bars are standard errors.)

To test for significant effects, a repeated measures ANOVA with age as the within-subject variable was carried out for each communicative use of 'want'. The tests show that, taken one by one, there is no significant effects for age and each of the communicative uses of 'want': $F(5, 55)=1.954$, $p>.05$, n.s. for behavioral requests, $F(5, 55)=1.411$, $p>.05$, n.s. for direct repetitions and $F(5, 55)=0.504$, $p>.05$, n.s. for idiomatic expressions. However, a repeated measures ANOVA with age and communicative uses as within-subjects variables yielded a signifi-

cant effect for both age $F(2.96, 32.61)^3=3.174$, $p=.038$ and communicative uses $F(2, 22)=12.708$, $p<.000$. No main effect was found for the interaction between age and communicative uses $F(3.36, 36.97)^2=0.682$, $p>.05$, n.s. However, when GPRDs are included, the picture changes dramatically. Now there is indeed a significant interaction between age and communicative use, $F(15, 165)=4.861$, $p<.004$. The results show that indeed there is a different developmental course for GPRDs than for the other communicative uses of 'want'.

Figure 5.7: Frequency of the different communicative uses of 'want' as a percentage of total utterance and as a function of MLU compared to GPRDs. (Error bars are standard errors.)

The results are quite similar when taking MLU as the independent variable.

³ F corrected for sphericity by Greenhouse-Geisser.

Both the first (MLU I) and last (MLU V) stages in Figure 5.7 are roughly the same when it comes to the frequency of communicative uses of 'want'. Behavioral requests (BRs) and direct repetitions (DRs) seem to progress and differentiate themselves from idiomatic expressions (IEs) by MLU II. By MLU III, all three categories seem at different levels, with DRs being the most frequent in the child's vocabulary, followed by BRs and lastly by IEs. By MLU IV and V, all communicative uses exhibit lower frequencies and children use each different communicative sense of 'want' at roughly similar frequencies. Even at MLU I, children are already talking about the psychological state of desire. At all other MLU stages, children's talk about the mental state of desire seem quite constant and always well above these children's talk about the other communicative uses of 'want', those that are not about the mental state of desire.

A battery of inferential statistics seem to corroborate a good part of the hypotheses put forward by visual inspection. A four (the use of 'want': BRs, DRs, IEs and GPRDs) by 5 (MLU Stage) independent groups ANOVA was carried out. Frequency of talk about the relevant communicative use of 'want' was the dependent variable and pairwise comparisons were carried out using the Bonferroni criterion for adjustment at the $p < .05$ level. The omnibus F tests yielded significant main effects for both communicative use, $F(3,164)=66.03$, $p < .001$ and MLU stage, $F(4,164)=10.79$, $p < .001$, and a significant interaction between MLU and communicative use, $F(12,164)=5.122$, $p < .001$. For our purposes here,

it is important to report the pairwise comparison corresponding to the different uses at the different MLU stages and, in particular, the comparisons between communicative uses (BRs, DRs and IEs) versus GPRDs. Interestingly, there are no significant differences between the three communicative uses of want (BRs ($M=0.09$, $SD=0.22$), DRs ($M=0.10$, $SD=0.15$) and IEs ($M=0.02$, $SD=0.09$)) and GPRDs ($M=0.47$, $SD=0.40$) at MLU I. However, there are significant differences between communicative uses of 'want' and GPRDs at all other MLU stages.

As a general remark, it seems that the NOS paradigm is indeed sensitive to different developmental pictures. Thus, H1 seems to be given some more empirical support by comparing GPRDs to the other "communicative uses" of the term 'want'. Yet, it may be the case that some of the utterances were mistakenly assigned to GPRDs, when they should have been assigned to BRs. Since the only time that this could happen is when children are referring to objects that are present, in front of the child at the time of the desire reference, then it could be telling to separate those instances; namely, physical objects, and see the developmental trend for other, more complex desires. The following study looks at this and a related issue.

5.3.5 Desires And Their Intentional Objects as a function of age and mean length of utterance

Given an unexpected success on the part of autistic children in understanding desires while failing to understand other, allegedly similar mental states – particularly, that of belief (Baron-Cohen, 1995; Baron-Cohen et al., 1985) – Tan and Harris (1991) have suggested that it could be the case that children with autism suffer from a delay in understanding those mental states that involve intentional objects that do not really exist, intentional objects that have what the authors call “intentional inexistence”, following Brentano (1874). For example, Tan and Harris (1991) claim that there is a difference between the following two thoughts: (a) *John wants his dog* and (b) *John wants a dog*; the latter does not imply any particular dog or any dog at all that would satisfy this desire (made explicit by the use of the indefinite article ‘a’); while the former talks about a particular dog that the child knew to exist (made explicit by the use of the possessive adjective ‘his’). Tan and Harris (1991) speculate that it may be the case that it is not necessarily the mental state that children with autism fail to grasp, but rather the intentional objects that those mental states are directed toward.

Although this hypothesis was put forward for the population of children with autism (about whom I will have something to say later on), it is not far-fetched to suppose that the same happens to non-autistic children somewhere along the developmental path. In other words, it might be the case that normally devel-

oping children fail to grasp the intentional objects of a given mental state – in the case in which the intentional objects do not actually exist physically– rather than the concept of the propositional attitude itself.

The next study aims at analyzing the intentional objects of desire utterances. First, I will analyze, in general terms, what kinds of intentional objects children talk about when they refer to desires at the different ages (whether, for example, they talk about physical object, actions or states of affairs indistinctly along the developmental path). Second, I will analyze the hypothesis that children may have problems referring to actions or states of affair that are not currently the case, as proposed by Tan and Harris (1991).

5.3.5.1 The Data

The data come from the identified GPRD utterances in the corpus already described in the previous chapter. To recapitulate: there are 347 GPRD utterances belonging to 12 children at six different age groups (in months): AGE1 (18-24); AGE2 (25-28); AGE3 (29-32); AGE4 (33-36); AGE5 (37-40); AG6 (41-44).

5.3.5.2 Procedure

I will take into account three different, linguistically determined kinds of intentional objects plus a “no object” category that children use when they communicate their own or refer to someone else’s desires. The categories are not arbitrary,

they were actually discussed, though not formalized in any way, in Bartsch and Wellman (1995).

Again, each occurrence of 'want' qua GPRD in the corpus was coded as belonging to one of four exhaustive and mutually exclusive categories: as talking about physical objects, as talking about states of affairs, as talking about actions and, lastly as talking about no particular object. Needless to say, these guidelines are just that, guidelines. Sometimes the context warranted deviating from their precepts, but it is assumed they helped reduce potential individual differences between the two raters.

Guidelines for recognizing physical objects

First, children could be talking about a desire for physical objects ("I want a marble"). This may be the simplest kind of intentional objects that children could be talking about; so much so that, sometimes, researchers want to exclude them from genuine psychological references to desire and assign them the name "behavioral requests" to imply that these are not really desires (see Bartsch & Wellman, 1995 and all the literature based on Shatz *et al.*'s (1983) studies).

Gpo1: want (+ 'some') (+ adjective) + noun

e.g. I want mommy. (want + noun)

e.g. wanna bike. (want + noun)

e.g. He wants some milk. (want + 'some' + noun)

e.g. He wants his dinner. (want + adj. + noun)

Gpo2: want + pronoun (+ prepositional complement)

e.g. I want this. (want + pro.)

e.g. I want one of these. (want + pro. + prep. comp.)

Gpo1 states that to recognize the object of the desire proposition as a physical object, the utterance should have at least the lexical item 'want' (as a verb) followed by an optional lexical item 'some', followed by an optional adjective and finally an obligatory noun (like 'milk', or 'car', etc.).

Gpo2, in turn, states that to recognize the object of the desire proposition as a physical object, the utterance should have at least the lexical item 'want' (as a verb) followed by an obligatory pronoun and an optional prepositional complement. Since these categories are a subdivision of GPRDs, it is assumed that the object the child is referring to is not directly accessible to him or her at the time.

Guidelines for recognizing actions

Second, children could also potentially be talking about desires for actions: an expression of desire would read something like "I want to run to school" or "He wants to run to school" as a reference to someone else's desire. This is a relatively straightforward category and one that is obviously either an expression or

a reference to a desire, given that if anyone utters such a sentence then he or she (or someone else) is either not running to school or has been stopped or is being told to stop running to school, with the final result of the desire not being satisfied. The only instances in which I can foresee the expression or the reference to a desire for an action as not being a genuinely psychological (i.e. using a desire term without having the concept of desire) is in either a direct repetition of someone else's utterance or as an idiomatic expression, like the lyrics of a song learned by heart or a poem or a saying. For the most part, though, I believe the communication of a desires for an action will be "genuinely psychological" almost all of the time. A goal of the present study is to analyze whether using 'want' standing proxy for the concept of desires for actions appear before, at the same time or later than expression or reference to desires for physical objects.

Gal: want + to + verb (+ complement)

e.g. I want to go home. (want + to + verb + comp.)

e.g. I want to go. (want + to + verb)

e.g. He wants to write there. (want + to + verb + comp.)

Gal states that to recognize the object of the desire proposition as a desire for an action, the utterance should have at least the lexical item 'want' (as a verb) followed by either a to-infinitive clause (i.e., to + verb like 'go', 'stay', etc.) and an optional complement (e.g. an adverb phrase).

Guidelines for recognizing states of affairs

Third, children could be communicating desires for states of affairs (talking about M&Ms, a child may say: "I want them red!"). Although resembling reference to desires for physical objects, desires for states of affairs are more complex both linguistically and conceptually. Only a particular state of affairs would satisfy the desire; for example, that all M&M's be red. If there is one that is blue, the desire will remain unfulfilled. This is the kind of desire that seems to separate desire qua mental states from simple internal states like hunger or anger. In hunger, any M&M would do, however, the desire for red M&Ms is much more specific, and specific in a sense that cannot be ascribed to simple bodily states like hunger (Wellman, 1990). It would be quite strange if, when in the middle of the Sahara dessert, not having drunk any liquid for three days, our body asked us for Ginger Ale and rejected drinking a passing Bedouin's water because it does not satisfy the desire! This issue will be pursued further in the discussion section. As is the case with reference to desires for actions, reference to some desired state of affairs can only be taken to be genuine psychological references to the mental state of desire. Again as in desires for actions, the only instances in which I can foresee the expression or the reference to a desire for an action as not being a genuinely psychological (i.e. using a desire term without having the concept of desire) is in either a direct repetition of someone else's utterance or as an idiomatic expression. For the most part, the communication of a desire for

an action will be considered to be “genuinely psychological”.

Gsoa1: want + accusative pronoun + prepositional complement

e.g. I want me in there. (want + acc. pro. + prep. comp.)

Gsoa2: want (+ pronoun) ((*adverb) $\wedge$ adjective)

e.g. I want it white. (want + acc. pro. + adj.)

e.g. I want it now. (want + acc. pro. + adv.)

e.g. I want some more. (want + adv. + adv.)

Gsoa3: want + (pronoun $\wedge$ noun) + ‘like’ + complement

e.g. I want it like Darren’s. (want + acc. pro. + like + comp.)

e.g. I want my milk like Darren’s. (want + noun + like + comp.)

Gsoa4: want + (pronoun $\wedge$ noun) + to + verb (+ complement)

e.g. I want him to go home. (want + acc. pro. + to + verb + comp.)

e.g. I want Peter to come. (want + object + to + verb)

Gsoa1 states that to recognize the object of the desire proposition as a state of affairs, so yielding a desire for a states of affairs, the utterance should have at least the lexical item ‘want’ (as a verb) followed by an accusative pronoun (it, him, her, etc.) and any prepositional complement (i.e. in, of, etc. plus a noun phrase).

Gsoa2 states that to recognize the object of the desire proposition as a state of affairs, the utterance should have at least the lexical item ‘want’ (as a verb)

followed by an optional pronoun plus either one or more adverbs ('some' or "some more" - that is what the star (*) means) or an adjective. The plus sign between the parentheses means that whatever is in parentheses with it is optional. On the other hand, if the plus sign is outside the parentheses then that means that there is an option between parentheses, but one of the items is in fact obligatory. Options are marked with the wedge symbol ($\wedge$).

Gsoa3 states that to recognize the object of the desire proposition as a state of affairs, the utterance should have at least the lexical item 'want' (as a verb) followed by either a pronoun or a noun (necessarily one of those) plus the lexical item 'like' plus an obligatory complement.

Finally, Gsoa4 states that to recognize the object of the desire proposition as a state of affairs, the utterance should have at least the lexical item 'want' (as a verb) followed by either a pronoun or a noun (and necessarily one of those) plus a to-infinitive clause (to go, to stay, etc.) and an optional complement.

Guidelines for recognizing "no object" utterances

Lastly, children may be talking about desires with no linguistically realized object, desires which I could have still categorized as being GPRDs. In this case, these instances will be considered in yet a fourth, different category, even though by context it might have been possible to answer whether the desire is for one or another kind of category (physical objects, actions or states of affairs). This is

done as a precaution so not to pollute the results with bogus data points and just in the spirit of being conservative as to what category the desire might belong to.

Gnol: want (+ to) + $\emptyset$

I want. (want + $\emptyset$)

I want to. (want + to + $\emptyset$)

Gnol states that to recognize the object of the desire proposition as a no-object, the utterance should have at least the lexical item 'want' (as a verb) followed by either a to-infinitive clause without the verb (to go, to stay, etc.) or nothing else. These are linguistic rules, so the empty set symbol ($\emptyset$) stands for the non-linguistic realization of the object, but this does not mean that the desire is for no-object, that would make little sense. I am talking here about desires whose objects are not specified.

To summarize, then, the present analysis aims at exploring the following empirical questions: What kinds of entities (physical objects, actions or states of affairs) do children talk about when they refer to desires? Is there any correlation between children's chronological ages and the kinds of entities they talk about when they either express or refer to desires? Are there any developmental characteristics that are worth discussing? For example, do children start talking about physical objects, move on to talk about states of affairs to end up talking

about all three? Is there any developmental difference at all among children's talk about the three kinds of entities? Do children prefer to express desires for physical objects while referring to desires about states of affairs or vice-versa? In other words, is there any correlation between expressing and referring to desires and the objects of those desire utterances? Given that, to the best of my knowledge, no study has been carried out in order to analyze desire objects, no hypotheses will be put forward. This will be, out of necessity, an exploratory study.

These guidelines for identifying desire objects are simply that and nothing more. There were instances in which these decision algorithms could not be applied. And even when they could be applied rather clearly, the whole transcript had to be analyzed anyway for possible nuances in the context of the utterance. These nuances would have yielded an utterance classified as an action, for example, instead of a state of affairs, had the context not been analyzed. The categories for the utterances about which we were unclear were resolved by discussion between the author and an independent rater, or conservatively categorized as a no-object utterance.

In any case, to be conservative, although identifying utterances as falling into one or another category was straightforward, a second coder was asked to code 100 utterances and *kappa* reliability tests performed, yielding a value of ($\kappa = .94$).

5.3.5.3 Results

Table 5.5 below shows the number of utterances containing a desire term for each of the categories of objects: physical objects, actions, states of affairs and desire terms with no particular object.

Object	Age Groups (in months)						Total
	18-24	25-28	29-32	33-36	37-40	41-44	
PHYs	14	10	31	55	15	28	153
ACTs	7	3	30	43	22	15	120
SOAs	2	1	9	20	9	17	58
NOBs	1	1	2	7	2	3	16
Total	24	15	72	125	48	63	347

Table 5.5: Number of desire utterances for physical objects (PHYs), actions (ACTs), states of affairs (SOAs) and GPRDs with no object in them (NOBs) by age

Table 5.6 shows the means and the standard deviations at the different age stages for all kinds of objects, except "no objects", which I left out because the number of GPRDs with no object was so low that it just evidenced floor effects. It is indeed quite rare to find GPRDs with no object associated to them.

Figure 5.8 on page 126, in turn, shows the different kind of desire objects as a percentage of the total number of GPRD utterances for that child at that age⁴. When talking about mental states of desires (GPRDs), children at the first age analyzed (18-24 months) refer to desires for physical objects around 70% of the time. Indeed, at this early stage, they already talk about actions and states

⁴There are no error bars to make the figure more readable. The descriptive statistics for this study are given in the results section.

of affairs, but considerably less, at 17% and 7% respectively. This same picture remains constant until AGE3 (29-32 months), when children start increasing the proportion of desire talk about actions and states of affairs. At this stage, children are talking about desires for physical objects around 50% of the time, while they refer to desires for actions and states of affairs around 35% and 13% of the time, respectively. Talk about the different objects of desire clauses seem to stabilize after AGE4 (33-36 months) and they are much the same at AGE6 (41-44 months, the last age group analyzed), when children refer to desires for physical objects around 36% of the time, 30% of the time they talk about actions and, lastly, 29% of the time they talk about states of affairs.

		Age Groups (in months)					
Object		18-24	25-28	29-32	33-36	37-40	41-44
PHYs	Mean	70.40	72.91	50.35	42.62	26.64	36.46
	Std. Dev	29.41	39.77	32.22	12.93	21.06	24.71
ACTs	Mean	17.68	14.58	35.50	35.89	50.12	30.09
	Std. Dev	25.34	27.36	35.11	16.17	26.00	16.53
SOAs	Mean	7.14	6.25	13.13	15.86	20.59	21.12
	Std. Dev	13.11	17.67	19.80	14.96	21.05	21.55

Table 5.6: Mean and standard deviations of the frequency of physical objects (PHYs), actions (ACTs) and states of affairs (SOAs)

Interestingly, when compared to the developmental picture of the other kinds of desire talk (talk about desires for physical objects, actions and states of affairs) talk about desires with no particular object in them is quite stable throughout the different ages analyzed; starting, at AGE1 with 4.7% and ending, at AGE6,

at 4.3% (and ranging from 1% to 6%).

A 3 (object: physical, actions or states of affairs) by 6 (AGE) independent groups ANOVA was carried out, with frequency of occurrence as the dependent variable (see Table 5.6). This analysis yielded a main effect for object, $F(2,156)=29.33$, $p<.001$ and an interaction effect between objects and age, $F(10,156)=4.43$, $p<.001$. The interaction seems to be due to the fact that while both ACTs and SOAs, talk about physical objects decreases. By visually inspecting the data, we see that the the three categories have several distinctive characteristics. PHYs begins with a higher frequency and reaches the highest level at 33-36 months of age; ACTs follow closely behind PHYs and also reaches its peak at 33-36 months; SOAs instead start with the lowest frequency and takes longer to increase and reaches a peak (albeit at a level roughly half that of PHYs) at 33-36 months. What happens after 33-36 months is harder to interpret. It's not clear that what happens 37-40 months is a sampling anomaly or a "real" measure of the frequencies of these categories. At 41-44 months, it looks like the relative ranking of frequencies from the first 4 sampling periods is maintained, but the differences between the three categories are significantly reduced ⁵. *Post hoc* comparisons using the Bonferroni procedure make evident that children refer to desire for physical objects significantly less by the last two age groups; i.e. by AGE5 ($M=26.64$, $SD=21.06$) and AGE6 ($M=36.46$, $SD=24.71$) when compared to the earlier four age groups, AGE1 ($M=70.4$, $SD=29.41$), AGE2 ($M=72.91$,

⁵I am grateful to John Logan for useful discussions about these results.

$SD=39.77$), AGE3 ($M=50.35$, $SD=32.22$) and AGE4 ($M=42.62$, $SD=12.93$).

There are no significant differences for either ACTs or SOAs when plotted across age. This may mean that children do refer to these objects by the earlier ages and that they prefer referring to physical objects at the earlier stages for only a pragmatic reason or maybe just a matter of frequency. By AGE3 however, the frequency of desire for all these different objects has stabilized. The only difference that remains is at AGE5 between ACTs ($M=50.12$, $SD=26$) and SOAs ($M=20.59$, $SD=21.05$).

Figure 5.8: The different kind of desire objects as a percentage of the total number of GPRD utterances for that child at that age (Error bars have been omitted for readability)

An important by-product of this study is that it permits carrying out some further analyses on the reliability of the results already discussed in Chapter 1.

The communication of desires for physical objects are usually met with some degree of doubt as to their status as reference to people's mental states. This hesitation stems from the fact that desires for physical objects, when the objects alluded to are in plain view, seem to be a sort of communicative act to request that object without it really being a communication of a mental state (see 5.3.5.2 and Tan & Harris, 1991). If this were the case, then what we would need is a way to separate physical objects (for which we are sometimes unsure of their representational status) from other kinds of propositional objects like actions and states of affairs, which do indeed stand for nonexistent propositional objects, as we discussed above in section 5.3.5 on page 113. What was done, then, was to add the elements in actions and states of affairs and look at the developmental trend. Arguably, if the results show roughly the same developmental trend as observed in 5.1 on page 90, I would be able to make a stronger case that, even without counting the (misleading) case of physical objects (that is, only counting actions and states of affairs, which are, we are sure, reference to mental states), the developmental trend is the same.

This is, in fact, exactly what happens. As we can see in figure 5.9, when we add actions and states of affairs, we get roughly the same developmental trend observed for the full GPRD dataset. As in my replication of B&W, in order to see whether there is, in fact, difference in children's reference to either desire for physical objects compared to desires for either SOAs and ACTs

(added together), a 2 (object: PHYs versus ACTs + SOAs) by 6 (Age) independent groups ANOVA was carried out. A significant interaction effect was found, $F(5,104)=8.72$, $p=.001$. As in Bartsch and Wellman's studies, ACT+SOA desire talk seems to be present at the earliest ages analyzed (18-24 months) but ACTs and SOAs seem to increase while PHYs decrease with age. Desire for SOAs plus ACTs seem to peak at around AGE5 to decrease from then on. There appears to be a stabilization at around AGE6 (at 41-44 months of age).

Figure 5.9: Physical objects (PHY) compared to states of affairs plus actions (SOA+ACT) as a percentage of the total number of utterances for each child at each particular age.

5.3.5.4 Conclusion

Prior research on children's reference to mental states had focused primarily on coding the function of whole utterances containing lexical items associated with certain internal (mental) states. In other words, researchers have focused on deciding whether or not these utterances communicate thoughts that involve someone's mental state (Shatz et al., 1983). However, only limited attention had been paid to the "innards" of these utterances, what I will call a propositional attitude's "intentional object", following Fodor (1992). In similar words, researchers have concentrated mostly on deciding whether "I believe it's raining" and "I want it to rain" stand proxy for a thought containing the concepts of desire or belief (rather than being behavioral requests, direct repetitions, etc.). These researchers have not paid attention to the propositions inside the main clause of the utterance (e.g. "it is raining" in both the 'want' and the 'believe' utterances above) (Harris, 1996); that is, their intentional objects.

The studies above suggest that children seem to start talking about physical objects rather than actions or states of affairs. This trend changes around AGE3 (29-32 months) and talk about actions and states of affairs start appearing more often to stabilize at the last age analyzed, AGE4 (41-44 months of age). At this age, children seem to be talking about physical objects, actions and states of affairs rather indistinctly; that is, not really preferring one or the other. Genuine desire talk without any intentional object is stable throughout the child's

development.

A further result is that, even if we consider physical objects to be a rather ambiguous category for GPRD, we cannot do so with actions and states of affairs. The latter (actions and states of affairs) are, inherently, more abstract than physical objects and are thus more readily categorized as showing a grasp of the concept of desire. Thus, the working assumption was that actions and states of affairs taken together and analyzed together would give a stronger index of children's development of the concept of desire. The developmental trend obtained and the results of the statistics replicated the results in Chapter 1 quite closely, which I believe lends further support to the findings in that chapter.

5.4 General Conclusions

I have described a multitude of different analyses in these pages. I will try to focus on the ones that are important for our purposes here while staying true to the spirit of the results of each of the analyses.

Firstly, the results of the canonical battery of studies on the development of expression and reference to genuine desires, where the child's age is taken as the independent variable, yielded results that are very similar to those reported by Wellman and Bartsch (1994) and Bartsch and Wellman (1995) on the major points. Children develop their capacity to talk reliably about desire more or less at 29-32 months, remaining relatively constant at later stages. However, although

with much less frequency, they do refer to the mental state of desire by the first age analyzed, and definitely by 2 years of age.

There is a particular point that is worth emphasizing: while it is true that children prefer to use 'want' to talk about their own desires rather than the desire of others, this effect is quite constant across ages. That is, there does not seem to be a developmental trend in reference to the desires for self vis-a-vis other, so again, while it is true that children talk more about their own desires, they do so across the board. I would like to take this finding to mean that children can be referring to other people's desire without much difficulty, but that it may be more pragmatically useful to communicate their own desires. They make them explicit so that adults in their environment can fulfill them (see section 5.3.3).

Now, what do the different extensions to Bartsch and Wellman's work tell us and how do they fit in the larger picture? The most important issue is the switch to mean length of utterance (MLU) as the independent variable, rather than age. The first striking result is that by MLU II (an average of more than 1.75 words but less than 2.26 words), children seem to have reliably acquired the linguistic competence required to express or refer to desire. After MLU II, children's use of 'want' in a way that shows the grasp of [desire]_p stays relatively stable.

The picture of the development of desire talk about self versus other presents roughly the same picture as observed when age served as the independent variable. In other words, when children have the linguistic capacity to refer to their own

desires, they have the capacity to refer to other people's desire as well. As previously argued, that they do not might simply involve a matter of pragmatic convenience, rather than conceptual shortcomings. Children usually refer to their own desires more often because that fulfills their needs at that particular point in time, and thus they are worth expressing.

So far, results suggest that (a) there is some sort of development of talk about genuine desires and that (b) genuine desire talk is, for all practical purposes, independent of whether utterances are about the child's own or other people's desires. The former result (a) would not be as interesting if we found out that other identified communicative uses of 'want' (rather than using 'want' to refer to [desire]_p) undergo a similar developmental change among themselves, but different from GPRDs. Interestingly, results show that there are no differences developmentally if the different communicative uses were considered individually. In other words, DRs, for example, do not differ from Age 1 to say Age 6 at any of the stages, with the same holding true for MLUs. Interestingly enough, except for MLU I, the frequency of GPRDs is always significantly higher than the other communicative uses. This suggests that GPRDs are somewhat different from the other communicative uses of desires. In fact, this result suggests two things: first, the NOS paradigm seems to be capturing a difference when comparing those 'want' words that refer to [desire]_p and those which do not. The second, and possibly most important finding is that, although behavioral requests should use the same

syntax, their frequency is always lower (except at MLU I) than GPRDs. This suggests that something different than mere syntactic structure is at work and thus would suggest some sort of conceptual development.

The last study in this series was designed in order to make sure that when children talked about a desire, they were in fact taking the proposition inside the desire clause as false, as we saw in previous chapters. To do so, they should take the intentional object of the desire proposition as an action or state of affair that is not currently the case. In this last study, the focus was shifted from all possible GPRDs to only those desire expressions that referred to actions or states of affair. If states of affairs and actions are combined and their frequency plotted by age and MLU, the same developmental picture as that obtained for all GPRDs taken together is observed. Results suggest that children start talking about desires for objects, rather than actions and states of affairs to finally, by age 44 months, talk about each kind of intentional object at about the same frequency.

CHAPTER 6

Desire in autistic children

The objectives of this chapter are twofold: the first objective is to provide, post-facto, a control group of normal children to Tager-Flusberg's (1993) studies in order to find out whether there are differences between non-autistic and autistic groups when it comes to talk about the mental state of desire. The rationale for this post-facto comparison of normally developing children with autistic children was that it may be the case that Down's syndrome children, to whom the autistic children were paired to in the original Tager-Flusberg studies, might themselves suffer from a developmental delay that may somehow mask a real difference between the autistic and Down's syndrome children when compared to normal children. The second objective was to extend the Tager-Flusberg's studies and apply the new comparisons introduced for normal children when it comes to talk about desires, comparisons that would hopefully shed some light on the representational status of the concept of desire in the autistic group. Thus, I compared normally developing and autistic groups of children in their communicative uses of the word 'want' and their use of the different kinds of objects of

desire: desire for physical objects, actions or states of affairs. The comparison itself between these two groups of children will have to wait until next chapter, however. Here, I will only characterize autistic children's talk about the mental state of desire.

6.1 Examining Bartsch & Wellman's three hypotheses with autistic children

A bit of history: the first analysis conducted (Tager-Flusberg et al., 1990) was designed to find out about the language abilities of the different groups and to chart the participants' MLU over time. According to the authors, MLU_m values were calculated automatically using the SALT program (Miller & Chapman, 1985). The authors concluded that "the MLU data from the 6 autistic children show widely different developmental patterns", with some children achieving normal levels, others showing a decline and still others showing a slow development. With regards to the Down's syndrome children, the findings were similar, except that no Down's syndrome child showed an irreversible decline in MLU as one of the children in the autistic sample did.

More in line with the present investigation, Tager-Flusberg (1993) reports findings on the relationship between MLU and talk about desire terms. She concluded that there was no difference between autistic and Down's syndrome

children when it comes to their talk about mental states of desires.

There are, to my understanding, two important issues with the Tager-Flusberg studies. First, I believe that while it is true that there is no difference in talk about desire for MLU Stages I to IV, there does seem to be a difference between the samples belonging to the autistic versus the Down syndrome children at Stage V. While Down's syndrome children evidence a small decline in desire talk (just like normal children in Bartsch & Wellman 1994, 1995, as well as in my own preliminary frequency studies), autistic children don't show this decline. On the contrary, it seems their talk about desires is ever-increasing, rather than declining.

Secondly, it is hard to make sense of these results because a normal control group was not included in the study. To be fair, Tager-Flusberg did have a control group, but this was a group of children with Down's Syndrome. This is a population that, although presenting several marked deficiencies like delayed development and lower IQ (even when compared to autistic children), they do not seem to show representational deficits. Although inherently interesting, the results in Tager-Flusberg et al. (1990) would have been easier to interpret had there been a control group of normal children.

6.2 Study 1: GPRDs in autistic children

The first analysis carried out in this chapter was a replication of those studies pioneered by Wellman and Bartsch, Baron-Cohen (1994, 1995) and by Tager-

Flusberg (1992, 1993). Here, I used the same sample of children and redid the analysis that Tager-Flusberg carried out some time ago. The same data were used, but a new coder recoded the whole database. If the same findings were obtained, then we could be sure that the NOS paradigm has indeed relatively good inter-rater reliability. Also, new studies were carried out with the same data in order to compare autistic children's grasp of the concept of desire to their normally developing counterparts.

6.2.1 Coding GPRDs: MLU as the independent variable

6.2.1.1 The data

The data came from the already identified 'want' utterances in the corpus described in Chapter 3. To recapitulate: there were 931 'want' utterances belonging to 6 children at four different MLU groups: MLU I (an average of less than 1.76 words); MLU II (an average of less than 2.26 words and more than 1.75 words); MLU III (an average of less than 2.76 words and more than 2.25 words); MLU IV (an average of less than 3.5 words and more than 2.75 words). There was, unfortunately, no data for MLU V.

6.2.1.2 Procedure

The same procedure used in recognizing GPRDs in previous studies was also used here. Each occurrence of 'want' in the autistic corpus was coded as belonging

to one of the five exhaustive and mutually exclusive categories defined in section 5.1 above; namely, GPRDs (section 5.1.0.1), BRs (section 5.1.0.2), DRs (section 5.1.0.3), IEs (section 5.1.0.4) and UU (section 5.1.0.5). To do this, utterances were not taken in isolation, but embedded in a window of the four previous and the four following utterances in the sample. However, sometimes this short context did not help defining which category the utterance belonged to. Thus, the whole transcript had to be analyzed in order to assign a category to the utterance in question. The box on page 87 shows an example of this contextual window. A second coder, unaware of the research hypotheses, carried out a second set of codings of 10% of the database. Cohen's kappa coefficients were calculated and yielded a value of $\kappa = .82$. The codings for the ambiguous utterances were resolved by discussion.

6.2.1.3 Results

Table 6.1 below shows the results of coding the utterances. It is telling that autistic children are already making genuine psychological references to desire at MLU I, lending some indirect evidence to the hypothesis that high MLU values are not a precondition to talking about mental states, at least not about the mental state of desire.

The data in Figure 6.1 in page 140 indicate that some of the children with

Mean Length of Utterance Stages					
Child(*)	MLU I	MLU II	MLU III	MLU IV	Total
Brett	–	–	7	171	178
Jack	–	36	15	–	51
Mark	12	1	–	–	13
Rick	9	2	26	14	51
Roger	–	8	68	14	90
Stuart	25	–	–	–	25
Sum	46	47	116	199	408

Table 6.1: Number of GPRD utterances available from each autistic child by MLU stage.

autism (those for which we have data at this stage) are already making a few genuine psychological references to desire (GPRDs) by MLU stages I and II. However, the frequency of GPRDs seems to grow considerably by MLU III and MLU IV. A one-way ANOVA with frequency of talk as dependent variable and MLU Stage as the independent variable was carried out and the effect of the latter variable was significant, $F(3,10)=18.362$, $p < .001$. *Post hoc* analyses using the Bonferroni criterion for significance indicated that the average frequency of talk about the mental state of desire for MLU I ($M=0.85$, $SD=0.44$), MLU II ($M=0.8$, $SD=0.59$) and MLU III ($M=2.28$, $SD=0.69$) were all significantly lower than frequency of talk about GPRDs at MLU IV ($M=4.62$, $SD=1.14$).

To repeat the note of caution provided earlier in the section on the philosophy of statistical tests: Unfortunately, due to missing data points at some of the MLU stages for both the normally developing children and the autistic children, the data had to be analyzed with an independent groups ANOVA. I am aware that this is not the optimal analysis and that I may be overlooking some interesting

Figure 6.1: GPRDs by MLU as a percentage of the total number of utterances for that autistic child at that MLU stage

differences in the data (I, in fact, may be making a Type II error). However, many of the results I have obtained, as I will show below, are indeed significant, minimizing the likelihood that a Type II error occurred. This is, I believe, a necessary step forward and although the statistical analysis may not be the optimal one, I think it gives a useful index of the reliability of the findings.

6.2.2 Self versus other

6.2.2.1 The data

The data come from the 'want' utterances that were coded as genuine talk about desires in the previous section. To recapitulate: there are 408 'want' utterances

that are genuine talk about desires and that belong to 6 children at four different MLU groups.

6.2.2.2 Procedure

The procedure was the same as that used in the analysis of Self versus Other for normally developing children (see section 5.3.3 in page 104).

6.2.2.3 Results

No developmental trend of talk about self versus other was observed. Children with autism are indeed talking about the desires of others at the first stage analyzed (MLU I). Similar to their normally developing counterparts, children with autism seem to talk about their own desires significantly more than about those of others but the overall frequencies do not seem to change significantly as a function of MLU.

A two-way independent groups ANOVA was conducted with self-other and MLU Stage as the independent variables and frequency of use as the dependent variable. A significant main effect was found for self-other, $F(1,20)=69.89$, $p=.000$. Neither MLU Stage or the interaction between the variables were found to be significant, $p > .05$.

Figure 6.2: Frequency of GPRDs to self versus other in children with autism as a function of MLU

6.2.3 Desire and communicative uses

6.2.3.1 The data

The data come from the already identified ‘want’ utterances in the autistic corpus described in Chapter 4 and described in more detail above. In this case, however, I report results from BRs, DRs and IEs compared to GPRDs.

6.2.3.2 Results

Talk using ‘want’ in its communicative function rather than its GPRD function reveals a marked decrease at MLU II. After this decline, the frequency of communicative uses seem to increase again by MLU III and then probably stabilize

at MLU IV.

Figure 6.3: Average frequency of the different communicative uses as a percentage of total utterances as a function of MLU compared to age for autistic children (Error bars are standard errors.)

Figure 6.3 indicate that if we compared communicative uses of 'want' versus the term 'want' used to talk about genuine desires there seems to be practically no difference between communicative uses and genuine talk about desire until at least MLU IV, where there appears to be an increase in frequency of talk about genuine desire rather than other communicative uses. A four (uses of 'want': BRs, DRs, IEs and GPRDs) by four (MLU Stages) independent groups ANOVA was carried out on the frequency data (the dependent variable). Preliminary conclusions drawn from the visual inspection of Figure 6.3 are corroborated by the statistical analyses. An omnibus main effect was found for both MLU Stage $F(3,40)=6.30$, $p=.001$ and Communicative use of 'want' $F(3,40)=14.68$, $p=.000$. There was also an interaction effect $F(9,40)=3.78$, $p=.002$. Pairwise comparison revealed

that although the frequency of talk about genuine desires is not statistically different at the first two MLU stages (MLU I and MLU II), there is in fact a difference at MLU III. At this MLU stage, frequency of talk about genuine desires ($M=2.28$, $SD=0.69$) is significantly more than talk about both behavioral requests ($M=0.67$, $SD=0.43$) and idiomatic expressions ($M=0.43$, $SD=0.49$). By MLU IV, the frequency of talk about genuine desires ($M=4.62$, $SD=1.14$) has increased so as to make this kind of talk much more common than the other uses of ‘want’, BRs ($M=0.47$, $SD=0.56$), DRs ($M=1.23$, $SD=1.02$) and IEs ($M=0.31$, $SD=0.27$).

6.2.4 Desire and their intentional objects

6.2.4.1 The data

The data come from the already identified GPRD utterances in the autistic corpus. To recapitulate: there are 408 GPRD utterances belonging to 6 children at four different MLU groups.

6.2.4.2 Procedure

Every GPRD utterance in the autistic corpus was coded as belonging to one of three possible and mutually exclusive categories: as a desire having a physical object as its desired object (Fodor, 1992), as a desire having an action as its desired object or as a desire having a (different) state of affairs as its desired

object. For more details on how the different objects were recognized see Chapter 4. As in Chapter 4, those few utterances coded as having no objects were excluded from further analysis given that they were so low in frequency that they were subject to a floor effect.

6.2.4.3 Results

As we can see from Figure 6.4, physical objects as a percentage of total utterances stays relatively stable across MLU stages. The only noticeable pattern is the drop in frequency at MLU II, which is also present in the other MLU results from children with autism. It is interesting to notice also that the same pattern is not the case with actions and state of affairs. Actions do not show the decrease in frequency at MLU II but they continue increasing, seemingly after the last MLU stage analyzed. Actions increase in frequency at MLU III and particularly MLU IV. Desires for alternative states of affairs are more evident than for actions. The frequency of talk about states of affairs is extremely low at early MLU groups, but increases markedly at MLU IV.

A two-way independent groups ANOVA was carried out on the data, with MLU Stage and Objects as independent variables and frequency as the dependent variable. Significant effects were found for MLU Stage $F(3, 30)=25.84$, $p=.000$, the different objects in the condition (physical objects, actions and states of affairs) $F(2, 30)=42.11$, $p=.000$ and also for the interaction, $F(6, 30)=16.71$,

Figure 6.4: Average frequency of the different objects as a percentage of total utterances as a function of MLU for autistic children. (Error bars omitted for readability.)

$p=.000$. In order to see the differences between the means in the different conditions by themselves (i.e. whether there were some significant changes in the means of physical objects at the different MLU Stages, for example), several One-Way ANOVAs were carried out on the data, with the different kinds of possible desire objects as the independent variable and frequency of talk about these objects as the dependent variable. Interestingly, there was no main effect for physical objects $p > 0.05$, there is a strong main effect for actions, $F(3, 10)=53.30$, $p=.000$ and a small, marginal main effect for states of affairs, $F(3,10)=3.75$, $p=0.49$. Post-hoc tests using the Bonferroni criterion for significance showed that there is a difference in frequency of talk about desires for actions between MLU I

($M=0.078$, $SD=0.069$) and MLU III ($M=1.40$, $SD=0.43$) and IV ($M=3.51$, $SD=0.10$). There was also a significant increase in frequency talk from MLU II ($M=0.55$, $SD=0.49$) to MLU IV ($M=3.51$, $SD=0.10$) and between MLU III ($M=1.40$, $SD=0.43$) and MLU IV ($M=3.51$, $SD=0.10$). In the case of states of affairs, there seems to be a marginal difference between MLU II ($M=0.01$, $SD=0.02$) and MLU IV ($M=0.58$, $SD=0.56$).

6.3 General conclusions

Results from the sample of autistic children point to a development of the concept of desire at a late MLU stage, probably by MLU III or IV. Interestingly, there is no change whatsoever between MLU I and MLU II, while between MLU III and IV, talk about desire actually increases its frequency. This stands in sharp contrast to the sample of normally developing children. However, more conclusive results will be analyzed in the next chapter when these two samples are formally compared against each other with the same battery of inferential statistics.

When it comes to genuine desire talk about those desires belonging to self versus those belonging to other people, there is indeed a preference for talk about self desires rather than those of others. There does not seem to be an evident developmental trend when these two kinds of talk about desires are plotted against MLU stage. Children with autism do not seem to have problems talking about other people's desires, even though they prefer, pragmatically perhaps, as their

normally developing counterparts, to talk about their own desires most of the time.

When it comes to the different communicative uses of desire, it is evident that children with autism make use of direct repetitions quite regularly. Interestingly, behavioral requests lie at the bottom of the frequency count. There is no satisfactory story to tell about why children with autism do not make such behavioral requests. One explanation may have to do with the fact that they do not see other people as "helpful" in satisfying their desires. It is only at MLU IV that autistic children's begin to talk more about genuine desires. Before then, talk about desire is on a par with other communicative uses.

Lastly, autistic children's talk about the different kinds of desires indicates more talk about actions than any of the other two kinds of objects. In fact, in the end, children with autism are usually talking about desires for actions of the style "I want to run" rather than for physical objects or states of affairs. An explanation of what this might mean is difficult at this stage. This will have to wait until the next chapter, when the findings of the previous and the current chapters are combined in order to find some patterns in the comparison of autistic and normally developing children.

CHAPTER 7

Comparing normally developing and autistic children

7.1 A comparison of children's reference to [desire]_p by means of 'want'

With the data analyzed so far, it is not evident whether autistic children and normally developing children differ developmentally when they come to talk about other people and themselves as being in a mental state of desire. In this chapter, I compare samples of these two populations of children in order to find out whether there are, in point of fact, certain telling differences between them.

If, on the one hand, there are no differences between the groups, then one may be inclined to assume that, if desire talk is necessarily representational, as has been suggested here, children with autism do have representational abilities and their inability to reason by means of beliefs should be explained by alluding to other non-representational mechanisms.

If, on the other hand, some differences exist between the two groups of children: these differences could be of two sorts: a) there could be evidence of delay; that is, children with autism seem to talk about desires only later during their development, rather than never. This I will call the “developmental delay hypothesis”. The other hypothesis is that autistic children never acquire the concept of desire in the same way as normally developing children do, in which case it is not a matter of developmental delay, but of a more pervasive developmental characteristic of this population of children. I will call this hypothesis the “pervasive developmental hypothesis”.

7.2 Comparison 1: GPRDs

Based on a visual inspection of Figure 7.1, autistic children and normally developing children seem to start off more or less the same in terms of rate of talk about genuine desires. At MLU II, however, children with autism continue to make roughly the same number of references to desire as they did at MLU I while normally developing children have increased the frequency of talk about this mental state at MLU II. By MLU III, children with autism and normally developing children have again come to talk about desires at roughly the same rate. It does seem like there is a developmental delay in talk about the mental state of desire for children with autism, but this delay is over by MLU III. It seems frequency of talk about desires from autistic children at MLU IV is not

Figure 7.1: Frequency of talk about genuine desires for both autistic and normally developing children as a percentage of total utterances and as a function of MLU

radically different from their normally developing counterparts. Unfortunately, there are no data for autistic children at MLU V, which would have clarified the issue further.

Overall, the pattern of the data in Figure 7.1 points to a relative delay but not necessarily to a pervasive developmental problem.

In order to test these observations, a 2 (Condition: autistic versus normally developing children) by 4 (MLU Stage: MLU I vs. MLU I vs. MLU II vs. MLU III vs. MLU IV) independent groups ANOVA was conducted on these data, with *Condition* and *MLU Stage* as the independent variables and *Frequency* of talk about genuine desire as the dependent variable. There was no main

effect for *Condition*, $p > .05$. However, there was, not unexpectedly, a main effect for *MLU Stage*, $F(3,45)=11.33$, $p < .001$, and the interaction between the variables (*Condition* x *MLU Stage*) approached significance, $F(3,45)=2.27$, $p = .097$. Clearly, more data would be needed to clarify this interaction. As expected, pairwise comparisons yielded a significant result for MLU II. Children with autism are talking about desires significantly less ($M=.8$, $SD=.60$) than their normally developing counterparts ($M=2.68$, $SD=2.14$).

7.3 Comparison 2: Self versus other

When it comes to contrasting normally developing and autistic children, the difference between ascription of mental state to self and other becomes quite relevant. In fact, it has been assumed that children with autism are not able to *impute* mental states to others. Figure 7.2 below shows the data for both autistic and normally developing children's ascription of desire to both self and other as a percentage of total utterances coded as genuine psychological talk about desire. A note of caution, however: I might have found a difference if I had grouped children in MLU stages that covered a shorter MLU period; that is, casting a finer net. However, as discussed above, the staging was not arbitrary and is supported both in the literature on language acquisition (Brown, 1973) and in the studies on which these studies are based (Bartsch & Wellman, 1995; Tager-Flusberg, 1993). Thus, using the stages proposed by other researchers are

necessary to compare the results obtained here with other studies in the literature.

An inspection of Figure 7.2 suggests no meaningful difference between autistic and normally developing children's talk about desires when it comes to their own desires or those of others. A 2 (Condition: autistic vs. normally developing children) by 4 (MLU Stage) by 2 (self versus other) independent groups ANOVA was carried out on the data. Condition, MLU Stage and self/other talk were the independent variables and frequency of talk about using desire terms was the dependent variable. As predicted from a visual inspection of the data, there were no main effects for either condition or MLU Stage nor for the interaction between them, $p > .05$. Not unexpectedly, however, there was a significant difference in talk about self versus other. In both conditions, autistic and normally developing children use desire words to refer to their own desires more than they do so to refer to those of others, $F(1,86)=409.27$, $p < .001$.

The next comparison examined the differences between the autistic and normally developing groups of children when it comes to their communicative uses of 'want' (uses that need not involve the ascription of a mental state to others). The null hypothesis was that there would be no (developmental) changes in the frequency of occurrence of 'want' qua its communicative role as a function of MLU in each of the conditions (normally developing versus autistic children).

Figure 7.2: Frequency of talk about genuine desires for self versus other for both autistic and normally developing children as a percentage of total utterances and as a function of MLU. Figure A shows a comparison of self between autistic and normally developing children and Figure B shows a comparison of desire terms about others.

7.4 Comparison 3: GPRDs versus communicative uses of ‘want’

As has been discussed extensively in other sections of this dissertation (see §5.3.4 and §6.2.3), comparing all possible uses of ‘want’ has implications for what GPRDs are. If the utterances that have been coded as communicative uses do not show a clear developmental pattern as GPRDs (or the term ‘want’ as used to talk about the mental state of desire) do, then this may lend some empirical support to the hypothesis that the NOS paradigm is actually identifying a developmental trend in desire reference, not merely a change in the use of the word ‘want’ per se, but something that reflects a change in the child’s conceptual repertoire. The child is now able to assign some meaning to the word ‘want’ that it either did not have before or was not useful for some reason.

Figure 7.3 below shows how the frequency of talk of the different communicative uses of want changes as a function of MLU Stage for each of the two groups (autistic versus normally developing children). The leftmost plot clockwise corresponds to the most disputed one when it comes to its status as non-mental ascription or expression of desire. Researchers have actually called into question the hypothesis that these expressions are not inherently talk about mental states. Conservatively and in virtue of more comparative power to other studies in the literature, BRs will be taken not to refer to a mental state, as Bartsch and

Figure 7.3: Frequency of talk of the different communicative uses of 'want' for both autistic and normally developing children as a function of MLU. (Error bars are standard errors)

Wellman did in their own studies. More work is, of course, needed in this area. Both Direct Repetitions (DRs) and Idiomatic Expressions (IE) are taken to be unproblematic as to their status as communicative uses of 'want'.

Figure 7.3 indicates that autistic children start using BRs earlier than their normally developing counterparts, at MLU I. In turn, normally developing children start using the term 'want' qua a behavioral request considerably more than

their autistic counterparts somewhat later, at MLU II. Children with autism again repeat their utterances considerably more than their normally developing counterparts in MLU I, consistent with the finding that children with autism present echolalia at young ages (Baron-Cohen, Tager-Flusberg, & Cohen, 1993; Frith, 1989; Happe, 1994). However, this effect seems to diminish at higher MLUs. Finally, differences in IEs between the two groups are minimal at the different MLU stages.

An omnibus 2 (condition: autistic vs. normally developing children) x 3 (communicative use: BRs, DRs and IEs) x 4 (MLU Stage) independent groups ANOVA was carried out to test for significant differences. Results show only an interaction effect for condition versus MLU Stage, $F(2,135)=4.01$, $p < .01$. Given our previous studies, this result should not be very surprising: the interaction reflects the fact that many a times the autistic children increase their frequency of talk for certain categories while at the same time normal children do the opposite. This is particularly evident at MLU II, when children with autism seem to lower the frequency of talk about the categories studied here, while normal children seem to increase them.

In order to clarify the results of the omnibus ANOVA described above, two (condition: autism vs. normally developing children) by four (MLU Stage) independent groups ANOVA were carried out on each of the three kinds of communicative uses of 'want'. In each ANOVA, frequency of talk about the relevant

communicative use of 'want' (BRs, DRs, IEs) was the dependent variable.

For BRs, there were no main effects for either Condition or MLU Stage, $p > .05$; however, there was an interaction between these variables, $F(3,45)=4.98$, $p = .005$. For the purposes of the present investigation, it is important to notice that the interaction effect is brought about by significant differences between the normally developing and autistic groups at different MLU Stages. As previously hypothesized, there were significant differences between the autistic ($M=0.94$, $SD=0.68$) and normally developing children ($M=0.09$, $SD=0.22$) children at MLU I and for the autistic children ($M=0.22$, $SD=0.19$) and the normally developing children ($M=0.81$, $SD=0.70$) at MLU II. A Bonferroni adjustment for multiple comparisons was used. For DRs a main effect for condition was obtained, $F(1,45)=5.83$, $p = .020$. This is due, mostly, to a significant difference between autistic ($M=1.61$, $SD=1.30$) and normally developing children ($M=0.1$, $SD=0.15$) at MLU I. Finally, the IE analysis indicated a main effect for MLU Stage $F(3,45)=3.55$, $p = .022$, no main effect for the interaction $p > 0.1$, with *Condition* approaching significance, $p = .064$. This latter effect can be accounted for by a small significant difference between autistic ($M=0.43$, $SD=0.49$) and normally developing children ($M=0.11$, $SD=0.20$) at MLU III, as obtained using pairwise comparisons.

7.5 Comparison 4: the objects of GPRDs

The last in these series of analyses comparing autistic versus normally developing children's talk about the mental state of desire deals with the frequency of talk about the different objects of desire as a function of MLU. To recapitulate: desires for physical objects (e.g., "I/He want(s) a cookie") are considered ambiguous as to their status as talk about the child's own or someone else's desire; while, as has been discussed, talk about desires for actions or states of affairs is considered as talk about the mental state of desire. According to several authors (either explicitly, see Tan & Harris, 1991; Harris, 1996 or implicitly in their treatment of the representational status of desire, see Bartsch & Wellman, 1995), communication of desires for actions or states of affairs should be more "representational" than talk (communication) of desire for physical objects. There is still controversy about this issue, but no researcher so far has looked at the differences in the objects of desire when they are coded as GPRDs, least of all a comparison between a group that is hypothesized to have representational deficits and one that is not.

A 2 (condition: autistic versus normally developing children) by 4 (MLU stage) by 3 (GPRD objects: physical objects, actions and state of affairs) independent groups ANOVA was carried out on the frequency of talk about these different objects. This omnibus ANOVA reported significant interactions for all comparisons: normally developing children versus autistic by MLU, $F(3,135)=3.93$,

Figure 7.4: Frequency of talk of the different objects of desire for both autistic and normally developing children as a function of MLU

$p = .01$, normally developing children versus autistic by object of desire, $F(2,135)=8.01$, $p = .001$, MLU stage by the object of desire utterances $F(6,135)=9.34$, $p = .001$ and finally the three way interaction of condition by MLU stage by object, $F(6,135)=4.14$, $p = .001$.

In order to explore these interactions, as in the previous section, a two (condition: autistic vs. normally developing children) by four (MLU Stage) independent groups ANOVA was carried out on each of the three kinds of identified desire objects. In all these ANOVAs, frequency of talk about the relevant object

type (physical objects, actions and states of affairs) was the dependent variable and pairwise comparisons were carried out using the Bonferroni criterion at the $p < .05$ level.

For physical objects, there was a main effect for Condition, $F(1,45)=4.95$, $p = .031$. There was no main effect for MLU Stage nor was there an interaction between Condition and MLU Stage. However, the interaction approached significance $F(3,45)=2.69$, $p = .057$ and, given the restrictive nature of independent groups ANOVA, the possibility of a Type II error cannot be discounted. Pairwise comparisons yielded a significant difference between the autistic ($M=0.22$, $SD=0.19$) and normally developing children ($M=1.25$, $SD=0.66$) at MLU II. For desires for actions, there were significant main effects for each of the variables and a significant interaction. There was a main effect for Condition, $F(1,45)=6.13$, $p = .017$, , a main effect for MLU Stage, $F(3,45)=21.79$, $p = .000$ and an interaction effect for Condition by MLU Stage, $F(3,45)=4.98$, $p = .005$. Of interest here is the comparison between the autistic and normally developing groups. Pairwise comparisons yielded a significant difference between autistic ($M=3.51$, $SD=0.10$) and normally developing children ($M=1.50$, $SD=0.78$) at MLU IV. Finally, for SOAs (state of affairs), there was a significant main effect for MLU Stage, $F(3,45)=6.97$, $p = .001$ and an interaction effect, $F(3,45)=5.57$, $p = .002$. Pairwise comparisons only yielded a difference between autistic ($M=0.14$, $SD=0.32$) and normally developing children ($M=0.10$, $SD=0.16$) at MLU IV.

7.6 Discussion and conclusions

The resulting picture is, expectedly, rather complex in nature. The following discussion should clarify some of the most important points derived from the analyses above.

Comparison 1 - GPRDs in autistic and normally developing children:

Overall, no difference in frequency of GPRDs was observed between autistic and normally developing children. GPRDs generally increased as MLU increased. No systematic relationship between the two groups and MLU in terms of GPRD frequency was observed, with the exception of MLU II. These results lend some support to the hypothesis that the normally developing and autistic groups do not differ dramatically from each other at different MLU Stages. However, at MLU II, children in the autistic group talked about the mental state of desire significantly less than their normally developing counterparts. At later MLU Stages, though, autistic children appear to have caught up with the normally developing group.

Comparison 2 - GPRDs for self and other in autistic and normally developing children: Overall, no difference in frequency of GPRDs was observed between autistic and normally developing children. Talk about the child's own desires versus those of others did not increase significantly as MLU increased. No systematic relationship between the two groups and MLU in terms of frequency of Self/Other reference was observed. The lack of differences between the nor-

mally developing and autistic groups is an important one for TToM research. This lends support to the claim that children with autism do not have a problem in (explicitly) ascribing a mental state to others even if this mental state includes a proposition that is not the case, as with desire. In other words, desires seem to involve a proposition that is in fact false. When the child says that he or she wants a red car, the proposition "I have that red car" is obviously false. Not only children with autism refer to these false propositions when referring to their own desires, but they also refer to others in the same way. This leads us to think that children with autism are in fact able to ascribe a mental state to other people such that the proposition that that attitude is about is a counterfactual state of affairs, much as beliefs are.

To beat a dead horse: The finding that children with autism can ascribe desires to other people stands in sharp contrast with many of the presuppositions of TToM research; namely, that children with autism cannot entertain a proposition that stands for a counterfactual state of affairs, as is the case with false beliefs (Wimmer & Perner, 1983; Wellman et al., 2001). To clarify: similarities between autistic and normally developing children when ascribing desires to self and other means that the former group can hold thoughts that involves a representation of an epistemic relation between an agent and a proposition and that the proposition is in fact false. This bears some resemblance to the thoughts that children should be able to have about beliefs when they are around 4 years of age.

Comparison 3 - communicative uses of 'want': Overall, the ANOVA showed a difference in frequency of the different communicative uses of 'want' between autistic and normally developing children. Communicative uses generally increased as MLU increased. A systematic relationship between the two groups and MLU in terms of the frequency of communicative uses was observed. The finding that autistic and normally developing children differ from each other as a function of MLU only in terms of behavioral requests suggest that this use of 'want' is indeed somewhat different from the other two communicative uses of 'want'. Several researchers, among them particularly Tan and Harris (1991) and Harris (1996), have suggested that behavioral requests are in fact a form of talk about the mental state of desire. Although the studies reported here lend some support to this idea, it is hard to draw conclusions about what the nature of the difference might be. One might assume that behavioral requests are themselves a proto-form of fully-fledged desire talk; indeed, when Bartsch and Wellman's (1995) proposed a non-mental desire category, they were talking about behavioral requests rather than genuine talk about the mental state of desire (or GPRDs),

Another interesting and well-known phenomenon is that children with autism repeat the adults' utterances more than normally developing children. Indeed, children with autism present highly noticeable echolalia at younger ages, as previously discussed (Happe, 1994). This is quite evident at MLU I, when there is

a significant difference between the normally developing and autistic groups in terms of direct repetitions (DRs).

Comparison 4 - the objects of desire: Interestingly, both normally developing children and autistic children start referring more to physical objects than to actions or state of affairs, but they do so indistinctly. That is to say, there are no differences between these two groups when it comes to using the word 'want' to talk about any of these different objects. This is in fact puzzling in the sense that we are sure that actions and state of affairs are arguably desires that involve talking about an nonexistent object, an object that they have to somehow imagine in order to talk about it. It is also interesting to notice that children with autism have a sort of decline in the frequency of talk using the desire term 'want' at MLU II for physical objects, but not for any other kind of desire objects. In fact, when it comes to actions and states of affairs, children with autism end up outperforming their normally developing counterparts. The same does not happen for physical objects, a category in which normally developing children outperform autistic children, with a very different developmental trend (notice that until MLU III, frequency of talk using the term 'want' to talk about desires for actions and state of affairs for normally developing children and autistic children is similar. However, there is a frequency increase from MLU III to MLU IV for autistic children). It seems that talk using 'want' to refer to a desire for a physical object is a mirror image for autistic and normally devel-

oping children. This effect is quite peculiar given that children with autism are in fact talking with 'want' about actions and states of affairs, presumably more "abstract" objects, if you will.

CHAPTER 8

Conclusions, speculations and future work

8.1 H1: Making sense of all the results

H1 is the hypothesis that children make genuine psychological references to the mental state of desire before they do so to the mental state of belief. This hypothesis has been put forward by many researchers in a number of different ways (Harris, 1996). One of the main objectives of this dissertation was to put the hypothesis that children have the concept of desire (even for actions and states of affairs) to a number of empirical tests by using one of the methodologies available for TToM research; namely, the Naturally Occurring Speech paradigm. This methodology involves analyzing a large database containing children's utterances at different stages during their development and searching for certain words that are supposed to stand as proxies for the concept of $[\text{desire}]_p$.

In order to set the basis for later studies, the first logical step was to replicate Bartsch and Wellman (1995)'s studies. This was done in Chapter 5. The concept of $[\text{desire}]_p$ seems to be present at the youngest ages analyzed, consistent

with Bartsch and Wellman's results. This is quite early developmentally in TToM terms. In fact, not only was the early presence of the concept of desire replicated, but also the general developmental picture. As with Bartsch and Wellman, although children refer to desires quite early, there is a peak of genuine desire talk at around the 33rd month of life and a slight decrease in frequency thereafter. Even the percentages are roughly the same in both studies. Thus, at least when it comes to this replication, B&W's results seem to be quite robust, considering that this replication consisted of 12 different children and two different coders who had not been involved in the original studies. To summarize, B&W's study about desires seems to be extremely replicable, which I believe says quite a bit about the reliability of the methodology, if not yet about how the results may or may not be an artifact of the instrument. To this discussion I now turn.

It could be the case that development of the concept of $[\text{desire}]_p$ can be given a linguistic explanation. The data are inconclusive in this respect. We see the same kind of development in MLU (which is a good measure of linguistic (read syntactic) ability) as we see with age. However, this does not really answer the question of whether age or MLU is a better predictor of desire reference. The data are equivocal about which is the better predictor of the frequency of utterances with the concept of $[\text{desire}]_p$ embedded in the proposition triggering the linguistic expression. What is relevant, I think, is that even at MLU I children are actually producing talk with the concept of $[\text{desire}]_p$ embedded in the proposition "giving

birth” to the linguistic expression. Thus, as Bartsch and Wellman (1994, 1995) would put it, it is more likely that MLU is related to the frequency with which children talk about desires, not really whether they have already acquired the concept or not. Children seem to be able to refer to desires even at the lowest MLU stages. Figure 6.1 provides some evidence for this last claim. Children with autism do not have the peak that non-autistic children present until MLU IV. MLU stages I, II and III are not statistically significant from one another. What is interesting is that there seems to be no change between MLU I and II, more evidently so when compared to the frequencies of their non-autistic counterparts. I take this to mean that even if they have the same linguistic abilities as their non-autistic counterparts, children with autism take a bit longer in developing the concept of [desire]_p. Figure 7.1 makes this difference more evident.

Significantly, both for age and MLU stage, utterances coded as GPRDs increased in frequency, while those that were coded as non-GPRD did not do the same. Actually, neither group of children had any significant differences relative to the total number of utterances at the different stages (MLU nor age) when they used ‘want’ in a merely communicative sense. This pattern of results suggests that the NOS paradigm is really picking out a qualitative difference between GPRDs and non-GPRDs utterances. If this is the case, then we have children using desire talk at the earliest age analyzed and the instrument seems to identify real desire talk. In short, the evidence supports H1. Thus, the lag between belief

explain behavior in terms of desires. In other words, for Loreto to explain her own behavior (or that of others) she must entertain a belief about the organism's (maybe her own organism's) desire state (Davies & Stone, 1995c). To explain or predict behaviors and actions, it is not enough that we are able to be in desire states ourselves (unlike, for example, what Simulation-Theory of Mind says, see Gordon, (1995a)). What is a *conditio sine qua non* is that we are able to entertain beliefs about an organism's desire states (Dennett, 1987; Sperber, 1999). For example, one way to explain why Loreto is buying a ticket to Chile this morning is to entertain a belief with the embedded proposition in [3] above.

Here is an example of the discussion above: that Loreto wants to be in Chile doesn't necessarily mean that she has articulated that desire herself. She may simply be in a desire state (an internal state such that she wants to be in Chile), without having referred to it or even acted to bring it about. Now, the minute I (or indeed anyone, Loreto included) ascribe(s) that desire to Loreto in order to explain her behavior, I am not (indeed no one is, Loreto included) in just a desire state. There are now also belief states about what Loreto wants. I am now (or indeed Loreto is) in a belief state such that the proposition "Loreto wants to be in Chile" is true of Loreto (herself).

This is also true of $[\text{belief}]_p$. In TToM, to entertain a belief that someone else believes that it's raining, I need two kinds of beliefs, a belief state that is about a $[\text{belief}]_p$ and the $[\text{belief}]_p$. We are indeed always in one belief state or another,

but what is at issue in TToM is that we are not always able to be in a belief state *about* a belief. Likewise, what I want to say here is that when ascribing a desire to someone I need to be in a belief state such that that belief state has a desire concept embedded in it.

This is what I believe is happening in the analyses carried out so far. Notice that children by the third age group in the analyses presented here are not only in a desire state but they are in a belief state *about* a desire; i.e. using $[\text{desire}]_p$. These children are referring to those states by means of the word 'want' (when utterances are real GPRDs, obviously). Together with the fact that the children analyzed here talk about non-existent objects and that they do not really change the frequency of talk about their own and other people's desires as a function of age, this makes me believe that, by AGE3 (29–32 months) these children do have the ability to refer to and reason by using the mental states of desire. In yet other words, when children utter sentences with the term 'want' in its GPRD sense, they are not only in a desire state (the first, trivial interpretation discussed and exemplified in the previous two paragraphs), but they are in a belief state that has, as part of the belief proposition, an instance of $[\text{desire}]_p$. The non-trivial interpretation.

A puzzling implication about entertaining thoughts and beliefs including $[\text{desire}]_p$: the results of my analyses showed that by the 33rd month of life, children are referring to desires for non-existent objects (actions and state of affairs). If so,

these children are referring to a desire for a counterfactual state of affairs. That Abigail wants to be ten years old or that Benjamin wants a red car means that Abigail is not ten years old and that Benjamin got a blue, rather than a red car. Thus, in order to entertain beliefs about these kinds of desires, the desirer needs to believe that the state of affairs referred to by the proposition he or she is holding the attitude (desire) towards does not obtain. It would be paradoxical or irrational in Dennettian terms to think of it otherwise; namely, that Loreto wants to be in Chile when she believes she is in Chile, that Abigail wants to be ten years old when she believes she is ten years old or that Benjamin wants a red car when he believes that he has the same red car that he wants.

There is the case that we sometimes express a desire when it is in fact the case that P (the proposition that our desire concept is an attitude of), but this is quite peculiar: it may be indeed just a plain communicative act (I am so happy that P that I say "I want P " because I want to stress the fact that I'm happy about it) or it may be the case that we know that Loreto is in Chile, wants to be in Chile and she believes herself not to be there, though in fact she is. But this is a special case that would mean analyzing a lot of other beliefs: the belief that she may be asked to leave Chile or that she will just leave Chile because she has stuff to do in Canada, etc. In any case, I am analyzing the canonical case and I yearn for simplicity in this issue, so these particular cases will be left for further study elsewhere. However, in many important respect, when using

[desire]_p, children are entertaining a belief about someone's desire (maybe their own) such that the embedded proposition of that desire is believed by the desirer to stand for a counterfactual state of affairs.

If the above discussion is somewhat on the right track then even at younger ages, every time children refer to their own or other people's desires, they should be in a belief state such that it involves the use of [desire]_p. It would be hard to characterize the thoughts that the child is entertaining when explaining or reporting behaviors by means of desires merely as "subjective connections" to objects (connections to those objects and not others, particularly actions and state of affairs). In other words, some general connection (such as a subjective connection) is not enough to entertain GPRDs. There has to be a fully-fledged intentional connection (in the philosopher's sense of being *about* something) and thus that something has to be represented in the mind. Suppose, for the sake of argument, that children do grasp desires as a "subjective connection" between the organism they are trying to explain the behavior of and the object that this organism desires. This could be relatively easy to see for desire ascription to organisms other than self. However, it would be hard to believe that when talking about their own desires and explaining their own behaviors by means of desires ("because I wanted to go to the park"), children think of *themselves* as just holding a subjective connection to a state of affairs that does not hold. It is in fact very hard to believe that when reasoning about their own behaviors by

means of desires, children are not representing themselves as wanting something in particular; to be in the park, for example.

One implication of this is that anything children say (with the exception, of course, of fixed expressions, direct repetitions and some seemingly unanalyzed utterances that should be taken as mere expressions of desire, what B&W would call behavioral requests) about their own (or indeed whoever's) desires is right away an ascription. If a child says: "I want to be ten years old" or she says "I want an apple", she is, in my framework, already ascribing a desire (to herself). The mere reference to a genuine desire (even its own) demands of the organism that it be using [desire]_p and be in a belief relation to that concept.

As said before, several of the analyses reported in these pages have a bearing on this issue. The most relevant, I believe, is the lack of difference in the relative frequency of desire ascription for self versus other; namely, children's reference to other people's desires (rather than reference to their own) at the different stages of their TToM development is quite constant, suggesting that children can in fact ascribe desire to others from the earliest ages (i.e. children can ascribe genuine desires, desires that involve all the nuances discussed in the above paragraphs, including entertaining a proposition that the desirer believes stands for a counterfactual state of affairs). In other words, there seems to be no significant change in frequency of reference to other people's desires either at the different ages or at the different MLU stages analyzed.

Another relevant piece is that children, even at the youngest ages and lowest MLU stages analyzed, are able to talk about nonexistent objects. To represent desire for propositions that are not the case, $\neg P$, children have to be first in a mental state such that there is a representation of the desired state of affairs or action. These propositions are different from desires for physical objects in that it is completely unambiguous in that children should be somehow mentally representing them, i.e. that it is $\neg P$ in the context of the desirer. That age and MLU stage do not have a bearing in the developmental curve for these two special objects means that very young children are able to be in mental states such that the object of the state has to be represented. But if they also utter 'want' utterances with these special objects embedded in them, it means that they can entertain the concept of desire and ascribe (to themselves, at least) the representation (by using the $[\text{desire}]_p$) of whatever action or state of affairs they want. They are entertaining a representation of a representation. This effect is evident both for age and for MLU state as the dependent variables. But if so, these mental state ascriptions demand this child to entertain thoughts that may include non-existent objects and a proposition standing for a counterfactual state of affairs. Curiously again, this has all the complexity of the concept of belief, but at a much earlier age (in a TToM development scale). What should we now take to be the difference between these two concepts such that $[\text{belief}]_p$ seems to be acquired later than $[\text{desire}]_p$?

The NOS paradigm shows that in normally developing children, belief ascription starts at around four years of age (around a year later than desire ascription). The NOS paradigm also provides some evidence that children with autism never really get to the stage of being able to ascribe beliefs (to other people at least). But if children had had an inability to form meta-representations, the analyses presented in this investigation would have made it evident. Even the fact that desire ascription to others seems to appear reliably (the child's third reference) two to three months after expression of their own desires shows that these ascriptions are in place almost a year and a half before the concept of belief (Wellman et al., 2001). The claim, then, is that there must be something else, other than meta-representation, that is, that accounts for the difference between children's ability to ascribe desires and their ability to ascribe beliefs.

The child being able to form belief meta-representations need not be in contradiction with the evidence coming from the literature on belief ascription. A hypothesis worth investigating could be that children's understanding of $[\text{belief}]_p$ depends on their abilities to calculate possible worlds, so to speak, i.e. differentially calculating the truth values of the propositions that the different propositional attitudes (belief and desire, in this particular case) are an attitude of. In summary, children can be in beliefs states about desires whose propositions stand for counterfactual states of affairs, but they still cannot be in beliefs states about beliefs. This hypothesis actually begs the question of whether $[\text{belief}]_p$ is there

even if children cannot use it because they do not have enough computational resources to use it.

The hypothesis that children with autism lack metarepresentational abilities of some sort has been quite favored in the literature even though another account of the lag has been put forward by Fodor (1992). Although quite a few empirical studies have been carried out to back up the idea of a metarepresentational deficit, all the evidence supports is that there is a difference between autistic and non-autistic children when it comes to performing in false-belief tasks such as the Sally-Anne experiment, also called a visible displacement task (Wimmer & Perner, 1983). Autistic children may indeed have a problem with false-belief, but that does not necessarily imply that they have a metarepresentational deficit, as is assumed in the literature. If the hypothesis about desire ascription I have proposed in these pages is right, then even children with autism can ascribe the mental state of desire to other people in roughly the same ways as their non-autistic counterparts and so can meta-represent. At least that is what the evidence presented in this work tells us. Moreover, by MLU III, children with autism seem to catch up (after a brief plateau in their MLU II stage) with their non-autistic counterparts in both GPRDs and desires ascriptions of actions and states of affairs, rather than physical objects. This evidence points to the presence of an ability to meta-represent, as discussed in the paragraphs above. What may be happening is that children with autism can't predict behavior by

means of those beliefs because of a deficit in computational mechanisms that work on mental state (meta)-representations. Research on the nature of these computational mechanisms will be left for future research.

8.2.1 Problems, future work and last thoughts

The present investigation is not without problems. The objective of this section is to lay out some of these problems and to discuss possible solutions to them. These problems are mainly of two sorts: first, methodological, the issue of insufficient data, for example, and second, what could be called a semantic problem, the issue of expression versus reference to desire.

8.2.1.1 Data, analysis and controlled experiments

Evidence presented in these pages suggests that the NOS paradigm can capture relatively subtle phenomena in children's language. However, although the results that this methodology can reveal are quite reliable, as we have seen, the use of the NOS paradigm should be restricted to a context of discovery (such as the present one), rather than to context of justification. In other words, given preliminary results obtained using the NOS paradigm, further work should utilize controlled experiments. Carrying out laboratory experiments would have the advantage of dealing with missing data and yielding, in turn, statistical analyses less prone to the probability of various types of statistical errors associated with the present

analyses. This notwithstanding, the use of methods such as the NOS paradigm has the important advantage of “naturalness”, of being able to actually catch important phenomena in the wild. However, the NOS paradigm seems to be more suited to a context of discovery rather than a context of justification. In order to reach this context of justification, the same phenomena the NOS paradigm is making evident should be tested under more constrained conditions in the laboratory but as a complement, not as a substitute for it.

In this vein, one useful experiment that is a logical extension of the present work would be eliciting children’s reference to someone else’s desire. The fact that children do not increase or decrease their reference to other people’s desire across different conditions is an important result. Most of the stronger claims made in this dissertation depend on this ability. Thus, one focus of future research in the area would be to evaluate this ability. Though there are exceptions (Repacholi & Gopnik, 1997), most of the laboratory experiments in the literature on desire are about the child participant’s own desire, rather than those of somebody else. It would be of much interest to test the child on how well they can entertain a desire of someone else for an object that is non-existent (an action, or a state of affairs). Given the characteristics of the NOS paradigm, the data reported in this dissertation could not be used to shed light on this issue. There are simply not enough data to cross-study objects by age by MLU and by self versus other in a meaningful way, without being subject to floor effects. If children can refer to other

people's desires for actions and state of affairs reliably, then that would mean that they can definitely be in a belief state such that there is a mental concept involved ($[\text{desire}]_p$), that the mental concept is ascribed to someone else and that the proposition inside the mental concept stands for a counterfactual state of affairs. If more evidence appears concerning this hypothesis, then it would be hard to maintain that belief and desire ascriptions differ in their metarepresentational status.

Language is another important factor in the analyses reported here and thus one of the main source of problems of the NOS paradigm. It is not clear up to what extent language plays a role in helping children acquire the different concepts of their folk psychology. I have tried to provide some evidence that MLU does not really play a big role in ascribing desires to self versus other for example in both normally developing children and autistic children. However, laboratory experiments should try to rely on other communicative behaviors, such as pointing or language-unrelated methodologies, such as non-nutritive sucking in infants order to test the understanding of desire (mostly other people's) by very young children. By means of these methodologies, we should be able to find out whether using $[\text{desire}]_p$ depends on the child's ability to use language, to talk about it or not. The relationship between the acquisition of TToM concepts and language development is an extremely important issue and should be investigated further.

8.2.1.2 Reference versus expression of desire

Let me come back to an issue I raised right at the very beginning of this dissertation. The problem of whether or not the NOS paradigm, as understood in these pages, is not identifying real *references* to desire. Of course GPRDs do that. The issue is identifying when a child has and when a child has not *grasped the concept* of desire. The careful reader will have noticed that there is concern about a possible distinction between reference and expression of desires.

The difference between these two categories is difficult to capture propositionally, but maybe a couple of examples will help making the point. Take the following (pure) expressions of desire: “hmm, that looks really good” or “I could kill for a new desktop PC” and compare them to ascriptions of desire of the following kind “I have seldom felt a stronger desire to eat something” or “I want to finish this dissertation so badly”. Now, all of these are expressions or ascriptions of our own desires, but since we have to meta-represent others’ desires to deal with them at all, yet can represent our own desires without meta-representing them, then we cannot *just* express someone else’s desire, we must ascribe it. Thus, expression of desire (which, as discussed, can be only about the child’s own desire, given that there could be no expression of other people’s desires) may not necessarily be saying much about the child’s grasp of the concept of desire, while ascription is different in that it can¹. To be sure, this is a knotty issue that should

¹I owe this distinction to Andrew Brook (personal communication)

be addressed in future research. One way to test this issue is by carrying out experiments eliciting reference to other people's desires and expressions of desire in controlled laboratory conditions. I believe separating BRs from GPRDs both analytically and empirically should be the object of future research in this field. The results of this investigations, I believe, could lead to a more refined set of categories.

Even with these problems, I think the analyses reported in these pages clarify some of the concerns raised about the NOS methodology and lead to new hypotheses about the concept of desire, mainly about its metarepresentational status in children's minds.

REFERENCES

- American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders-dsm iiii*. Washington, DC: American Psychiatric Association.
- Astington, J. (2001). The future of theory-of-mind research: Understanding motivational states, the role of language, and real-world consequences. Commentary on "Meta-analysis of theory-of-mind development: The truth about false belief.". *Child Development*, 72(3), 685-687.
- Astington, J., & Gopnik, A. (1991). Developing understanding of desire and intention. In A. Whiten (Ed.), *Natural theories of mind* (p. 39-50). Oxford, UK: Blackwell.
- Astington, J., Harris, P., & Olson, D. (Eds.). (1988). *Developing theories of mind*. Cambridge: Cambridge University Press.
- Bakeman, R., & Gottman, J. (1997). *Observing interaction (2nd. edition)*. New York, NY: Cambridge University Press.
- 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. Cohen (Eds.), *Understanding other minds: Perspectives from autism*. Oxford: Oxford University Press.

- Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind*. Cambridge, MA: MIT Press.
- Baron-Cohen, S., Leslie, A., & Frith, U. (1985). Does the autistic child have a theory of mind? *Journal of Child Psychiatry*, 30, 285-297.
- Baron-Cohen, S., Tager-Flusberg, H., & Cohen, D. (Eds.). (1993). *Understanding other minds: Perspectives from autism*. Oxford: Oxford University Press.
- Bartsch, K., & Wellman, H. (1995). *Children talk about the mind*. New York: Oxford University Press.
- Blackburn, S. (1995). Theory, observation and drama. In M. Davies & T. Stone (Eds.), *Folk psychology* (p. 274-290). Oxford: Blackwell.
- Brentano, F. (1874). *Psychology from an empirical standpoint*. New York, NY: Routledge/Kegan Paul.
- Bretherton, I., & Beeghly, M. (1982). Talking about internal states: The acquisition of an explicit theory of mind. *Developmental Psychology*, 18(6), 906-921.
- Bretherton, I., McNew, S., & Beeghly-Smith, M. (1981). *Infant social cognition*. Hillsdale, NJ: Earlbaum.
- Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.

- Bryson, S., Clark, B., & Smith, I. (1988). First report of a canadian epidemiological study of autistic syndrome. *Journal of Child Psychology and Psychiatry*, 29, 433-445.
- Carruthers, P., & Smith, P. (Eds.). (1996). *Theories of theories of mind*. Cambridge: Cambridge University Press.
- Cassidy, K. (1998). Preschoolers use of desires to solve theory of mind problems in a pretense context. *Developmental Psychology*, 34(3), 503-511.
- Chomsky, N. (Ed.). (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Davies, M., & Stone, T. (1995a). Introduction. In M. Davies & T. Stone (Eds.), *Mental simulation*. Oxford: Blackwell.
- Davies, M., & Stone, T. (1995b). Introduction. In M. Davies & T. Stone (Eds.), *Folk psychology*. Oxford: Blackwell.
- Davies, M., & Stone, T. (Eds.). (1995c). *Mental simulation*. Oxford: Blackwell.
- Dennett, D. (1978a). Beliefs about beliefs:. *Behavioral and Brain Sciences*, 1(4), 568-570.
- Dennett, D. (1978b). *Brainstorms: Philosophical essays on mind and psychology*. Montgomery, VT.: Bradford Books.
- Dennett, D. (1987). *The intentional stance*. Cambridge, MA.: MIT Press.

- Flavell, J., Flavell, E., Green, F., & Moses, L. (1990). Young children's understanding of fact beliefs versus value beliefs. *Child Development*, 61, 915-928.
- Fodor, J. (1975). *The language of thought*. New York, NY: Crowell.
- Fodor, J. (1987). *Psychosemantics*. Cambridge, MA: MIT Press.
- Fodor, J. (1990). *A theory of content*. Cambridge, MA: MIT Press.
- Fodor, J. (1991). Fodor's guide to mental representation: The intelligent auntie's vade-mecum. In J. Greenwood (Ed.), *The future of folk psychology*. Cambridge, MA: Cambridge University Press.
- Fodor, J. (1992). A theory of the child's theory of mind. *Cognition*, 44, 283-296.
- Frith, U. (1989). *Autism: Explaining the enigma*. Oxford: Blackwell.
- Goldman, A. (1989). Interpretation psychologized. *Mind & Language*, 4, 161-185.
- Goldman, A. (1993). The psychology of folk psychology. *Behavioral and Brain Sciences*, 16, 15-28.
- Goldman, A. (1995). In defense of the simulation theory. In M. Davies & T. Stone (Eds.), *Folk psychology*. Oxford: Blackwell.
- Gordon, R. (1995a). Folk psychology as simulation. In M. Davies & T. Stone (Eds.), *Folk psychology*. Oxford: Blackwell.

- Gordon, R. (1995b). Simulation without introspection or inference from me to you. In M. Davies & T. Stone (Eds.), *Mental simulation*. Oxford: Blackwell.
- Happe, F. (Ed.). (1994). *Autism: An introduction to psychological theory*. Cambridge, MA: Harvard University Press.
- Harris, P. (1995a). From simulation to folk psychology: The case for development. In M. Davies & T. Stone (Eds.), *Folk psychology*. Oxford: Blackwell.
- Harris, P. (1995b). Imagining and pretending. In M. Davies & T. Stone (Eds.), *Mental simulation*. Oxford: Blackwell.
- Harris, P. (1996). Desires, beliefs and language. In P. Carruthers & P. Smith (Eds.), *Theories of theories of mind*. Cambridge: Cambridge University Press.
- Hickling, A., & Wellman, H. (2001). The emergence of children's causal explanations and theories: Evidence from everyday conversation. *Developmental Psychology*, 37(5).
- Leiter, A. (1974). International performance scale. In O. Buros (Ed.), *Tests in print ii. an index to tests, test reviews and the literature on specific tests*. Highland Park, NJ: Gryphone.
- Lotter, V. (1966). Epidemiology of autistic conditions in young children: I. prevalence. *Social Psychiatry*, 1, 124-137.

- MacWhinney, B. (2000). *The chldes project : Tools for analyzing talk* (3rd ed.). Hillsdale, N.J.: L. Erlbaum.
- Miller, J., & Chapman, R. (1985). Systematic analysis of language transcripts: User's guide. *University of Wisconsin Language Analysis Laboratory*.
- Morton, A. (1980). *Frames of mind: Constraints on the commonsense conception of the mental*. Oxford: Clarendon Press.
- Perner, J. (1991a). On representing that: The asymmetry between belief and intention in children's theory of mind. In D. Frye & C. Moore (Eds.), *Children's theories of mind* (p. 139-155). Hillsdale, NJ: Erlbaum.
- Perner, J. (1991b). *Understanding the representational mind*. Cambridge, MA: Bradford Books/ MIT-Press.
- Perner, J., & Howes, D. (1995). "He thinks he knows": And more developmental evidence against the simulation (role-taking) theory. In M. Davies & T. Stone (Eds.), *Folk psychology*. Oxford: Blackwell.
- Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *The Behavioral and Brain Sciences*, 4, 515-526.
- Repacholi, B., & Gopnik, A. (1997). Early reasoning about desires: Evidence from 14- and 18-month-olds. *Developmental Psychology*, 33(1), 12-21.

- Rieffe, C., & Terwogt, M. (2000). Deaf children's understanding of emotions: Desires take precedence. *Journal of Child Psychology and Psychiatry*, 41(5), 601-608.
- Rudder-Baker, L. (1999). Folk psychology. In R. Wilson & F. Keil (Eds.), *The MIT encyclopedia of the cognitive sciences* (p. 319-320). Cambridge, MA: MIT Press.
- Rutter, M. (1978). Diagnosis and definition of childhood autism. *Journal of Autism and Childhood Schizophrenia*, 8, 139-161.
- Searle, J. (1983). *Intentionality*. Cambridge, NY: Cambridge University Press.
- Shatz, M., Wellman, H., & Silber, S. (1983). The acquisition of mental verbs: A systematic investigation of the first reference to mental state. *Cognition*, 14, 301-321.
- Sperber, D. (1999). Metarepresentation. In R. A. Wilson & F. C. Keil (Eds.), *The MIT Encyclopedia of the Cognitive Sciences*. Cambridge, MA: MIT Press: Bradford Books.
- Stich, S. (1985). *Folk psychology: The case against belief*. Cambridge, MA: MIT.
- Stich, S., & Nichols, S. (1995a). Folk psychology: Simulation or tacit theory? In M. Davies & T. Stone (Eds.), *Folk psychology*. Oxford: Blackwell.

- Stitch, S., & Nichols, S. (1995b). Second thoughts on simulation. In M. Davies & T. Stone (Eds.), *Mental simulation*. Oxford: Blackwell.
- Tager-Flusberg, H. (1988). On the nature of a language acquisition disorder: The example of autism. In F. Kessel (Ed.), *The development of language and language researchers: Essays in honor of Roger Brown*. Hillsdale, NJ: Lawrence Erlbaum.
- Tager-Flusberg, H. (1992). Autistic children's talk about psychological states: Deficits in the early acquisition of a theory of mind. *Child Development*, 63, 161-172.
- Tager-Flusberg, H. (1993). What language reveals about the understanding of minds in children with autism. In S. Baron-Cohen, H. Tager-Flusberg, & D. Cohen (Eds.), *Understanding other minds: Perspectives from autism*. Oxford: Oxford University Press.
- Tager-Flusberg, H., Calkins, S., Nolin, T., Baumberger, T., Anderson, M., & Chadwick-Dias, A. (1990). A longitudinal study of language acquisition in autistic and down syndrome children. *Journal of Autism and Developmental Disorders*, 20(1), 1-21.
- Tan, J., & Harris, P. (1991). Autistic children understand seeing and wanting. *Development-and-Psychopathology*, 3(2), 163-174.
- Wellman, H. (1990). *The child's theory of mind*. Cambridge, MA: MIT Press.

- Wellman, H., & Bartsch, K. (1994). Before belief: Children's early psychological theory. In C. Lewis & P. Mitchell (Eds.), *Children's early understanding of mind*. Hove: Lawrence Erlbaum.
- Wellman, H., Cross, D., & Watson, J. (2001). Meta-analysis of theory of mind development: The truth about false-belief. *Child Development*, 72(3), 655-684.
- Wellman, H., & Estes, D. (1987). Children's early use of mental verbs and what they mean. *Discourse Processes*, 10, 141-156.
- Wellman, H., & Gopnik, A. (1995). Why the child's theory of mind really is a theory. In M. Davies & T. Stone (Eds.), *Folk psychology* (p. 232-258). Oxford: Blackwell.
- Wellman, H., Harris, P., Banarjee, M., & Sinclair, A. (1995). Early understanding of emotion: Evidence from natural language. *Cognition and Emotion*, 9, 117-149.
- Wellman, H. M. (2001). Theory of mind and conceptual change. *Child Development*, 72(3), 702-702.
- Wells, C. (1981). *Learning through interaction: The study of language development*. Cambridge, UK: Cambridge University Press.
- 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.