A STUDY OF IDIOM COMPREHENSION
IN CHILDREN WITH SEMANTIC-
PRAGMATIC DIFFICULTIES

DEBRA KERBEL

A thesis submitted in partial fulfilment of the requirements of De Montfort University for the degree of Doctor of Philosophy

April 1996
CONTAINS PULLOUTS
ABSTRACT

In the light of inconclusive evidence regarding the comprehension of idioms by children with semantic-pragmatic difficulties, this study aimed to investigate the extent to which difficulty with idiom comprehension was characteristic of a group of primary school children considered to have semantic-pragmatic difficulties. It further aimed to explore whether such difficulty constitutes a valid means of discriminating children diagnosed as having "semantic-pragmatic disorder", Asperger syndrome or high-functioning autism, from children considered to have language disorders not primarily of a semantic and/or pragmatic type. It also compared these two groups with normally-developing children aged 6;6 to 7;6 or 10;6 to 11;6.

In view of the limitations of available procedures for assessing idiom comprehension in children with semantic-pragmatic difficulties, a new play-based methodology was developed for this study. Supplementary information was obtained from an additional definition task and from a symptom checklist.

The results indicate that the children with semantic-pragmatic difficulties did, as a group, demonstrate significantly fewer appropriate idiomatic interpretations and significantly more inappropriate interpretations than did any of the other three groups. However, the higher level of inappropriate scores reflected a larger number of "fuzzy" responses rather than significantly higher rates of literality.

Despite relative weakness, the children with semantic-pragmatic difficulties displayed appropriate interpretations considerably more often than they evidenced inappropriate ones. Within-group analysis reveals that the children diagnosed with Asperger syndrome or high-functioning autism performed less well than did those diagnosed with "semantic-pragmatic disorder". Nevertheless, both of these subgroups encompassed a considerable
range of comprehension ability. This variation appears to reflect essential differences in the critical semantic and pragmatic skills underpinning idiom comprehension. In combination with definition task data and broader knowledge of particular children, the play task may be used to identify the sites and sources of idiom comprehension breakdown in individual children.
ACKNOWLEDGEMENTS

This thesis bears only one name but owes a great deal to many contributors. In particular, I wish to thank my supervisor, Pam Grunwell, who has guided me through the PhD process, shared her extensive knowledge and experience, devoted innumerable hours to reading and discussing my work and supported me through all the ups and downs of the past three years. I thank too my second supervisor, Kim Grundy, for her clinical input, her editorial advice and her sustained encouragement.

Much support and many hours of thought-provoking discussion have also been contributed to this study by my clinical advisor, Rae Smith. My statistical advisor, Dave Rowley, has given generously of his time and expertise and has remained tirelessly patient and enthusiastic. Jude Balding not only performed all of the videoing for this study, but has also driven hundreds of miles on my behalf and has put up with listening to the story 270 times! Claire Norris undertook the arduous task of second-coding and has prepared me psychologically for every phase of the PhD. I extend my grateful thanks to them all. I thank too my friends, family and colleagues for their sustained interest and enthusiasm, and for bearing with me through the periods of crisis and frustration! In particular, I thank my husband, David, who constructed the play set for this study, and who has been my sounding board and the very best friend through every stage of this PhD.

I am further indebted to the staff of the Leicestershire Autism Outreach Service and of the following schools: Alderwasley Hall, Cherryleas Assessment Centre, Craven Lodge, Dawn House, Glenfield County Primary, Marriott Primary, Sherard County Primary, St Mary's Fields Infants, Sutherland House, Thurnby Lodge Primary and Uffculme School. Finally, I thank the teachers and clinicians who have supported this project, the parents who have entrusted me with their children and, especially, the ninety children who have each given a part of themselves to this study.
# CONTENTS

<table>
<thead>
<tr>
<th>ABSTRACT</th>
<th>ii</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>iv</td>
</tr>
<tr>
<td>TABLES</td>
<td>viii</td>
</tr>
<tr>
<td>FIGURES</td>
<td>xiii</td>
</tr>
<tr>
<td>CHAPTER 1: INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>CHAPTER 2: LITERATURE REVIEW</td>
<td>3</td>
</tr>
<tr>
<td>1. Definition of semantic-pragmatic difficulties in relation to the autistic continuum</td>
<td>3</td>
</tr>
<tr>
<td>2. Definition of semantic-pragmatic difficulties in relation to specific language impairment/language disorder</td>
<td>12</td>
</tr>
<tr>
<td>3. Definition and subclassification of idioms</td>
<td>13</td>
</tr>
<tr>
<td>4. How are idioms processed?</td>
<td>23</td>
</tr>
<tr>
<td>5. Idiom comprehension in normally-developing children</td>
<td>36</td>
</tr>
<tr>
<td>6. Anecdotal accounts of literality in children with semantic-pragmatic difficulties</td>
<td>62</td>
</tr>
<tr>
<td>7. Studies of idiom comprehension in children considered to have semantic-pragmatic difficulties or language disorders</td>
<td>66</td>
</tr>
<tr>
<td>8. Formal tests of idiom comprehension</td>
<td>75</td>
</tr>
<tr>
<td>9. Idiomatic usage in educational settings</td>
<td>87</td>
</tr>
<tr>
<td>CHAPTER 3: AIMS AND HYPOTHESES</td>
<td>89</td>
</tr>
<tr>
<td>CHAPTER 4: METHODOLOGY</td>
<td>93</td>
</tr>
<tr>
<td>1. Subjects</td>
<td>93</td>
</tr>
<tr>
<td>2. Materials</td>
<td>98</td>
</tr>
<tr>
<td>3. Procedure</td>
<td>108</td>
</tr>
</tbody>
</table>
CHAPTER 5: RESULTS ...................................... 129
1. Play task: between-groups analysis ................... 130
2. Definition task: between-groups analysis .......... 139
3. Between-tasks analysis .................................. 151
4. Within-group results .................................... 156
5. Analysis of the "SPD" and "ASP" subgroups of the SP group ............................................. 169
6. Between-idioms analysis .................................. 192
7. Miscellaneous statistical analyses ................... 197

CHAPTER 6: DISCUSSION ................................... 204
1. Idiom comprehension in the mainstream groups ......... 204
2. Idiom comprehension in the LD group as compared with the mainstream groups .................... 209
3. Idiom comprehension in the SP group as compared with the mainstream groups .................... 215
4. Idiom comprehension in the SP group as compared with the LD group ................................ 227
5. Accounting for the SP group's considerable competence on the play task .......................... 230
6. Accounting for the SP group's poor play task performance relative to the mainstream and LD groups: consideration of play task demands ................ 242
7. Sources of idiom comprehension impairment in children with semantic-pragmatic difficulties .... 253
8. Where and why did the SP children break down in comprehending idioms? Evidence from the within-SP group results ........................................... 274
10. Between-tasks differences ............................... 304
CHAPTER 7: CONCLUSIONS AND IMPLICATIONS ................. 314
1. Conclusions of the study ........................................ 314
2. Implications for the characterisation of children
   with semantic-pragmatic difficulties ......................... 316
3. Implications for differential diagnosis ..................... 317
4. Implications for differential diagnosis ..................... 317
5. Implications for teaching and therapy ...................... 318
6. Implications for research ..................................... 320

REFERENCES .......................................................... 322

APPENDIX A: Kerbel, Grunwell and Grundy (1996) ............ 339
APPENDIX B: Details of subjects ................................. 358
APPENDIX C: The Checklist ........................................ 366
APPENDIX D: Most frequently-occurring idioms in teacher
   and television language samples .............................. 376
APPENDIX E: The story ............................................. 377
APPENDIX F: Play set scanning and play task trials ....... 381
APPENDIX G: Prompts used on the play task and
   definition task .................................................. 384
APPENDIX H: Example of data transcript and coding ...... 386
APPENDIX I: Guidelines for data coding ....................... 391
APPENDIX J: Sentences repeated during the play task ..... 419
APPENDIX K: Ideas for therapeutic intervention
   focusing on idiom comprehension ............................. 420
TABLES

CHAPTER 4: METHODOLOGY

TABLE 1: Between-idioms analysis of inter-coder
agreement ........................................ 121
TABLE 2: Between-groups analysis of inter-coder
agreement ........................................ 122
TABLE 3: Relationship between broad and narrow codes .... 126

CHAPTER 5: RESULTS

TABLE 4: Play task - between-groups broad codes results . 130
TABLE 5: Play task - between-groups differences on
broad codes ........................................ 132
TABLE 6: Play task - breakdown of "appropriate" items
into "appropriate" narrow codes ............... 133
TABLE 7: Play task - between-groups differences on
"appropriate" narrow codes ..................... 134
TABLE 8: Play task - breakdown of "inappropriate" items
into "inappropriate" narrow codes ............. 135
TABLE 9: Play task - between-groups differences on
"inappropriate" narrow codes ................... 136
TABLE 10: Play task - discriminant analysis by group ..... 138
TABLE 11: Definition task - between-groups broad codes
results ............................................. 139
TABLE 12: Definition task - between-groups differences
on broad codes ..................................... 141
TABLE 13: Definition task - breakdown of "appropriate"
items into "appropriate" narrow codes .......... 143
TABLE 14: Definition task - between-groups differences
on "appropriate" narrow codes .................. 144
TABLE 15: Definition task - breakdown of "inappropriate"
items into "inappropriate" narrow codes ....... 145
TABLE 16: Definition task - between-groups differences on "inappropriate" narrow codes ................. 146
TABLE 17: Definition task - breakdown of "ambiguous" items into "ambiguous" narrow codes .............. 148
TABLE 18: Definition task - between-groups differences on "ambiguous" narrow codes .................... 149
TABLE 19: Definition task - discriminant analysis by group ......................................................... 150
TABLE 20: Between-tasks differences - broad codes ........ 151
TABLE 21: Between-tasks differences - narrow codes ....... 153
TABLE 22: Between-tasks analysis - differences between "appropriate" and "inappropriate" scores on each task ................................................................. 155
TABLE 23: Play task - MJ group results ..................... 157
TABLE 24: Play task - MI group results ..................... 158
TABLE 25: Play task - LD group results ..................... 159
TABLE 26: Play task - SP group results ..................... 160
TABLE 27: Play task - correlation between age and broad codes performance ............................. 161
TABLE 28: Play task - correlation between TROG scores and broad codes performance .................... 162
TABLE 29: Definition task - MJ group results ............. 163
TABLE 30: Definition task - MI group results ............. 164
TABLE 31: Definition task - LD group results ............. 165
TABLE 32: Definition task - SP group results ............. 166
TABLE 33: Definition task - correlation between age and broad codes performance ......................... 167
TABLE 34: Definition task - correlation between TROG scores and broad codes performance ............. 168
TABLE 35: Play task - "SPD"/"ASP" subgroups: frequencies and differences on broad codes .......... 169
TABLE 36: Play task - "SPD"/"ASP" subgroups: frequencies and differences on "appropriate" narrow codes .. 170
TABLE 37: Play task - "SPD"/"ASP" subgroups: frequencies and differences on "inappropriate" narrow codes ................................................................. 171
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>38</td>
<td>Play task – differences between &quot;SPD&quot;/&quot;ASP&quot; subgroups and MJ, MI and LD groups on broad codes</td>
<td>173</td>
</tr>
<tr>
<td>39</td>
<td>Play task – differences between &quot;SPD&quot;/&quot;ASP&quot; subgroups and MJ, MI and LD groups on &quot;appropriate&quot; narrow codes</td>
<td>174</td>
</tr>
<tr>
<td>40</td>
<td>Play task – differences between &quot;SPD&quot;/&quot;ASP&quot; and MJ, MI and LD groups on &quot;inappropriate&quot; narrow codes</td>
<td>175</td>
</tr>
<tr>
<td>41</td>
<td>Play task – discriminant analysis by group with SP group broken down into &quot;SPD&quot; and &quot;ASP&quot; subgroups</td>
<td>176</td>
</tr>
<tr>
<td>42</td>
<td>Definition task – &quot;SPD&quot;/&quot;ASP&quot; subgroups: frequencies and differences on broad codes</td>
<td>177</td>
</tr>
<tr>
<td>43</td>
<td>Definition task – &quot;SPD&quot;/&quot;ASP&quot; subgroups: frequencies and differences on &quot;appropriate&quot; narrow codes</td>
<td>178</td>
</tr>
<tr>
<td>44</td>
<td>Definition task – &quot;SPD&quot;/&quot;ASP&quot; subgroups: frequencies and differences on &quot;inappropriate&quot; narrow codes</td>
<td>179</td>
</tr>
<tr>
<td>45</td>
<td>Definition task – &quot;SPD&quot;/&quot;ASP&quot; frequencies and differences on &quot;ambiguous&quot; narrow codes</td>
<td>180</td>
</tr>
<tr>
<td>46</td>
<td>Definition task – differences between &quot;SPD&quot;/&quot;ASP&quot; subgroups and MJ, MI and LD groups on broad codes</td>
<td>182</td>
</tr>
<tr>
<td>47</td>
<td>Definition task – differences between &quot;SPD&quot;/&quot;ASP&quot; subgroups and MJ, MI and LD groups on &quot;appropriate&quot; narrow codes</td>
<td>183</td>
</tr>
<tr>
<td>48</td>
<td>Definition task – differences between &quot;SPD&quot;/&quot;ASP&quot; subgroups and MJ, MI and LD groups on &quot;inappropriate&quot; narrow codes</td>
<td>184</td>
</tr>
<tr>
<td>49</td>
<td>Definition task – differences between &quot;SPD&quot;/&quot;ASP&quot; subgroups and MJ, MI and LD groups on &quot;ambiguous&quot; narrow codes</td>
<td>186</td>
</tr>
</tbody>
</table>
TABLE 50: Definition task - discriminant analysis by group with SP group broken down into "SPD" and "ASP" subgroups ............................ 187
TABLE 51: Between-tasks differences on broad codes - "SPD"/"ASP" subgroups .......................... 188
TABLE 52: Between-tasks differences on narrow codes - "SPD"/"ASP" subgroups .......................... 190
TABLE 53: "SPD"/"ASP" subgroups - differences between "appropriate" and "inappropriate" scores on each task ...................................... 191
TABLE 54: Play task - level of "appropriacy" for individual idioms by group ..................... 192
TABLE 55: Definition task - level of "appropriacy" for individual idioms by group ..................... 193
TABLE 56: Play task - proportion of groups scoring "appropriate" on each idiom .................... 194
TABLE 57: Definition task - proportion of groups scoring "appropriate" on each idiom .......... 195
TABLE 58: Within-task concordance of group performances on idioms ranked in terms of "appropriacy" levels ......................................... 196
TABLE 59: Correlation of idiom rankings across tasks (by group) .................................... 196
TABLE 60: Play task - between-groups differences excluding data for "on the spot" for "inappropriate", "literal" and "inferred literal" .................. 198
TABLE 61: Definition task - between-groups differences excluding data for "on the spot" for "inappropriate", "literal" and "inferred literal" .................. 199
TABLE 62: Relationship between SP items coded "appropriate" on the play task and the codes "repeat", "narrative" and "repeat-narrative" on the definition task ............................. 200
TABLE 63: Correlation between Checklist language features and play task performance .......... 202
TABLE 64: Play task - correlation between number of SP trials and play task performance ............... 203

APPENDICES

TABLE 65: MI children - personal details and TROG scores . 358
TABLE 66: MJ children - personal details and TROG scores . 359
TABLE 67: LD children - personal details ......................... 360
TABLE 68: Language test results: LD1-LD5 ....................... 362
TABLE 69: Language test results: LD6-LD10 .................... 363
TABLE 70: Language test results: LD11-LD15 .................... 364
TABLE 71: SP children - personal details ......................... 365
TABLE 72: Proportion of subjects marked positive for Checklist language features (LD, SP, "SPD" & "ASP") ....................................... 375
TABLE 73: MI group - play set scanning and number of play task trials .................................................. 381
TABLE 74: MJ group - play set scanning and number of play task trials .................................................. 382
TABLE 75: LD group - play set scanning and number of play task trials .................................................. 382
TABLE 76: SP group - play set scanning and number of play task trials .................................................. 383
TABLE 77: Children for whom sentences were repeated during the play task ......................... 419
FIGURES

CHAPTER 2: LITERATURE REVIEW

FIGURE 1: Test of Language Competence - Expanded
(Wiig & Secord, 1989) Level 1 Item 11 ........... 77

CHAPTER 4: METHODOLOGY

FIGURE 2: Play set ................................... 103
FIGURE 3: Detail of play set - house ................. 104
FIGURE 4: Detail of play set - vet's office .......... 105
FIGURE 5: Detail of play set - police station ....... 106

CHAPTER 5: RESULTS

FIGURE 6: Play task - group results (broad codes) .... 131
FIGURE 7: Definition task - group results (broad codes) ........................................ 140
FIGURE 8: Play task - group/subgroup results (broad codes) ........................................ 172
FIGURE 9: Definition task - group/subgroup results (broad codes) ........................................ 181

APPENDICES

FIGURE 10: Checklist results - sections A-F .......... 371
FIGURE 11: Checklist results - section G .......... 372
CHAPTER 1: INTRODUCTION

Given that as much as two-thirds of the English language comprises idiomatic expressions (Boatner & Gates, 1969 cited in Conley, 1976), difficulty in comprehending idioms may have considerable impact on social interaction, literacy and learning. The literature suggests that substantial impairment in using and understanding idioms is commonly observed in children considered to have semantic-pragmatic difficulties (see below). The only systematic study of idiom comprehension in such children to date (Vance & Wells, 1994) failed to find any evidence of impairment in this group.

The aims of this study were to investigate the extent to which idiom comprehension impairment characterises children considered to have semantic-pragmatic difficulties, and to evaluate the validity of such disability as a diagnostic indicator. To this end, a group of children variously diagnosed with "semantic pragmatic disorder/difficulties", Asperger syndrome or high-functioning autism was identified. This group was compared with a second group of children considered by speech-language therapists to have language disorders not primarily of a semantic and/or pragmatic type. Two further groups included mainstream children aged 6;6-7;6 or 10;6-11;6.

In view of the serious limitations associated with conventional idiom assessment methodologies, a play-based idiom comprehension task was developed for this study. This procedure involved the videoing of subjects as they used props to act out a story played on audio-tape. Into the single narrative were embedded twelve common idioms drawn from teacher language samples and children's television. Following the play task, the children watched their videos and were asked to define each idiom as it occurred. Following the session, the demonstrations and definitions for each idiom were transcribed and coded. The background to the study and the development of the methodology, along with the
results and discussion of the findings are presented in this report.

Chapter 2 begins by outlining the diagnostic controversies against which this study was undertaken. The literature regarding the definition, subclassification and processing of idioms is then reviewed. Thereafter, the themes that have characterised research on normal development of idiom comprehension are traced and the anecdotal and research literature on idiom comprehension in children with semantic-pragmatic difficulties and/or language disorders is reviewed. Formal tests of idiom comprehension are critiqued and the literature on idiom usage in educational settings is considered.

Chapter 3 outlines the aims and main hypotheses of this study.

Chapter 4 sets out the methodology. It traces the development of the play and definition tasks, considers their reliability and validity, and presents the scoring categories devised for these procedures. A set of more detailed hypotheses based on the scoring codes is outlined.

Chapter 5 presents the findings of the play and definition tasks. The results are summarised and statistical analyses are reported.

Between-groups and within-groups results are discussed in Chapter 6 and are considered in the light of previous research and the hypotheses of the study. This chapter further explores possible sources of idiom comprehension disability, the relationship between idioms, and differences in the play and definition task results.

Finally, Chapter 7 summarises the conclusions of this study and their implications for clinical practice and future research.
CHAPTER 2: LITERATURE REVIEW

This chapter reviews and discusses the literature pertinent to the development of the present study. It begins by outlining the relationship between semantic-pragmatic difficulties, Asperger syndrome, high-functioning autism and specific language impairment/language disorder. Against this background, the sense in which these terms are used in this study is clarified. This is followed by a definition of the term "idiom" and an overview of linguistic subclassifications of idioms. The literature on idiom processing is then summarised. This is followed by a review and discussion of the literature on the acquisition and development of idiom comprehension in normally-developing children. Thereafter, the anecdotal discussion of literality in the literature on semantic-pragmatic difficulties is explored. This is followed by a review of systematic investigations of idiom comprehension in children considered to have semantic-pragmatic difficulties and/or language disorders. Formal idiom comprehension assessment tools are then critiqued. Finally, a review of the literature on idiom usage in educational settings is presented.

1. DEFINITION OF SEMANTIC-PRAGMATIC DIFFICULTIES IN RELATION TO THE AUTISTIC CONTINUUM

Research in the field of semantic-pragmatic difficulties is inevitably conducted within a context of ill-defined boundaries. Outlining the dual requirements of a valid diagnostic category, Szatmari (1992) stresses that it should be

possible to measure, or operationalize, the category in the first place.... (Furthermore, the category) should differ from related disorders on attributes other than the behavioral descriptors by which the disorder was originally defined (p. 584).

In the case of "semantic-pragmatic disorder", the case for
diagnostic validity is undermined by two factors. Firstly, it is not at all clear that individuals currently considered to have problems with semantics and pragmatics comprise a single, homogeneous group. Not only do individuals present with varying combinations of semantic and pragmatic problems, but it is further hypothesized that disability in these areas is underpinned by a range of different aetiologies (Fey & Leonard, 1983; Leinonen, 1991; Leinonen, Letts & Parke, 1993; McTear, 1991; McTear & Conti-Ramsden, 1992; Prutting & Kirchner, 1983; Smith & Leinonen, 1992). In view of the apparent heterogeneity in symptomatology and aetiology, the term "semantic-pragmatic difficulties" is used in preference to "semantic-pragmatic disorder" in this report. The use of this term is not intended to imply that semantic and pragmatic disabilities are necessarily co-existent.

The second factor complicating the consideration of semantic-pragmatic difficulties as a separate diagnostic category concerns its "fuzzy" boundaries with related disorders; autism in particular. The term "semantic-pragmatic syndrome without autism" was coined by Rapin and Allen in 1983. In 1987, the authors acknowledged that "semantic-pragmatic deficit syndrome" is frequently seen in verbal autistic children. Nevertheless, not all of their autistic sample were considered to have semantic-pragmatic deficit syndrome. The authors reported that the co-occurrence of semantic-pragmatic deficit syndrome and autism produced a more severe clinical picture than that of semantic-pragmatic deficit syndrome in the absence of autism.

While a diagnosis of "semantic-pragmatic disorder" may not be appropriate for all individuals with autism, it is generally agreed that all autistic individuals display some semantic and pragmatic difficulties. (See, for example, Leinonen, 1991, and Smith & Leinonen, 1992, for a discussion of the differing language profiles of individuals on the autistic continuum.) More controversial is the question of whether or not semantic-pragmatic difficulties may present in the absence of autism.
Rapin and Allen (1987) retained a distinction between the two diagnoses. Bishop and Rosenbloom (1987) asserted that most children with "semantic-pragmatic disorder" would never be called autistic: they do not have the avoidance of eye-contact and aloofness typical of infantile autism, and only mild tendencies to ritualistic and obsessional behaviour (p.34).

While recognising that children with semantic-pragmatic difficulties "tended to have mild autistic features", Bishop (1989) still maintained that these were "typically not severe enough or extensive enough to merit a diagnosis of autism" (p.273). She argues that it is unhelpful to use a single label to encompass very different types of difficulty (p. 274). Bishop thus retains the distinction in diagnosis between autism (defined in terms of the DSM-IIIR, APA, 1987, and Rutter's, 1978, criteria) and "specific semantic-pragmatic disorder" which applies to children who are not autistic but who initially present with a picture of language delay and receptive language impairment, and who then learn to speak clearly and in complex sentences, with semantic and pragmatic abnormalities becoming increasingly obvious as their verbal proficiency increases (Bishop, 1989, 118).

Despite retaining a diagnostic distinction between autism and "semantic-pragmatic disorder", Bishop nevertheless highlights the overlap between "semantic-pragmatic disorder", Asperger syndrome and autism. Bishop's model is concerned not only with severity, but also with patterns of symptoms. She represents the relationship in terms of a venn diagram set on two axes: interests and social relationships, and meaningful verbal communication. Placed in the area of greatest severity on both dimensions, autism is characterised by persisting social and language deficits. In contrast, Asperger syndrome is characterised by greater social deficits and fewer language problems, while semantic-pragmatic difficulties are characterised by persisting communication problems with milder social difficulties.
Notably, Bishop's model retains the concept of semantic-pragmatic difficulties as a disorder analogous to autism and Asperger syndrome. This differs from the linguistic approach proposed by Adams (1991) and McTear (1985). Adams observes that the terms "autism" and "semantic-pragmatic disorder" derive from different models. Whereas autism

is a categorical/diagnostic medical term, based on behavioural criteria which encompass far more than language behaviours; semantic-pragmatic disorder is essentially a descriptive linguistic term. Thus we should be able to apply it to language-impaired children, regardless of any medical diagnosis (p.73).

That "semantic-pragmatic disorder" can be validly or usefully distinguished from "autism" is strongly disputed by Aarons & Gittens (1992). The authors suggest that while "certainly, these children may not be classically autistic ... to regard children displaying such communication deficiencies as merely language disordered is too simplistic" (p.63). They assert that expanding our notions of the autistic continuum to take in semantic-pragmatic difficulties places these language problems in the only context which allows the children's underlying problems to be recognised. They thus propose that "semantic-pragmatic disorder" should be equated with "high-level autism".

Similarly, Wing (1991) asserts that "semantic-pragmatic disorder" is "virtually identical" to Asperger syndrome. The latter, according to Wing's thesis, falls within the "autistic continuum" (Wing, 1988) and is characterised by less severe manifestations than is Kanner's (1944) autism of the triad of social impairments: social interaction, social communication and social imagination. (Wing stresses the additional importance of variable skill and disability in language, non-verbal communication, reading, writing, calculation, visuo-spatial skills, fine and gross motor co-ordination and psychological functioning.)

Commenting on Wing's (1988) approach, McTear and Conti-Ramsden (1992) observe that an emphasis on the common triad of
impairments in autism, Asperger syndrome and "pragmatic difficulty" makes it difficult to compare and contrast these behaviours across these groups without reducing them all to a single syndrome. The authors express greater support for the approach of Bishop (1989) and Rapin and Allen (1983, 1986 cited in McTear & Conti-Ramsden, 1992) who acknowledge and address both the similarities and the variations in behaviour and communication style across these disorders. McTear and Conti-Ramsden highlight the importance of exploring the relationship between pragmatic abilities and other language and cognitive skills, as well as considering whether or not similar pragmatic difficulties are underpinned by similar causes. At this stage, the authors support the retention of a distinction between the disorders along the lines described by Bishop (1989).

Interestingly, while adhering to the view that "semantic-pragmatic disorder" / Asperger syndrome should be considered part of the autistic continuum, Wing (1991) does reflect on the value of studying the variations in which the autistic continuum can be manifested. Indeed, she cautions against the unquestioning equating of Asperger's syndrome with high-functioning Kanner's autism, noting that this might cause us to overlook differences which provide clues to underlying dysfunctions.

The close relationship between Asperger syndrome and Kanner's autism is underlined by the finding that some children with classic, Kanner-type autism appear to develop into teenagers and subsequently into adults with Asperger syndrome (Gillberg & Steffenburg, 1987 cited in Ozonoff, Rogers & Pennington, 1991; Wing, 1981, 1988, 1991). Further evidence for a link between the two derives from studies which have revealed the presence of both of these disorders within single families and even within sets of triplets (Bowman, 1988; Burgoine & Wing, 1983; De Long & Dwyer, 1988 all cited in Happé, 1995; Gillberg, 1991a).

While the similarity of autism and Asperger syndrome is widely acknowledged, numerous studies have sought to identify and
validate subgroups within the autistic spectrum\textsuperscript{1}. Comparisons of abilities in hypothesized subgroups such as Asperger syndrome, high-functioning autism and classic Kanner autism have thus far yielded inconclusive results (Klin, Volkmar, Sparrow, Cicchetti & Rourke, 1995; Manjiviona & Prior, 1995; Ozonoff et al., 1991; Szatmari, Bremner & Nagy, 1989). As noted by Happé (1995) and Ozonoff et al. (1991), the interpretation and comparison of these studies is made difficult by their variability in diagnostic criteria, approaches to grouping and choice of controls.

Evidence from cluster analysis and epidemiological prevalence studies generally appears to support a division of the spectrum into a number of subtypes (see Szatmari, 1992, for a comprehensive review of the studies up until 1992, as well as Eaves, Ho & Eaves, 1994; Ehlers & Gillberg, 1993). In a review prepared for the American Psychiatric Association DSM-IV Advisory Group on Pervasive Developmental Disorders, Szatmari (1992) concluded that at least three subgroups could be distinguished from "classical" autism on clinical grounds. These include "a low-functioning atypical group", "a high-functioning atypical group" and "Asperger syndrome". Szatmari stresses that while "the clinical differences appear reasonably robust across studies, samples, and methods of analysis, (the) evidence that these distinctions carry inferences with respect to etiology, clinical course and treatment is, however, only suggestive" (p. 596).

On the basis of current knowledge, the American Psychiatric Association (APA), like the World Health Organisation (ICD-10, 1992), has recently conferred separate status on Asperger syndrome (DSM-IV, 1994). It is acknowledged in the ICD-10, however, that Asperger syndrome is "a disorder of uncertain nosological validity" (p. 258).

Interestingly, the criterion on which the DSM-IV and ICD-10

\textsuperscript{1} Even the term "autistic spectrum" is used alternatively to include (Aarons & Gittens, 1992) or to exclude (Szatmari 1992) autism itself.
discriminate Asperger syndrome from autism is widely held to be inappropriate. ICD-10 adopts the stance that Asperger syndrome "differs from autism primarily in that there is no general delay or retardation in language..." (1992, 258). This view is echoed by the DSM-IV (1994) statement that "in contrast to Autistic Disorder, there are no clinically significant delays in general language (in Asperger syndrome)" (p. 77). By excluding semantic-pragmatic difficulties and prosodic abnormality, these classifications omit characteristics considered fundamental to Asperger syndrome by other authors (Asperger, 1944; Burd & Kebeshian, 1987 cited in Happé, 1995; Gillberg, 1991b; Gillberg & Gillberg, 1989; Szatmari et al., 1989; Wing, 1981, 1991). A more subtle language-based distinction is implicated in Wing's (1991) differentiation of Asperger syndrome and Kanner's autism:

the young classic Kanner's child has good visuo-spatial skills, good manual dexterity when engaged in his or her preferred activities but has delayed and deviant language development as well as social impairment of the aloof kind. Those with typical Asperger's syndrome have good grammatical speech from early in life, passive, odd or subtly inappropriate social interaction and poor gross motor co-ordination shown in gait and posture. They also tend to be in the mildly retarded, normal, or superior range of intelligence, while Kanner's group covers a wider range of the IQ scale with many being severely retarded.... (p. 115).

Wing's more detailed outline of the language difficulties typical of the upper end of the autistic continuum includes grammatical but long-winded speech and repetitive usage along with literal interpretations. (See Wing, 1991, for examples of more and less severe manifestations of each of the central elements of the autistic continuum.)

As evidenced in this review, the validity and operationalisation of the distinction between autism, Asperger syndrome and semantic-pragmatic difficulties remains controversial. Differences of opinion are evident too with regard to the clinical and educational value and limitations that result from
retaining or eliminating a distinction in terminology\(^1\). With regard to research, Rutter and Schopler (1992) point out soberingly that it is rather unlikely that the present criteria, in any of our systems, will exactly match up with the cause or causes of autism when they become known, or with distinctions in terms of course or response to intervention as may be demonstrated by future research. Accordingly, ... it is highly undesirable for research to be constrained by any one classification system (p. 474).

In the absence of clear and valid differentiating criteria, an attempt to assign children with semantic-pragmatic difficulties to discrete groups on the basis of symptom profiles was not considered useful for this study. Instead, it was decided that subjects with any of three diagnoses - "semantic-pragmatic disorder / difficulties", Asperger syndrome and high-functioning autism - would be included in a single semantic-pragmatic difficulties group (referred to as the "SP group"). The grouping together of these three diagnoses is not intended to imply that the terms are synonymous. This approach reflects the author's belief that a child with any of these three diagnoses is likely to display considerable difficulty in the areas of semantics and pragmatics. The author is of the opinion that one single processing fault is unlikely to underpin all of the semantic-pragmatic difficulties displayed in this group of children. It is not possible at this time, however, to draw any conclusions with regard to whether or not aetiological differences will be found to correspond to the groups of children to whom different labels are currently assigned.

In order to identify the range of children required for the study and to ensure a spread of symptom types, the labels "semantic-pragmatic difficulties / disorder", "Asperger syndrome" and "high-functioning autism" were used at the point of subject selection. The analysis considered the performance of the SP group as a whole as well as using the a priori diagnoses as a

\(^1\) For a discussion of these, see, for example, Aarons & Gittens (1992), Bishop (1989), Happé (1995), Szatmari (1992), Wing (1991).
starting point for investigating within-group differences. For this purpose, children diagnosed with "semantic-pragmatic disorder/difficulties" are dubbed the "SPD" subgroup, while children diagnosed with Asperger syndrome or high-functioning autism are referred to as the "ASP" subgroup.

In summary, the phrase "semantic-pragmatic difficulties" is used in this report as a descriptive term to denote problems at the levels of semantics and pragmatics which occur in autism and may or may not occur independently of autism or Asperger syndrome. This group is assumed to be symptomatically and aetiologically heterogeneous. Where the term "semantic-pragmatic disorder / difficulties" (in inverted commas) is used, this reflects the use of the term by a given author, or refers to the clinical use of this diagnosis.

The "SP group" comprises all subjects with semantic-pragmatic difficulties. This group breaks down into two subgroups on the basis of a priori diagnoses assigned to the given children by specialist speech-language therapists (SLTs), psychologists or psychiatrists. The "SPD" subgroup (always in inverted commas) includes those children diagnosed as having "semantic-pragmatic difficulties / disorder". The "ASP" subgroup (in inverted commas) includes children diagnosed as having Asperger syndrome or high-functioning autism.

---

1 The diagnoses discussed in relation to individual SP group children are those assigned by specialist SLTs, psychiatrists or psychologists prior to the study. In most instances, the diagnosis associated with a given child in this study is that recorded on the child's Statement of Special Educational Needs. It is acknowledged that these labels may well have been assigned inconsistently.

2 The grouping together of the diagnoses "Asperger syndrome" and "high-functioning autism" reflects the fact that many children in this subgroup were described to the author as having "Asperger syndrome / high-functioning autism" or "Asperger syndrome / autism".
2. DEFINITION OF SEMANTIC-PRAGMATIC DIFFICULTIES IN RELATION TO SPECIFIC LANGUAGE IMPAIRMENT / LANGUAGE DISORDER

The question of the validity of semantic-pragmatic difficulties as a construct distinct from autism and Asperger syndrome has obvious implications for the relationship between semantic-pragmatic difficulties and specific language impairment (a term synonymous in this report with "language disorder"). The term "specific language impairment" (SLI) or "language disorder" has been defined as a delay or deviance in receptive or expressive language which cannot be attributed primarily to other impairments such as autism or hearing loss (Bishop, 1992; Lees & Urwin, 1991; Maddick, 1986). Clearly, the view that semantic-pragmatic difficulties represents a subgroup of SLI (Rapin & Allen, 1983, 1987) entails acceptance of the belief that not all individuals with semantic-pragmatic difficulties lie within the autistic spectrum. Instead, in at least some cases, the primary impairment must be viewed as being of a linguistic nature.

Given that the validity of semantic-pragmatic difficulties as a construct distinct from autism, Asperger syndrome and neurological damage (Shields, 1991) remains a moot point, the status of semantic-pragmatic difficulties relative to SLI inevitably remains debatable. It does appear, however, that there is some basis on which children considered to have semantic-pragmatic difficulties can be distinguished from those said to have (other) SLIs. Unlike (other) SLIs, these children are characterised by a range of semantic and pragmatic difficulties in the absence of persistent major problems in phonology, morphology and syntax (Bishop & Rosenbloom, 1987; Rapin & Allen, 1983, 1987; Vance & Wells, 1994). Additionally, the nature of the social interaction problems witnessed in children with semantic-pragmatic difficulties are considered by some authors to differ from those seen in other language-impaired children (Bishop, Hartley & Weir, 1994).

¹Some relatively minor syntactic problems may persist.
For the purposes of this study, the term "language disorder" (abbreviated to LD) is applied to a heterogeneous group of children, drawn from language units, who have been identified by SLTs as showing primary impairment in phonology, morphology or syntax (see Appendix B). While it is acknowledged that these impairments might have secondary effects on the children's semantic, pragmatic, social and emotional functioning, the children's primary difficulties are considered by SLTs and other clinicians not to lie in any of these areas. It is further acknowledged that, owing to the definitional problems associated with the constructs "SLI" and "language disorder", these labels may not be consistently applied by all clinicians and researchers.

3. DEFINITION AND SUBCLASSIFICATION OF "IDIOMS"

For the purpose of this study, an idiom is defined as a short, multi-word expression, the meaning of which is fixed, is included in dictionaries and cannot be deduced from the meaning of the individual words.

This definition serves to distinguish this class of expressions from other forms of non-literal language including metaphors, similes, proverbs and indirect speech acts. A considerable range of expressions is, nevertheless still encompassed by this

---

1 The term "language unit" is used here to describe both special schools for children with language disorders as well as classes for language-disordered children within mainstream schools. With the exception of LD4, all LD subjects were drawn from the former.


3 Unlike idioms, metaphors are prevented from dictionary inclusion on the grounds that their meanings are not fixed and because they do not "express a different meaning of a lexeme, but rather (represent) the use of a lexeme with some features violated" (Strässler, 1982, 80). A simile also represents a relationship between meanings rather than a different meaning. Unlike idioms, proverbs typically encompass a whole sentence, may be judged true or false in their literal sense, and are invariably associated with a particular function (e.g. giving advice, admonishing or encouraging). As noted by Searle (1979 qtd in Strässler, 1982), "in indirect speech acts, the speaker means what he says. However, in addition, he means something more" (p.130). For a discussion of the relationship between indirect speech acts and idioms, see Strässler, 1982.
widely accepted definition of the term. As will become clear in
the review of idiom comprehension and processing studies, the
psychological literature has confined itself almost exclusively
to a very small group of idioms within this range. In contrast,
linguistic and lexicographic discussions have highlighted a broad
spectrum of idiom subtypes.

The notion of a spectrum of idiomaticity is central to a range
of diverse approaches to describing the relationship between
idiom types. It is widely accepted that idiomaticity (and,
indeed, literality) are more appropriately construed in terms of
a continuum than in terms of discrete groupings with clear
boundaries.

At the broad level, the model proposed by Wood (1981 cited in
Nattinger & DeCarrico, 1992) positions "true idioms" at one
extreme of a continuum that stretches all the way to "free
combinations" at the other pole. Wood's continuum is divided up
further by "collocations" (lying closer to "true idioms") and
"colligations" (closer to "free combinations"). This model is
based on the dimensions of compositionality and productivity
(i.e. the degree to which meanings are derivable from the
component words and the degree to which the syntactic form of the
expression constitutes the basis of other phrases). Thus, as the
most non-compositional and non-productive combinations,
expressions like "hell for leather" and "by and large" are
considered to be "true idioms". The meanings of "collocations",
such as "take umbrage" and "pay attention" are more easily
derived from their component parts and many of these expressions
allow a degree of lexical substitution (eg. "pay attention" or
"pay heed"). More compositional and more productive yet are
"colligations" such as "off with his head". Such expressions are
characterised by having a meaning which is largely predictable
by componential analysis and a syntactic form which is the basis
for a limited group of other similar expressions. (The form
"directional particle + with + NP" also yields the phrases "down
with the king" and "away with all bureaucrats", for example.)

14
At the extreme literal end of the continuum lie such "free combinations" as the literal sentence, "see the river". Such combinations are fully compositional and fully productive.

Based on related but different semantic and syntactic criteria, the continuum described by Cowie, Mackin and McCaig (1985, Oxford Dictionary of Current Idiomatic English, henceforth ODCIE) focuses on the idiomatic range of this spectrum. At one extreme of their continuum lie the fully petrified "pure idioms" exemplified by such expressions as "kick the bucket" and "blow the gaff". The relationship between their idiomatic and literal meanings is lost in history and the idioms remain lexically and syntactically invariable. Only slightly further along the continuum, this group of idioms shades into "figurative idioms" like "act the part" and "a narrow shave". These expressions, more semantically transparent than "pure idioms", may also allow a very restricted degree of lexical variation ("act the part/role", "a narrow/close shave") and pronoun substitution (eg. "Bill had a narrow shave and Fred an even narrower one"). Nevertheless, the meaning of a "figurative idiom" relates to the expression as a whole. In this sense, these can be distinguished from "restricted collocations" or "semi-idioms" in which one word carries a figurative sense not found outside of the expression while another element is used in its literal sense. In "jog one's memory" for example, "jog" is figurative while "memory" retains its usual literal meaning. Occupying the border zone between idiomaticity and literality, such expressions are included in the ODCIE on the grounds that the particular sense of the figurative word is unique to that context. Such "restricted collocations" contrast with "open collocations" like "fill the sink" or "a broken window", all the elements of which are literal and "freely recombinable".

Despite being granted idiom-like status by the ODCIE, "restricted collocations have not been included in previous studies of idiom comprehension. This study has, likewise, adopted the conservative stance of investigating only idioms of the "pure"
and "figurative" type. The idiomatic status of the items selected for inclusion was validated by the ODCIE. Unlike most other investigations, however, the present study has not confined itself to clause idioms (eg. "kick the bucket" and "spill the beans") and complex idioms (eg. "take the bull by the horns"). (Following the ODCIE Vol. 1, "complex idioms" are defined as units of meaning made up of a verb and/or a particle or preposition as well as other fixed elements.) In view of this study's concern with the comprehension of common idioms, it was considered appropriate that other idioms of different grammatical types should be included. Recordings of teacher and television language (detailed below) confirmed that phrase idioms (eg. "on the spot") and phrasal verb idioms (eg. "run over") are very commonly used in these contexts. The centrality of phrasal verb idioms in the body of idiomatic language is further underlined by the dedication of an entire volume of the ODCIE to idioms of this type (Cowie & Mackin, 1993). Against this background, it was considered appropriate that phrasal verb idioms be well represented among the items included in this study.

As in the case of phrase and clause idioms, idiomaticity in relation to phrasal verbs is best considered in terms of a continuum. Combinations of a "verb + particle and/or preposition" are subjected by the ODCIE to a series of tests in order to assess their semantic unity. A "true idiom", such as "step up" (increase) or "take off" (mimic), is one that meets the criteria of two separate tests. Firstly, the expression should be replaceable with a single word of equivalent meaning. In the case of "took off", for example, "Bill took off the Prime Minister to perfection" is synonymous with "Bill mimicked the Prime Minister to perfection". Secondly, it should not be possible to break the unity of the expression either by removing the particle element or by replacing the verb component. In

---

1 In compiling the Oxford Dictionary of Phrasal Verbs (ODCIE Vol. 1), Cowie and Mackin's (1993) principal concern was with the distinction between idiom and non-idiom rather than that between particles and prepositions. The authors point out that the "idiom / non-idiom distinction cuts across (the particle / preposition) contrast..." (p. 424).
"took off", the particle cannot be deleted (consider *"Bill took the Prime Minister to perfection"), nor can the verb be replaced (*"Bill snatched/grabbed the Prime Minister to perfection"). While not necessarily the case, many phrasal verb idioms will also be amenable to nominalization (e.g. "take off" can be converted into the noun "take-off"; "break down" can be converted into "breakdown").

In addition to containing the "true idioms" which pass at least the first two of these three tests, the ODCIE also includes a number of "semi-idioms" which lie on the border between idiomaticity and literality. When the first two of the tests outlined above are applied to such expressions as "put up (prices)" or "muck up", the results are inconclusive either way. In the case of "put up", the expression can be replaced with a single word -"raise" or "increase" - and the particle cannot be deleted. Thus far, the expression qualifies as an idiom. However, the verb can be replaced by "send" for example ("send up the prices") and thus it fails on this test. The ODCIE includes those "semi-idioms" from which the particle or preposition cannot be deleted without changing the sense or creating nonsense.

The difficulty of assigning either idiomatic or literal status to expressions that occupy the grey middle ground is widely acknowledged in the idiom literature (Cacciari, 1993; Dascal, 1987 cited in Levorato, 1993; Gibbs, 1984, 1989 both cited in Levorato, 1993; Levorato, 1993; Strässler, 1982). The phrasal verb idioms included in this study are all indisputably idiomatic. In some cases, however, their counterparts may be seen to fall on the border between idiomaticity and literality. (These are included in the ODCIE.) In such instances, the phrasal verb idiom is considered in this study to have a "borderline counterpart". This contrasts with phrasal verb idioms having a "clearly literal counterpart".

As evidenced by the varying bases on which different authors have
divided up the idiomatic continuum, idioms can be seen to differ from one another in terms of many different linguistic dimensions.

One of the earliest idiom variables to be identified related to what Fraser (1970) described as the "frozenness hierarchy". The degree of frozenness is determined by the extent to which an idiom can be syntactically transformed while still retaining its figurative meaning. The expression "lay down the law" can be viewed as less frozen (or more flexible) than "kick the bucket", for example. The former retains its figurative sense under gerund nominalization ("Her father's laying down the law prevented her from going to the dance"), adverb insertion ("The boss will quickly lay down the law if anyone shows up late"), particle movement ("They will lay the law down if the party gets too wild"), passivization ("The law will be laid down when Jane's boyfriend finds out where she's been"), and action nominalization ("The supervisor's laying down the law was just what the staff needed") (Gibbs & Gonzales, 1985, 245). In contrast, the figurative sense of "kick the bucket" ("to die") is lost when these transformations are applied (consider *"The kicking of the bucket occurred at 8pm"; *"The bucket was kicked by Jim").

Following Fraser, whose hierarchy was based on his own intuitions, Gibbs & Gonzales (1985) constructed a frozenness hierarchy based on speakers' intuitions. Both lists of relative frozenness have been found to correlate with the date of origin of the respective idioms (Cutler, 1982; Gibbs & Gonzales, 1985).

In addition to syntactic frozenness, idioms have also been considered to vary in lexical frozenness (Gibbs, Nayak, Bolton & Keppel, 1989a). An idiom such as "button your lip" permits the substitution of a synonym for the verb (eg. "zip your lip" or "fasten your lip"). Other idioms, like "kick the bucket" are not amenable to lexical substitution (consider *"punt the bucket" or

---

1 For further discussion about the syntactic behaviour of idioms, see Huddleston (1984), Newmeyer (1974) and Quang (1971).
*"kick the pail"*). As discussed above, this distinction forms part of the basis for that drawn between "figurative idioms" and "pure idioms" in the ODCIE (Cowie et al., 1985).

In order to explain the variability in the syntactic and lexical flexibility of idioms, Gibbs and his colleagues have argued that it is necessary to move away from the traditional noncompositional view of idioms (upheld by such authors as Chomsky, 1980; Fraser, 1970; Heringer, 1976; Katz, 1973 all cited in Gibbs & Nayak, 1989; Johnson-Laird, 1983). Following Nunberg (1978 cited in Gibbs & Nayak, 1989), these authors highlight a group of idioms in which the individual parts do contribute to the overall idiomatic sense (Gibbs & Nayak, 1989; Gibbs et al., 1989a). In the idiom "pop the question", for example, the verb "pop" refers to "suddenly ask", while "the question" refers to "a marriage proposal". In combination, these two parts produce the expression's overall idiomatic meaning of "propose marriage". This idiom qualifies as decomposable by virtue of its parts having a literal or figurative meaning which contribute to the idiomatic sense of the expression as a whole. In contrast, nondecomposable idioms cannot be broken down into parts which relate to the overall meaning. In the case of "kick the bucket", for example, neither "kick" nor "bucket" contributes to the overall idiomatic meaning, "to die". This is therefore a nondecomposable idiom.

A further distinction is drawn between "normally decomposable idioms" and "abnormally decomposable idioms" (Gibbs et al., 1989a; Gibbs, Nayak & Cutting, 1989b; Gibbs & Gonzales, 1985; Gibbs & Nayak, 1989; Nunberg, 1978 cited in Fellbaum, 1993). Each of these involves a subtly different relationship between the overall idiomatic meaning and the component parts. "Normally decomposable idioms", like "pop the question", break down into components which directly contribute to their overall figurative meanings. In "abnormally decomposable idioms", the relationship between the parts and the whole is mediated by a metaphor. Thus, "carry a torch (for someone)" is an abnormally decomposable idiom.
because the overall meaning "have admiration for (someone)" is related to the component part "torch" via a metaphor (torch is a conventional metaphor for warm feelings).

The concept of a continuum of compositionality is supported by the finding that adults can reliably judge whether certain idioms are more decomposable (or "analyzable") than others, and that some analyzable idioms are normally decomposable whereas others are abnormally decomposable (Gibbs & Nayak, 1989; Gibbs et al., 1989a). The validity of this concept is further strengthened by the correlation that has emerged between the degree of analyzability and both the syntactic and lexical frozenness of idioms (Gibbs & Nayak, 1989; Gibbs et al., 1989a). These studies have found that idioms viewed as normally decomposable are also considered to be more productive syntactically than either abnormally decomposable or nondecomposable idioms (Gibbs & Nayak, 1989).

Particularly in view of the explanatory potential of this concept, it is important to observe that Gibbs' compositionality theory and the traditional non-compositional view are not necessarily mutually exclusive. The confusion appears to stem from subtly different uses of the term "compositionality". Gibbs considers it untrue to state that the meaning of the expression is not equal to the sum of its parts because a degree of correspondence between the parts and the overall meaning can sometimes be identified. Gibbs terms this phenomenon "semantic analyzability" (my stress) but it could be argued that the relationship he describes is more syntactic than semantic. Indeed, at the syntactic level, it is sometimes the case that the whole is equal to the sum of the parts. The statement, "the whole is not equal to the sum of the parts", can also be interpreted at the semantic level, however. That is to say, it is not possible to derive the overall meaning of any idiom simply by putting together the literal meanings of the individual words. This remains true even in relation to a decomposable idiom like "pop the question". Despite the syntactic correspondence between
"pop" and "ask", "the question" and "marriage proposal", "pop" nevertheless does not actually mean "ask". As such, it would seem that Gibbs' concept of decomposability is not incompatible with the definition of idioms used in this and other studies. Rather than arguing for one or the other use of the term "compositional" being more accurate, it would seem preferable to introduce a new term for Gibbs' "semantic analyzability" which would allow the two senses of the term "compositional" to be more clearly distinguished. The term "syntactic correspondence", for example, might be used to describe Gibbs' notion of "semantic analyzability".

The internal structure of idioms has been further considered by Rose (1978). Rose differentiates between "semantically intact" idioms in which the usual lexical meanings of the components are maintained (e.g. "bite the dust", "kick the bucket") and "non-lexical" idioms whose internal semantic relationships have been altered (e.g. "go crazy", "take a walk", "run a fever"). Semantically intact idioms involve a wide variety of verbs which maintain their ordinary meaning. Non-lexical idioms employ a small number of verbs such as "do", "go", "make" and "have", which are semantically empty. The meaning of these idioms is assigned primarily by their direct objects.

Looking beyond internal semantics, researchers have observed additional variations based on the semantic relationships that hold between overall idiomatic and literal meanings. An idiom is considered "metaphorically transparent" if there is a clear relationship between the idiomatic and literal meanings (Gibbs, 1987; Wood, 1986 cited in Stock, Slack & Ortony, 1993). This is demonstrated by the expression "skating on thin ice": here the literal meaning of the expression is actually an example of its idiomatic meaning ("to be in a precarious position"). It is, therefore, metaphorically transparent. In contrast, the literal and figurative meanings of "kick the bucket" are not clearly related and the idiom is thus regarded as being metaphorically opaque. It has been noted that the figurative-literal
relationship for at least some opaque idioms was once transparent but is now buried in history (Ortony, Schallert, Reynolds & Antos, 1978; Stock et al., 1993). It is thus perhaps more accurate to refer to relative metaphorical transparency at a given point in historical time.

The transparency/opacity dimension, like many of the other variables discussed above, is notable in that it relates solely to those idioms which have a literal counterpart (what Van Lancker and Canter (1981) refer to as "ditropic idioms"). In fact, a sizeable group of English idioms do not have literal counterparts. These "literally ill-formed" idioms (Gibbs, 1993) break down into those that are semantically ill-formed and those that are syntactically ill-formed (Botelho da Silva & Cutler, 1993). The former usually involves a violation of selection restrictions. "Swallow your pride", for example, violates the restriction that the verb "swallow" must take a physical object. Syntactically ill-formed idioms typically violate subcategorization restrictions. In the case of the idiom "in the know", "know" functions as a noun. In non-idiomatic language, "know" cannot function as anything but a verb (Botelho da Silva & Cutler, 1993).

While some expressions clearly have no meaningful literal counterpart, many idioms have a possible literal counterpart that is nevertheless unlikely to be encountered. Thus, some context may be constructed in which expressions like "stew in your own juice" or "on cloud nine" might be used literally, but there is a far greater likelihood that these expressions will be used figuratively. Popiel and McRae (1988), Cronk, Lima and Schweigert (1993) and Van Lancker and Canter (1981) have all observed that idioms do in fact vary along an axis of literalness.

1The idiom, "kick the bucket", for example, derives its meaning from the old English slaughtering practice of tying pigs in such a way that their back legs kicked a beam, known as a "bucket" (Ortony et al., 1978).

2Levorato (1993) observes an additional aspect of the relationship between the literal and idiomatic senses of an idiom like "kick the bucket". She suggests that the use of this idiom to describe a quick death rather than a drawn-out one is motivated by the fact that the act of kicking takes only an instant.
and that the familiarity of the literal sense is independent of the familiarity of the idiomatic sense. An idiom may thus be characterised in terms of high or low frequency of its idiomatic sense alongside high or low frequency of its literal sense.

Associated with these ideas is a further continuum outlined by Mueller and Gibbs (1987). They note that expressions vary in terms of the number of meanings they can convey. Thus, some expressions (eg. "pop the question") have only one idiomatic sense and no literal sense. Others have one idiomatic and one literal sense (eg. "kick the bucket"). Others, such as "give a hand", have more than one idiomatic sense in addition to their literal sense ("applaud" or "assist"). Furthermore, Mueller and Gibbs note, metaphorically transparent idioms have literal and idiomatic senses that are isomorphic (eg. "skating on thin ice"). These may be viewed as falling somewhere between those without a literal counterpart and those with clearly distinct literal and figurative senses.

As revealed by this review of the linguistic literature, idioms are most accurately viewed as a heterogeneous class of related items. The history of idiom processing and comprehension research has highlighted the importance of taking these dimensions into consideration when assessing idiom-related behaviour. Indeed, as more and more of these factors are controlled for in research studies, their influence on comprehension and processing is increasingly being recognised.

4. HOW ARE IDIOMS PROCESSED?

Research on idiom processing has focused primarily on two issues: how idioms are stored in the mind and whether idiomatic and literal interpretations are retrieved simultaneously or in sequence. Three major models have emerged in response to these questions. These are based on studies of reaction times on
literal and figurative processing tasks. The "literal first hypothesis" (Swinney & Cutler, 1979) or "literal processing model" (Schweigert & Moates, 1988), based on the work of Bobrow & Bell (1973) and Clark & Lucy (1975 cited in Johnson, 1985), proposes that idioms are stored in a special list which is not part of the general mental lexicon. According to this hypothesis, a literal interpretation is always constructed first. If this interpretation is found to be incongruent with the context, an idiomatic meaning is then retrieved from the idiom list.

An alternative view, variously referred to as the "lexical representation hypothesis" (Swinney & Cutler, 1979) or "simultaneous processing model" (Schweigert & Moates, 1988), proposes that both idiomatic and literal processing are activated simultaneously as soon as the first word of an idiomatic expression is encountered (Fraser, 1974; Heringer, 1976 both cited in Glass, 1983). According to this model, idioms are stored in the general lexicon and are accessed as single units by a direct retrieval process. This process is quicker than the componential analysis required for literal interpretation. This, it is argued, accounts for the more rapid processing of idiomatic than literal meanings reported by Swinney & Cutler (1979) and others (Cronk & Schweigert, 1992; Gibbs, 1980, 1985; Gibbs & Gonzales, 1985; Glass, 1983; Ortony et al., 1978).

The processing time advantage for idiomatic meaning is alternatively explained by the "idiomatic first" hypothesis (Gibbs, 1980, 1985, 1986), also referred to as the "idiomatic processing model" (Schweigert & Moates, 1988). According to this theory, the idiomatic meaning is always retrieved first. If the idiomatic meaning is appropriate in the context, no literal, componential analysis takes place. Only if the figurative meaning is found to be incongruent with the context is literal processing activated. Gibbs concedes that listeners may momentarily examine the conventional meanings of individual words in the expression, but proposes that they do not combine these
meanings to form a literal interpretation. (The timing and mechanics of this termination of literal processing is not discussed by Gibbs but has been inconclusively considered by other authors including Peterson and Burgess (1993) and Needham (1992).)

While some degree of support has been mustered for each of the three hypotheses, no one model has been consistently or completely supported (Bobrow & Bell, 1973; Brannon, 1975 cited in Schweigert, 1991; Cacciari & Tabossi, 1988; Cronk & Schweigert, 1992; Estill & Kemper, 1982; Gibbs, 1980; Gibbs & Gonzales, 1985; Glass, 1983; Mueller & Gibbs, 1987; Needham, 1992; Schweigert, 1986, 1991; Schweigert & Moates, 1988; Swinney & Cutler, 1979; Tompkins, Boada & McGarry, 1992). The discrepancies in these findings are probably due, in large part, to the inconsistency of controls. Early research tended to consider idioms as a homogeneous class of items. Increasing awareness of the heterogeneity of idioms (as discussed above) has yielded a large crop of studies investigating the effects on processing of such factors as idiom familiarity, the likelihood of an expression occurring in the literal rather than idiomatic sense, the number of figurative meanings associated with a given phrase, and the frequency with which individual words within the phrase occur (Cacciari, 1993; Cronk et al., 1993; Cronk & Schweigert, 1992; Estill & Kemper, 1982; Flores d'Arcais, 1993; Forrester, 1995; Gibbs, 1980; Gibbs & Gonzales, 1985; Mueller & Gibbs, 1987; Popiel & McRae, 1988; Schraw, Trathen, Reynolds & Lapan, 1988; Schweigert, 1986, 1991; Schweigert & Moates, 1988).

The results of these studies indicate that none of the three models outlined above is likely, on its own, to account for the processing of all idioms.

The insight gained through studies of various idiom subgroups has given rise to several, more complex reformulations of earlier models. These have invariably involved revised conceptualisations of the mental lexicon. Having found that syntactically frozen idioms are processed more quickly than
flexible ones, Gibbs and Gonzales (1985) outlined three possible lexical models. Following Morton (1969, 1979 cited in Gibbs & Gonzales, 1985), their "threshold model"\(^1\) views idioms as logogens with different threshold levels corresponding to their respective degrees of frozenness. Frozen idioms are considered to have a lower threshold than flexible ones and, hence, to require less information from the input stimulus in order to be accessed. The lower threshold associated with frozen idioms would account for the authors' finding that these are retrieved faster than flexible idioms.

The second possibility outlined by Gibbs and Gonzales (1985) is a "search model". According to this model, each syntactic form of an idiom is entered separately in the mental lexicon. It is proposed that lexical access is a process by which an input stimulus is matched to the representation of a word or phrase in the lexicon. Partial analysis of a stimulus results in the generation of a set of candidates, which is then searched serially until a match with the stimulus is found. The set of candidates is... searched in order of frequency of occurrence of the item represented.... If the idioms have multiple entries, it would be expected that the entries for the most frequent syntactic forms would be found first from the search. In this way, idioms with restricted syntactic forms, which are most frequent, would be recognized faster than idioms that could be seen in a variety of syntactic forms (p. 257).

One weakness in this proposal relates to the assumption that "restricted syntactic forms (are) most frequent". In fact, a comparison of relative idiom frozenness as rated in Gibbs & Gonzales (1985) with rates of frequency (Popiel & McRae, 1988) indicates that there is no consistent relationship between them. The idiom "turn over a new leaf", for example, is rated as being more flexible syntactically than is "put on a good face". The former is, however, judged to be the more frequent idiom. In cases like this, each form of the flexible idiom may be more frequent than is the single form of the more frozen idiom. Gibbs

\(^1\)See Garnham (1985) for a discussion of "threshold models" and "search models" of lexical acquisition.
and Gonzales' search model fails to explain why frozen idioms should be processed more quickly in such cases.

It is further notable that this model is partially contradicted by another proposed by Mueller and Gibbs (1987). Here the authors argue that idioms with more entries in the mental lexicon are accessed more quickly than those with fewer entries because "it is more likely that one of their meanings will be encountered in a short time" (p. 63). An expression like "pop the question", for example, has only one meaning (its figurative sense). It therefore has only one entry in the lexicon. "Hit the sack" has two entries (one literal and one figurative). "Give a hand" has three entries (one literal and two figurative). According to this model, "give a hand" should be accessed more quickly than "hit the sack" which, in turn, should be accessed more quickly than "pop the question".

Interestingly, it seems that it would be possible to reconcile Gibbs and Gonzales (1985) and Mueller and Gibbs (1987) if a different mechanism were assumed to underlie the processing advantage for multi-entry idioms. Expressions that are used in a greater number of senses (idiomatic and literal) would tend to occur (and thus be encountered) more frequently than expressions with only one possible meaning. The higher frequency of a multi-sense expression would, in accordance with Gibbs and Gonzales' theory, result in its being matched to the input stimulus more quickly than would be the case for a single-sense expression. (It would, of course, be necessary to prove that multi-sense expressions do in fact occur more frequently than single-sense expressions.)

As it stands, Gibbs and Gonzales' (1985) search model fails to address an important issue. Interpretation involves not just retrieving possible meanings, but also selecting from among all the candidates. In the case of a single-sense idiom, there is no choice involved. Hence this stage can be eliminated. Multi-sense idioms do however necessitate a selection process. Thus,
even though the initial accessing of multi-sense idioms may be quicker than for single-sense idioms, the additional selection process would cause the processing of multi-sense idioms to be delayed. In order to account for their results (multi-sense idioms being processed quicker than single-sense ones), Mueller and Gibbs must assume that the combined process of retrieval and selection is faster for multi-sense idioms than for single-sense ones.

A third mental lexicon structure proposed by Gibbs & Gonzales (1985) considers the possibility that idioms are not represented by each of their permissible syntactic forms. Instead, each idiom is entered only once in the lexicon. A body of abstract knowledge guides speakers with regard to the syntactic operations permissible for any given idiom.

On the basis of the proposals outlined in Gibbs and Gonzales (1985), this theory appears to require some further elaboration. In particular, the authors do not discuss whether idioms with multiple idiomatic meanings (such as "give a hand") would have one or more idiomatic entries in the lexicon. Nevertheless, support for this theory derives from Bothelho da Silva's (1993) Portuguese study. The subjects read a set of 20 idioms which were presented in context and had been submitted to a range of syntactic transformations. In a subsequent recall task, it was noted that most of the transformed idioms were recalled in untransformed form.

Further support for this view of idiom processing derives from a series of experiments conducted by Gibbs and his colleagues. Firstly, evidence has emerged for a correlation between the degree of semantic analyzability (as defined above) and syntactic flexibility (Gibbs & Nayak, 1989). Idioms that are normally decomposable also tend to be syntactically flexible, while idioms that are nondecomposable tend to be more syntactically frozen. Secondly, it appears that readers are consistently able to differentiate degrees of semantic analyzability (Gibbs & Nayak,
1989; Gibbs et al., 1989a). It is therefore conceivable that speakers' knowledge regarding the semantic analyzability of individual idioms enables them to predict which syntactic transformations are permissible for a given idiom.

In addition to its relationship with syntactic flexibility, semantic analyzability has also been found to correlate with lexical flexibility. Normally decomposable idioms allowed words to be substituted with synonyms more often than did abnormally decomposable or nondecomposable idioms (Gibbs et al., 1989a). Thus, language users' sensitivity to the semantic analyzability of idioms might enable them to predict both the syntactic and lexical behaviour of individual idioms.

In order to account for these proposed operations, Gibbs et al. (1989b) outline a new processing model dubbed the "idiom decomposition hypothesis". In contrast to Gibbs' earlier model, this hypothesis abandons the assumption that all idioms are stored and retrieved as single units. Instead, the authors suggest that when an idiom is encountered, listeners automatically attempt a compositional analysis. If the idiom is decomposable, then the listener can assign independent meanings to the individual parts. The listener then recognises how these meaningful parts combine to form the overall idiomatic meaning. For example, when encountering the phrase "pop the question", the listener first recognises that "pop" refers to "suddenly ask" and that "the question" refers to "a marriage proposal". These two meaningful components are then combined to produce the overall idiomatic meaning, "propose marriage". When a nondecomposable idiom is encountered (e.g. "kick the bucket"), the listener first attempts to assign meaning to the individual components. When no meaning is found in the lexicon for the individual components, the lexicon is then searched for the entire string and the idiomatic meaning is retrieved. These nondecomposable idioms are considered to be more lexicalised because they are used in few

1Stock et al. (1993) use an artificial intelligence framework to address the mechanism by which idioms are parsed in line with theories of semantic analyzability and transparency.
syntactic forms (Gibbs & Gonzales, 1985; Gibbs & Nayak, 1989).

Importantly, while decomposable idioms are broken down in comprehension, the meanings assigned to the individual parts are not the literal meanings of the individual words (Gibbs et al., 1989b). In processing "pop the question", for example, listeners do not assign the meaning "burst" to "pop" and "request for information" to "the question". With regard to this aspect of the relationship between idiomatic and literal processing, Gibbs et al. (1989b) are very clear. Their stance on the more general relationship between idiomatic and literal interpretation is, however, a little less clear. Commenting on the idiom decomposition hypothesis, Gibbs et al. (1989b) note that the comprehension of idioms is "similar to (the) comprehension of 'literal' language (in that idioms undergo) a good deal of syntactic processing" (p.590). The authors claim to be "neutral with respect to whether or not understanders actually activate the 'literal' senses of individual words during idiom processing" (p.589). Gibbs and Nayak (1989) go so far as to suggest that "there may be little utility...in maintaining the widely held belief that different processes drive the comprehension of idiomatic and literal language..." (p.132). Despite this assertion, however, the authors remain committed to the view that there are important differences in the processing of idiomatic and literal language.

They propose, firstly, that different kinds of information are activated in literal and idiomatic interpretations. Following Lakoff (1987) and Lakoff and Johnson (1980), they suggest that idiomatic meanings are based on a metaphorical mapping of information across two domains that have common structural properties (Gibbs et al., 1989b; Gibbs & Nayak, 1989; Nayak & Gibbs, 1990). Thus, in processing the sentence "John got angry", the listener retrieves the literal meaning of each of the individual words. In interpreting an idiom like "John let off steam", however, listeners activate underlying conceptual metaphors such as THE MIND IS A CONTAINER. Listeners thus "map
their knowledge of a pressurized container with steam escaping to a person who is releasing tension because he is angry" (Gibbs and Nayak, 1989, 132). Nayak and Gibbs (1990) propose that idioms whose meanings are generated from the same conceptual metaphor may be linked together in the mental lexicon (e.g. "let off steam", "flip your lid", "blow your top", "hit the roof")

A second feature distinguishing idiomatic and literal processing relates to the difference in reaction times between the two. Even when a literal expression comprises words used more frequently, the literal expression is still processed more slowly than an equivalent idiomatic expression (Gibbs et al., 1989b). The authors suggest that the componential analysis of a familiar combination of words (such as occurs in idioms) is a quicker process than a componential analysis of a novel, literal combination of words.

As noted above, many studies have supported the view that idioms are accessed more quickly than literal expressions. Significantly, however, all of this support has emerged from off-line studies i.e. the tasks employed (e.g. pictorial multiple choice or definition tasks) are "distant in time (and correspondingly in terms of cognitive operations) from the initial retrieval or computation of meaning" (Tompkins et al., 1992). In contrast, on-line tasks involve a response during the comprehension process. These include tasks in which subjects monitor a read or heard sentence for a target word. For example, subjects might hear the sentence "By the fourth day in hospital, Orville was climbing the walls to get out". They would press a button as soon as they registered the target word. This may be the same as that in the sentence (e.g. "walls"), a rhyming word ("falls") or a category ("part of a building") (Estill & Kemper, 1982). It has been suggested that on-line tasks are more

---

1 This theory is supported by Forrester (1995) who presented subjects with phrases related to real idioms semantically but composed of words not usually associated with those idioms (e.g. "fly in the ointment" - "smudge on the sheepskin"). When encountering these "semantic equivalent phrases" in context, adult listeners interpreted them as they would the associated real idiom. For a discussion of the limitations of the metaphorical mapping theory, see Kreuz & Graesser (1991) and Stock et al. (1993).
appropriate than off-line ones for studying the relatively automatic processes involved in language comprehension (Marslen-Wilson & Tyler, 1980; Tyler, 1988; Tyler & Marslen-Wilson, 1982 all cited in Tompkins et al., 1992).

Significantly, studies using on-line tasks (Cacciari & Tabossi, 1988; Estill & Kemper, 1982) have reported less conclusive results with respect to processing time advantages. Estill and Kemper (1982) reported no reaction time advantages for either literal or idiomatic expressions on any of their monitoring tasks. Cacciari and Tabossi (1988) note that subjects could identify the idiomatic meaning of Italian expressions more quickly than the literal meaning when they were able to guess the last word of the idiom on the basis of the first few words. For example, in the case of "he was in seventh heaven", the word "heaven" can be predicted once one has heard the words "he was in seventh". Significantly, however, the literal meaning appeared to be primed to a greater extent than was the idiomatic meaning in the case of less predictable idioms.

This finding forms a cornerstone for Cronk et al.'s (1993) revised simultaneous processing model. In addition to its reliance on Cacciari and Tabossi's (1988) results, the model is based on the authors' investigation of the relationship between idiom familiarity (of the idiomatic meaning), literalness (i.e. the familiarity of the literal meaning) and the frequency of the individual words (as rated by Kucera & Francis, 1967 cited in Cronk et al., 1993). The model retains the assumption that idioms are stored in the mental lexicon as single units. However, the idiomatic meanings of less familiar idioms are viewed as being more difficult to locate than those of familiar idioms. The authors propose that when an idiom string is encountered, the phrase processor begins processing it from left to right. At some critical point in the string, it becomes possible to identify it as an idiom. At this point, the process splits into parallel literal and figurative subprocesses. These operate simultaneously. The general lexicon is searched for both
the whole figurative meaning as well as for the meanings of the individual words. The time it takes for figurative meanings to be found depends on how familiar the idiom is. The time it takes for literal meanings to be computed depends on the literalness of the phrase. Whenever either process retrieves a potential meaning, a context check occurs. If a constructed literal meaning is found to be appropriate, it immediately serves as the output. Figurative candidates are, similarly, checked against the context, but then undergo an additional check for the frequency of the individual words. If the phrase contains low-frequency words, the figurative meaning is treated as if it were contextually inappropriate. When contextually inappropriate meanings are encountered (be they literal or figurative), the phrase processor waits for the alternative meanings to be retrieved. At this point, a further context check occurs and all available meanings are considered before the most appropriate meaning is selected. Thus, for the literal sense of a given phrase, the processing time is dependent only on its literalness. In order for an idiomatic sense to be accessed quickly, however, it must be both familiar and contain high-frequency words.

While Cronk et al. (1993) cite Cacciari and Tabossi (1988) in support of their model, the latter have sketched an alternative hypothesis which differs from that of Cronk et al. in an important sense. Cacciari and Tabossi propose that idioms are not encoded in the lexicon as separate entries at all. Instead, the meaning of an idiom is associated with a particular configuration of words and becomes available as soon as that configuration is recognised. These configurations are composed of the same lexical items that are activated during the comprehension of literal language. Thus, "take" for example, occurs in the configurations "take the bull by the horns" and "take to heart", as well as being a lexical entry that is activated in order to understand a literal sentence like "The boy

---

1 The processing time advantage for frequently-occurring as opposed to infrequently-occurring idioms and for those more likely to occur idiomatically than literally has been generally supported (Cronk & Schweigert, 1992; Forrester, 1995; Schweigert, 1986; Van Lancker & Canter, 1981.
took the book"1. When an idiom is encountered, literal processing begins. At some key point (yet to be identified for individual idioms but discussed briefly by Cacciari and Tabossi, 1988, and by Flores d'Arcais, 1993), the configuration emerges and any remaining words in the string may not be literally processed. If the point of idiom identification occurs early on in the expression, only the idiomatic meaning is likely to be available when the expression ends. If the idiom cannot be identified before the end of the string, then the literal meaning should become available before the idiomatic one is accessed. (As noted by Cacciari and Tabossi (1988) and Flores d'Arcais (1993), this "point of idiom uniqueness" is not fixed for any given idiom but may vary in accordance with context and familiarity. For example, the meaning of the idiom "spill the beans" is available after "spill" if the prior context introduces the theme of secrets. Without such a context, the meaning may only become available after the last word.)

Cacciari and Tabossi acknowledge that this model is "incomplete and tentative" (1988, 680) but it is a significant contribution to processing theory in that it is able to account for their on-line data (discussed above). Furthermore, it addresses the processing of idioms with a literal counterpart as well as those without one (both being processed in the same way). It is also economic in that each word is represented only once in the lexicon and need not be marked "literal" or "idiomatic". Unlike Gibbs et al.'s decomposition hypothesis (1989b), it is, however, unable to account for how language users know which syntactic transformations are permitted for a given idiom. By what mechanism does the passive form of an idiom like "kick the bucket" prevent its being recognised as an idiomatic configuration while the passive form of "lay down the law" allows this configuration to be accepted as idiomatic?

1 For some British speakers, "take the book" may be idiomatic. Cacciari and Tabossi do, however, indicate that their use of the sentence, "The boy took the book", is intended to illustrate literal usage of the verb, "take".

34
Possibly the most important aspect of Cacciari and Tabossi's model is that it breaks away from the assumption that idiomatic processing is distinct from literal processing. Levorato (1993) suggests that there are many reasons to reject the assumption that there are two separate processes. Firstly, such a separation goes against the principle of economy on which, she claims, cognitive functioning is based. Secondly, differentiated processing would require a mechanism to determine when and whether an expression should undergo normal processing (literal sense) or whether it requires special processing. Thirdly, the high proportion of figurative language present in normal conversation would mean that the special processing would operate almost as often as the "normal" processing. Fourthly, non-literal language can take a variety of forms ranging from expressions that are clearly figurative (like proverbs, metaphors, idioms, and similes) to expressions in which the discrepancy between what is said and what is meant is more subtle (e.g., irony and indirect speech acts). Finally, the literal-figurative relationship is most appropriately viewed not in terms of discrete categories but in terms of a continuum.

While the idea of a single overarching process explaining literal and idiomatic comprehension is appealing, it is clear that more research is required in order to address the shortcomings of the available hypotheses.

In addition, processing theory needs to address the issue pertaining to the encoding, storage and retrieval of pragmatic information associated with individual idioms. Idioms vary with respect to their appropriacy in specific social contexts, for example. Thus, the sentence "I hear John kicked the bucket" would be construed differently depending on whether the addressee was John's widow or John's rival. Similarly, it might be appropriate to use this idiom in the pub but not in a court of law. The literature does not, at present, account for how such features such as [-respectful] and [-formal] might be encoded, stored and retrieved alongside idiomatic expressions.
Furthermore, the theories reviewed here have been developed in the context of adult studies. It is not at all clear that any conclusions regarding adult processing can be extrapolated to children. Indeed there is reason to believe that the type of factors which influence processing and the manner in which they do so may well differ for children as opposed to adults (Levorato, 1993; Nippold & Rudzinski, 1993). To date, there is a notable dearth of literature regarding the processes involved in children's storage and retrieval of idiomatic meaning.

5. IDIOM COMPREHENSION IN NORMALLY-DEVELOPING CHILDREN

Studies of idiom comprehension in children have concentrated predominantly on the development of idiom comprehension and the factors impacting on it. Despite a convergence of objective, the literature is in fact somewhat disparate. Individual studies are difficult to compare since they tend to use different sets of idiom data and employ a range of methods and materials. Most authors have investigated the effects of age, but studies vary greatly with respect to which additional factors have been systematically controlled. Since age-related findings cannot validly be generalized beyond the idioms and conditions of individual studies, no clear age-related norms have emerged.

Despite the variation however, studies have largely concurred on the general trend of normal idiom comprehension development. It appears that literal interpretations of idioms predominate in early childhood, giving way to a greater proportion of figurative interpretations through a protracted process which continues into late adolescence or even adulthood (Ackerman, 1982; Brasseur & Jiminez, 1989; Brinton, Fujiki & Mackey, 1985; Cacciari & Levorato, 1989; Douglas & Peel, 1979; Gibbs, 1987; Johnson, 1985; Levorato & Cacciari, 1992; Lodge & Leach, 1975; Nippold & Martin, 1989; Nippold & Rudzinski, 1993; Nippold & Taylor, 1995; Prinz, 1983; Strand & Fraser, 1979; Thorum, 1986; Wiig & Secord, 1989).
Indeed, Nippold (1991) concedes, "there seems to be no clear point in human development when it can be said that idioms have been mastered" (p. 101), although the evidence suggests that a plateau of development may be attained around the age of eleven (Johnson, 1985; Strand & Fraser, 1979).

The youngest group of children as yet studied ranged in age from 3;6 to 6;6 (Abkarian, Jones & West, 1992). Ten idioms were presented in isolation or at the end of a three-line story. Comprehension was judged on a pictorial multiple-choice measure and on subsequent rationale-for-choice statements. The multiple choice options included idiomatic, literal, verb-maintaining and object-maintaining pictures. The rationale-for-choice statements were scored in terms of four categories: correct idiomatic explanations, explanations that were not entirely correct readings of the idiom but demonstrated some figurative awareness, literal explanations, and variants.

The results indicated that children aged 3;6 to 6;6 were already aware of the non-literal nature of some idiomatic expressions. Surprisingly, the multiple choice task revealed a significant linear trend for idioms to be interpreted more literally with increasing age. In contrast, the rationale-for-choice task yielded an increase in idiomatic and figurative-but-incorrect answers as a function of increased age. A pretest confirmed that these results were not due to response bias. The authors acknowledge, however, that literal responses may reflect children's attempts to choose the least risky option when faced with "what they thought to be the bizarre requests of an adult examiner" (p. 585). In support of this hypothesis, the authors cite Johnson's (1985) finding that children aged between 5 and 9 displayed a high level of literal responses to idioms presented in neutral contexts. Johnson's result may, of course, also be argued to reflect an automatic tendency in young children to use a literal strategy rather than a figurative one whenever this is not prohibited by the context. Unfortunately, no other studies known to this author have investigated idiom comprehension in
children under 5 years of age. Until such time as further evidence emerges, it is necessary to consider Abkarian et al.'s findings cautiously.

An important feature of Abkarian et al.'s study regards the finding of significant between-idioms differences. Correct responses (across all subjects) to individual idioms ranged from 2% ("butterflies in my stomach") to 49% ("caught a cold"). Similar levels of between-idioms variability are reported in two studies of children aged 5 to 11 (Brinton et al., 1985; Strand & Fraser, 1979) and in a study of normally-developing and "mentally retarded" children (Ezell & Goldstein, 1991). Unfortunately, direct comparisons across studies cannot be made since there was little overlap in the idioms included in each. Brinton et al. (1985) suggest that such discrepancies between idioms may reflect differential salience, frequency of exposure and ease of production. Performance on individual items may also relate to the extent to which the context is biased towards an idiomatic rather than literal interpretation and to the relative likelihood of a phrase being used literally rather than figuratively. An expression like "lend me a hand", for example, is more plausible in its idiomatic than literal sense. This characteristic might sway uncertain subjects to select the idiomatic interpretation. Having included a range of different idiom types in their materials, Strand and Fraser (1979) were unable to account for their large between-idioms variation in terms of linguistic structure. The authors conclude that idioms used frequently with children and useful to them are acquired early on whereas other, less frequently used and less useful to the child are learned later. A further explanation of between-idioms variation is offered by Gibbs (1987) who found that 5- to 9-year old children had a better grasp of syntactically frozen idioms than of those judged to be more flexible. Metaphorically transparent idioms were also better understood than were metaphorically opaque ones (assessed on an explanation and a forced-choice task).
High levels of between-idioms variability clearly have important implications for the validity and generalizability of studies in this field. This is highlighted by the differential outcomes of two early studies on idiom comprehension. Lodge and Leach (1975) assessed the comprehension of children aged 6, 9 and 12 on a pictorial multiple-choice test. For each of ten idioms, subjects selected the two out of four pictures which meant the same as the idiom string. The options included an idiomatic representation (for example, for "kick the bucket" it showed a boy lying on his back with a flower between his hands), a literal representation (a boy kicking a bucket), an idiomatic variation (a boy jumping), and a literal variation (a boy throwing a bucket). Having found that the youngest children chose very few idiomatic pictures and many literal or literal-variation pictures, the authors conclude that this group lacked the capacity for semantic duality.

A subsequent study conducted by Strand and Fraser (1979) reported considerably better performance among young children. Following a pictorial multiple choice task, the subjects were required to explain their choices. Judged on their explanations, even 5-year olds were consistently found to understand some idiomatic meanings. Examination of the idioms investigated by both studies reveals an overlap of five idioms. All of these were among those least often understood by Strand and Fraser's 5-year olds ("hit the sack", "spilled the beans"; "faced the music", "kick the bucket"; "broke the ice"). The idioms that were best understood by the 5-year olds ("feeling blue", "got ripped off", "caught redhanded", "cracked up") did not feature in Lodge and Leach's study. Thus, while there is at least some consistency across the studies, it would clearly be unwise to generalize from the overall results of either.

The discrepancy in the two sets of results highlights a further limitation on generalization from idiom research: the effect of task differences. In Lodge and Leach's study, pictures representing the idiomatic and literal meanings were presented simultaneously. The children were thus required to select the
idiomatic meaning over the literal one. Strand and Fraser (1979) considered simultaneous presentation to be confusing for young children. They thus opted to present one set of picture options for the figurative meaning and then another set for the literal meaning. In both instances, the subjects were instructed to choose the picture that best expressed the meaning from a set of four. Comments made by 5- and 7-year old children of the sort "I knew you were hiding the real pictures" and "There's the real 'hit the sack'" lead the authors to acknowledge that many may have selected the literal pictures if these had been presented alongside the idiomatic ones.1

The impact of task differences is even more vivid in studies which use two different tasks to assess the same children on the same idioms. Ackerman (1982) investigated idiom comprehension in 6-, 8- and 10-year old children using first an explanation task and then a task requiring the subjects to respond "yes" or "no" to statements about the 20 idioms. (The findings of this study as regards context and form are discussed below.) On the explanation task, 6-year olds attained a 30% correct response rate in what was found to be the most favourable condition (idiom in conventional form within an idiomatic context). With the same combination of conditions, the same children achieved 80% accuracy on the "yes"/"no" task. Even in the less favourable conditions (neutral or literal contexts and idioms being presented in modified form), the "yes"/"no" task yielded accuracy rates of between 33% and 53% for 6-year olds.

Combining the procedures of Ackerman (1982) and Strand and Fraser (1979), Johnson (1985) used Ackerman's idioms in the same contexts but, like Strand and Fraser, presented one set of multiple choice pictures for the literal meaning and another for the idiomatic meaning. Thus, the children were not required to choose between a literal and figurative depiction of the given expression. Both sets of pictures were presented for the neutral

1 What is further revealed by these comments is that these young children knew about literality and were able to differentiate between literal and non-literal meanings.
(ambiguous) use of the idioms. Johnson reports that 48% of the 5-year olds' selections in the idiomatic context were correct.

The comparison of this finding with those of Ackerman (1982) and Strand and Fraser (1979) indicates that the nature of both task and idioms investigated is likely to affect the outcome of comprehension research. The impact of these factors on outcome is further evidenced by other child idiom comprehension studies.

Ezell and Goldstein (1991) compared Ackerman's results with those for their own group of 6-year olds, tested on a pictorial multiple-choice task. In this study, the children selected from literal, idiomatic, literal variation and idiomatic variation pictures. Only 9% of the 6-year olds' selections were correct. The literal option was chosen in 70% of cases. Ezell and Goldstein suggest that in Ackerman's forced-choice task, the binary selection (may have) provided a more favorable guess rate for these children. It is also likely that hearing the idiomatic meaning stated as a choice was a more salient cue than were the pictures used in (their) study (p.817).

Another multiple choice task, this time requiring subjects to select from three written statements representing idiomatic, literal and associated meanings, was conducted by Cacciari and Levorato (1989). Their subjects were 7- and 9-year old Italian children. It was found that over 50% of the 7-year olds' responses were idiomatic even when the context allowed a literal interpretation (although the idiomatic meaning was still more appropriate). (The findings relating to context variables are discussed below.) While the discrepancy in findings between Ezell and Goldstein (1991) and Cacciari and Levorato (1989) may be partially explained by the one-year age gap and the use of different idioms, it is nevertheless likely that the choice of pictorial over written statements also played a role.

In justifying their own methodology, Cacciari and Levorato (1989) propose that the poor performance of Ackerman's 6-year olds on the explanation task may reflect the inappropriacy of this task
for this age group. The authors assert that an explanation task is too difficult, cognitively and linguistically, for young children. The relative difficulty of explanation tasks even for older children and adolescents is highlighted by a number of other studies. Prinz (1983) tested children ranging in age from 6 to 15 on a pictorial multiple-choice task and a verbal explanation task. The author reports a dramatic increase in idiom comprehension on the multiple-choice task between age 6 (10%) and age 9 (50%), and again between age 9 and age 12 (80%). Accuracy on the explanation task followed a parallel linear development but all groups performed significantly better on the multiple-choice than on the explanation task. Gibbs (1987) similarly reported an age-related increase in correct idiom explanations and choices (on a literal/idiomatic forced-choice task) in children aged between 5 and 9 years. Once again, the level of correct explanations was consistently lower than the level of correct choices.

The same pattern of discrepancy in older children and adolescents is reported in a set of papers by Nippold and her colleagues (Nippold & Rudzinski, 1993; Nippold & Taylor, 1995). The authors compare the results of two studies using the same idioms embedded in the same stories with subjects of the same age groups (Grades 5: 10; 5-11; 9, Grade 8: 13; 3-14; 11, Grade 11: 16; 5-17; 10). In the first study (Nippold & Rudzinski, 1993), subjects were required to provide a written explanation for each idiom. In the second study (Nippold & Taylor, 1995), subjects selected the best explanation from the four provided. The foils included explanations related to the story rather than highly implausible or literal choices. The results revealed significantly greater accuracy on the forced-choice task than on the explanation task. The authors conclude that while both tasks are metalinguistic in nature, the explanation task is more demanding in that it requires the subjects to generate an appropriate interpretation.

The relative difficulty of tasks requiring the subjects to generate rather than select a response is further evidenced by
another experiment described in Levorato and Cacciari's 1992 report. The authors compared the performance of 8- and 10-year old children on an idiom completion test and a multiple-choice test. They found a higher incidence of idiomatic answers on the multiple-choice task than on the idiom completion task across age groups.

While task-related discrepancies limit valid generalization from any one study, they do, nevertheless, yield valuable information. It appears that children at 6 or 7 years of age have difficulty actively generating definitions of idioms but demonstrate considerable passive comprehension of idioms (Ackerman, 1982; Cacciari and Levorato, 1989; Gibbs, 1987; Johnson, 1985; Levorato & Cacciari, 1992). Thus, by around 6 years of age (and to a lesser degree by the age of 5), children demonstrate a sensitivity to non-literal language even though they may as yet be unable to define many idioms.

With only one study covering the whole range 3;5 to 6;5 (Abkarian et al., 1992), it is not possible to make any firm claims with regard to when this transitional phase begins. In view of the great variation in tasks and controls that characterise this literature, it would be unwise to compare results across studies. However, further support for a transitional period occurring somewhere between 5 and 7 years of age does emerge from studies which code responses that are neither clearly literal nor clearly idiomatic.

The explanations produced by Strand and Fraser's (1979) subjects were coded in terms of the following categories:--
(1) conventional idiomatic meaning
(2) nonliteral but incorrect explanations:--
(a) nonliteral explanation by means of analogy (usually to another idiom)
(b) nonliteral explanation in the same context domain
(c) nonliteral explanation not obviously in the same context domain
(3) complete or partial conventional idiomatic explanation and additional literal explanation
(4) complete literal explanation
(5) other explanation

The groups of 5- and 7-year olds in this study both gave literal (type 4) explanations more often than any other single type. However, when each group's non-literal strategies were combined (i.e. types 1, 2 and 3), the total exceeded the number of literal explanations for each group. The authors conclude that even 5-year olds employ non-literal methods (albeit immature ones) for understanding idioms more than 50% of the time.

Cacciari and Levorato's (1989, 1992) multiple-choice options for each idiom related to the idiomatic, literal and an associate meaning. The latter was plausible within the context but was neither the literal interpretation nor the idiomatic one. Not surprisingly, the 7-year olds selected fewer idiomatic choices than did the 9-year olds. The 7-year olds were, however, found to select more associate answers than did the older children. Interference from context and familiarity were ruled out. The authors conclude that the children who selected associate answers were those who were beginning to employ a figurative strategy but were not yet able to analyze the text in a manner that would lead them to the most appropriate interpretation. Although able to reject the literal meaning, they were not yet consistently able to identify the most appropriate idiomatic meaning.

The same pattern of results emerged from Cacciari and Levorato's (1989) experiment employing an idiom completion task. The words filled in by the children were categorised as "idiomatic", "literal" or "figurative". "Figurative" answers were those in which the child revealed an understanding of the idiomatic meaning but did not know the exact idiom that expressed that meaning (e.g. "broke the fear/problem" for "break the ice" in the context of a boy dealing with solitude in a new place). Once again it was found that older children (aged 10) produced more
idiomatic completions (40%) than did younger 8-year olds (13.9%). Nevertheless, 39.2% of the younger children's completions were "figurative". In these, the child reveals an understanding of the meaning of the action or mental state of the character to which the idiomatic expression refers, and yet does not know the exact idiom that expresses this state.... ("Figurative completions") show that children perceive the figurative potentialities of language and prefer to try to exploit these figurative ways of expression than to use the literal completions orienting the story towards a concrete end (p. 399).

Thus, at least by the age of 7, there is evidence for an emergent idiomatic strategy. Cacciari and Levorato suggest that this represents a transitional phase marked by the ability to apply a non-literal strategy but hindered by under-developed linguistic processing skills. This results in the child frequently failing to arrive at the most coherent interpretation of idioms in context.

As observed above, it is not currently clear when this transition begins. The only idiom study that has, as yet, included children below the age of 5, has reported a steady increase in "figurative-but-incorrect" rationale statements between the ages of 3 and 6 (Abkarian et al., 1992). The pattern that thus emerges is one of predominant literality in very early childhood, followed by a gradual shift into a transitional phase. This phase is marked by some sensitivity to non-literal language in the absence of substantial semantic knowledge relating to individual idioms. Gradually, idiomatic competence and experience of specific idioms increase, continuing to develop well into adolescence (Nippold & Martin, 1989).

To the extent that it accepts that there is a shift from a literal strategy to an idiomatic strategy, the literature reveals widespread consensus. Differences of opinion emerge, however, with regard to explaining how and why this shift should occur in children of approximately 6 years of age.
Lodge and Leach (1975) propose that the lack of early idiomatic comprehension may be the result of adults avoiding idioms in their interactions with young children. The authors reject this hypothesis, however, on the grounds that children are also exposed to considerable adult-to-adult interaction. Indeed, young children are further likely to encounter idioms in the language of older children and of television.

An alternative explanation offered by Lodge and Leach (1975) is that idiomatic meanings often lack concrete referents. The largely abstract nature of idioms makes them difficult for young children to understand. The authors argue further that the processing of a string as a whole may be more difficult than is a sum-of-the-parts interpretation of the same string.

The accounts considered by Lodge and Leach are consistent with the "acquisition via exposure" hypothesis. According to this proposal, idioms are learned as "giant" lexical units and stored in semantic memory as separate entities, in the same way as are single words (Ackerman, 1982; Strand & Fraser, 1979). As with words, idioms may be learned through exposure either to explicit explanations or by a process of "contextual abstraction" (Werner & Kaplan, 1950 cited in Nippold & Martin, 1989), whereby the child gradually infers the meaning of unfamiliar idioms through repeated exposure to the expression in context. As listening and reading skills improve, so too does the child's understanding of idioms (Nippold, 1991).

Lending a small measure of support to the view that idioms are acquired as giant lexical items, Gibbs (1987) found that syntactically frozen idioms were better understood than flexible ones by 5- and 6-year old children. The degree of frozenness appeared to make no significant impact on children aged 8 and 9 years, however. Given that frozen idioms are heard in only one (or very few) syntactic forms, these strings may be more easily recognised as single units than are flexible idioms heard in many different forms. In view of the interactive effect of context
(discussed below), it is however difficult to draw any clear conclusions from this result. It is also notable that Lodge and Leach (1975) report that the degree of syntactic frozenness had no significant effect on the comprehension of their 6-, 9- and 12-year olds, despite idioms having been better understood in the active than passive voice.

The "acquisition via exposure" viewpoint finds clearer support in Ezell and Goldstein's (1991) study of normally-developing and "mentally retarded" children. The authors report that the 9-year old "mentally retarded" children achieved 17% accuracy on a pictorial multiple-choice test. Although matched on receptive vocabulary age (Peabody Picture Vocabulary Test, Dunn, 1965), normal 6-year olds scored only 9% accuracy. The authors speculate that the discrepancy may result from the "mentally retarded" children being three years older and thus having had more exposure to idioms. While differential exposure may be implicated in this result, it should be borne in mind that groups matched on receptive vocabulary age may nevertheless differ on a range of other linguistic and cognitive indices (see, for example, McTear and Conti-Ramsden (1992) and Plante, Swisher, Kiernan and Restrepo (1993) for a discussion of the issues pertaining to language matching).

Ezell and Goldstein's (1991) study does, however, yield another result which supports the "acquisition via exposure" hypothesis. The idioms interpreted accurately by 6- and 9-year olds in this study were also among those judged to be frequently-occurring idioms by the college students surveyed by Ezell & Goldstein.

The influence of familiarity (frequency of exposure) on children's idiom comprehension has been investigated by Levorato and Cacciari (1992). In this series of experiments, familiarity was not found to be a significant factor when idioms were presented to 7- and 9-year olds in an idiomatic context. When idioms were presented in a context which permitted a literal interpretation (despite the idiomatic one being more
appropriate), literal responses were more common for unfamiliar idioms than for familiar ones. The authors conclude that familiarity impacts only on children who are not yet able to use contextual information fully. Interestingly, the same authors revealed a significant effect for familiarity on a subsequent production (idiom completion) task (also reported in Levorato & Cacciari, 1992).

Findings in support of the relevance of familiarity were also made by Nippold & Rudzinski (1993) and Nippold & Taylor (1995) who point out that Levorato and Cacciari (1992) based their familiarity ratings on adult (teacher) judgements. Indeed, this is also true of Ezell and Goldstein (1991). Using adolescent familiarity ratings, Nippold and her colleagues found that familiarity played a significant role in facilitating comprehension in children aged 11, 13 and 17. (Although not examining the effect of familiarity, Strand and Fraser (1979) used only idioms judged familiar by at least 50% of a group of 12-year olds. Interestingly, these authors reported higher levels of idiom comprehension in young children than several other studies which did not use child familiarity ratings.)

While Nippold's familiarity-related finding supports the "acquisition via exposure" hypothesis, another of Nippold's findings suggests that idioms are not all acquired as giant lexical units. In addition to controlling for familiarity, Nippold & Rudzinski (1993) and Nippold & Taylor (1995) also controlled for relative metaphoric transparency of idioms as judged by adolescents. Using an explanation task with children in Grades 5, 8 and 11, the authors found that metaphoric transparency was moderately associated with idiom explanation for the 8th and 11th grade students, but not for the younger 5th grade students (mean age 11;2). When subjects of the same age ranges were tested on a forced choice task involving the same idioms, significant correlations between idiom comprehension and idiom transparency were found for all three groups. In all cases, greater accuracy was associated with greater transparency.
These results are consistent with Gibbs' (1987) finding that children from kindergarten to 4th grade (ages 5 to 9) were better at explaining metaphorically transparent idioms than opaque ones. They are also partially consistent with Cacciari (1993). In a study including the investigation of Italian 10-year olds' ability to paraphrase idioms, Cacciari controlled for three degrees of transparency: semantically transparent idioms (e.g. "cry over spilled milk"), quasi-metaphorical idioms based on a figurative comparison (e.g. "to be as two drops of water" = "to be very similar"), and opaque idioms (e.g. "break the ice"). The results revealed the highest level of accuracy for quasi-metaphorical idioms (69.9% correct), followed by opaque idioms (51.9% accuracy) and transparent idioms (47.9% accuracy).

In considering these results, it should be noted that Cacciari's study used Italian idioms. Without further investigation, one cannot rule out the possibility of differential frequency associated with the opaque and transparent idioms, respectively, in the two languages. More generally, it is notable that three out of the four experiments cited with regard to the effect of metaphoric transparency involved paraphrase tasks. It might be argued that certain types of idioms are easier to explain than others but that it does not necessarily follow that these are better understood. Significantly, however, the relationship between metaphoric transparency and comprehension was also borne out in the forced choice tasks reported by Gibbs (1987) and Nippold and Taylor (1995). This finding is interpreted by Nippold and Taylor as evidence that the child learner tries to infer the non-literal meaning from the literal meaning of an expression. The authors conclude that while opaque idioms may be learned as giant lexical units, metaphorically transparent idioms seem to be learned through a strategy of dissection. Gibbs, similarly, suggests that different acquisition strategies apply to different types of idioms: opaque and ill-formed idioms are probably acquired by rote whereas transparent idioms are more likely to be worked out on the basis of some metaphorical reasoning strategy.
While the authors do not draw the connection, their hypothesis is itself dependent on another view of idiom acquisition outlined variously by Ortony et al. (1978) and by Cacciari & Levorato (1989). According to this view, idiom comprehension is closely linked to the recognition of contextual anomaly. When the child encounters an expression whose literal sense is incongruent with the context, s/he is alerted to search for a meaning that does make sense. In describing their "global elaboration hypothesis", Levorato and Cacciari (1992) observe that context not only makes it possible to suspend the literal interpretation of the idiom, but it also gives the semantic information necessary in order to assign coherence to the text and to extract the figurative sense of the idiom; in other words, to go beyond the local piece of information and reach the global sense of the text (p.417).

While the "acquisition via exposure" hypothesis and the "global elaboration hypothesis" approach the subject from different perspectives, they are not necessarily mutually exclusive. In fact, Ortony et al. (1978) accept the "acquisition via exposure" argument, but propose that once an idiom is learned, the context becomes a key factor in determining whether a literal or figurative interpretation is appropriate.

The importance of context as a facilitator of accurate idiomatic comprehension has been extensively studied. Two branches of research can be identified. The first relates to the effect of idioms being presented in isolation as opposed to being presented in an idiomatic context. The second compares the differential effect on comprehension of idiomatic, literal and ambiguous contexts.

For the most part, discussion of the impact of presentation in context versus isolation derives from comparisons of separate studies. Since these also use different sets of idioms and different methodologies, the conclusions that can be drawn from such comparisons are limited. Ezell and Goldstein (1991) found that 9-year olds achieved 65% accuracy on a pictorial multiple choice task when the idioms were presented in idiomatic contexts.
As noted by the authors, this finding is consistent with an accuracy range of 56% to 96% reported for 9-year olds in other studies which have used a supporting context (Brinton et al., 1985; Cacciari and Levorato, 1989; Gibbs, 1987; Levorato & Cacciari, 1992). The authors also note that their 9-year olds performed considerably better than those in studies which presented idioms in isolation (Lodge & Leach, 1975: 20%; Prinz, 1983: 50%). The authors conclude that while the variability may be due to differences in the idioms used, this result nevertheless indicates that supporting contexts facilitate idiom comprehension in this age group. It is interesting that Ezell and Goldstein neglect to cite the contrasting results of Strand and Fraser (1979). Despite presenting idioms in isolation, these authors found that 9-year olds achieved 62% accuracy in their experiment. Ezell and Goldstein further fail to note that their 6-year olds actually performed less accurately in response to idioms in a supporting context than did the 6-year olds in both Lodge and Leach (1975) and Prinz (1983). All of these studies employed pictorial multiple choice tasks. Once again, however, differences in methodology and idioms prohibit the drawing of any clear conclusions.

A more systematic approach to the issue of context was taken by Gibbs (1987) who presented his sets of frozen/flexible and transparent/opaque idioms either in isolation or in a short story context to children aged 5 to 9 years. On the explanation task, many more literal explanations were found in the isolation condition than in the context condition. Interestingly, however, a third of the children who gave literal explanations in this condition went on to choose idiomatic interpretations for the same idioms on the forced-choice task. The number of idiomatic choices was still greater when the idioms were presented in context however. The author observes further that the extent to which context enhanced comprehension varied with the nature of the idioms. On the explanation task, context facilitated comprehension most for syntactically frozen idioms and metaphorically transparent idioms. On the forced-choice task,
the greatest degree of contextual facilitation was associated with frozen idioms.

In their study of Italian children, Cacciari and Levorato (1989) also compared the comprehension of idioms embedded in short narratives with that of the same idioms presented in isolation. The subjects selected the meaning of each of eight idioms from three written statements representing the idiomatic meaning, the literal meaning and an associate meaning. Idioms in context were presented within either an idiomatic-biasing context or a literal-biasing one. The authors found that 7- and 9-year old Italian children performed significantly better when idioms were presented in context rather than in isolation. This effect was evident irrespective of contextual bias. (The impact of context bias is discussed below.)

Further support for the facilitative effect of context derives from an adolescent study (ages 14-17) conducted by Nippold and Martin (1989). Comprehension of twenty idioms was tested using a modified version of the Fullerton Language Test for Adolescents idiom subtest (Thorum, 1980, 1986, discussed below). The subjects listened to an explanation of idioms and were then instructed to write down only the non-literal interpretations of the test items. Half of the items were presented in two-sentence contexts and half in isolation. Accuracy was found to be greater when idioms were presented in context. The difference, however, was relatively small. The authors suggest that a fuller context might have yielded a greater difference.

The differential impact of more or less context on idiom comprehension has not as yet been investigated experimentally in children. Findings from an adult metaphor comprehension study do however indicate a facilitative effect for long contexts (33-60 words) as opposed to short ones (3-11 words) (Ortony et al., 1978). Idiom comprehension researchers have, to date, limited the context into which idioms are embedded to four sentences and each idiom has been presented within a separate context. It may
well be that contrived experimental contexts do not function in the same way as natural conversation. Levorato & Cacciari (1989) suggest that to qualify as ecologically valid, "the experimental setting must be as informative as is the context in everyday life" (p. 390).

Findings from studies investigating contextual bias have furthermore confirmed that idiomatic interpretations are facilitated by a context which promotes the expectation of a figurative interpretation. Johnson (1985) found that 5-, 7-, 9-, 11- and 14-year olds consistently made more idiomatic selections on a pictorial multiple choice task when the idioms were presented in idiomatic, rather than neutral, contexts. Levorato and Cacciari (1992) presented idioms in either idiomatic-inducing narratives or in literal-inducing narratives to 7- and 9-year olds. (In the literal-inducing contexts, a literal interpretation was permissible but the idiomatic interpretation was still more appropriate.) They also varied the degree of familiarity of idioms as judged by teachers. On a multiple-choice task, the children selected the idiomatic interpretation, the literal interpretation or an associate answer (plausible in the context but different from the idiomatic and literal meanings). This methodology is illustrated in the following example pertaining to the idiom "caught between two fires" (p.420).

A. LITERAL-INDUCING CONTEXT

A bandit has betrayed his gang, who are looking for him in order to get revenge. All of a sudden the bandit realizes that both the sheriff and one of his fellow gang members are chasing him. He is afraid he will not be able to get away and so feels he is caught between two fires.

Why does the bandit feel as if he is caught between two fires?

1. Because he has to protect himself from the firing of two guns. (literal)
2. Because he does not know what to do. (idiomatic)
3. Because he loves his horse. (associate)
B. IDIOMATIC-INDUCING CONTEXT

Joan steals some apples from a neighbour's tree. Her brother Eric goes to tell their mother, who gets very angry and puts her in a corner. Now Eric is very sorry that Joan has been punished but at the same time he thinks that stealing is not fair. This is why he feels as if he is caught between two fires.

Why does Eric feel as if he is caught between two fires?

1. Because he has to protect himself from the fire. (literal)
2. Because he is in a very difficult situation. (idiomatic)
3. Because he wants to be punished instead of Joan (associate)

In this study, Levorato and Cacciari found that an idiomatic context facilitated idiomatic answers relative to a literal context. The capacity to use context to facilitate correct interpretation appeared to increase with age.

The impact of contextual bias on Levorato and Cacciari's 7- and 9-year olds was partially supported by a study of similarly-aged English-speaking children (Ackerman, 1982). Ackerman investigated performance in 6-, 8- and 10-year old children and a group of adults, who were required first to explain the meanings of idioms and then to respond "yes" or "no" to statements about the idioms. Twenty idioms were presented in contexts that biased a literal, an idiomatic or a neutral interpretation. In addition, half of the idioms were presented in their conventional form (presented first in the example below) while the other half were presented in a modified form (presented second in the example below). The idiom, "throw in the towel", for example, was presented in the following ways (p.444):

A. IDIOMATIC-BIASING CONTEXT

David's team was way behind. The coach called time out and began to send in substitutions. David said the coach was throwing in the towel/handing in the towel.

B. NEUTRAL CONTEXT

David was not doing very well. He was tired and sweaty. The coach threw in the towel/handed in the towel.
C. LITERAL-BIASING CONTEXT

David was not doing very well in the match. He couldn't see because the sweat kept coming into his eyes. The coach threw in the towel/handed in the towel.

QUESTIONS

1. What was the coach trying to do?
2. Did the coach think David was going to lose? OR Did the coach want to keep trying to win?

Ackerman found that 6 year-old children were able to produce idiomatic interpretations only when idioms were presented in their conventional forms and in an idiomatic context. In any other combination, these children produced no idiomatic explanations at all. The 8-year olds achieved some correct answers in other conditions, but also performed significantly better in idiomatic-biasing contexts when idioms were presented in their conventional forms. The facilitative effect of idiomatically-biasing contexts was further evidenced in the second, "yes"/"no" task. For all but the 10-year olds, idiomatic contexts were associated with the highest level of idiomatic interpretations. Neutral contexts produced a lower number of idiomatic interpretations and literal contexts produced the fewest.

While the presence of an idiomatically-biasing context clearly facilitated younger children's comprehension, Ackerman found that the context type had considerably reduced impact on 10-year old children. Relative to 8-year olds, the older children produced a much greater proportion of idiomatic explanations in neutral and in literal contexts (neutral context: 10-year olds 62.5%, 8-year olds 27.8%; literal context: 10-year olds 41.7%, 8 year-olds 8.3%). The older children did not respond as accurately to modified forms of the idioms as they did to conventional forms, but they nevertheless still performed significantly better than 8-year olds in the modified idiom condition.

With reference to these results, Ackerman proposed that contextual cues, important in guiding the young child to an
interpretation decision, become less significant once knowledge of the specific form and typical use of idioms is established. At least with regard to this particular set of idioms, 10-year old children appeared to have achieved recognition of the idiom form, itself. Ackerman does not suggest how idiom forms come to be recognised, but others have proposed that children may be sensitive to a range of non-verbal cues including tone of voice, rate of speech and eye blinking (Ezell & Goldstein, 1991). Indeed, prosodic differences corresponding to literal and idiomatic utterances, respectively, have been reported in the literature (Van Lancker & Canter, 1981; Van Lancker, Canter and Terbeek, 1980 cited in Prinz, 1983).

Commenting on Ackerman's results, Nippold and Martin (1989) suggest that the familiarity of the particular idioms included may underlie the apparent lack of contextual influence in the 10-year olds. These children may have relied on context to a greater extent if the idioms had been less familiar.

Indeed idiom familiarity may have contributed to these results in another sense, not discussed in the literature. Apparently unintentionally, Ackerman's contextualising narratives include a number of idioms other than the target ones. In the case of "throw in the towel", exemplified above, three idiomatic expressions occur in addition to the target item ("was (way) behind", "call time out", "send in"). Other examples quoted in Ackerman's report similarly involve idioms other than the target expressions. It cannot be assumed that these apparently-overlooked idioms are familiar to any of the age groups studied, but it is likely that they would be better understood by the 10-year olds than by either of the other age groups. Whether or not this oversight accounts for the difference in contextual influence across the age groups, it nevertheless threatens the validity of Ackerman's results.

Interestingly, Johnson's study, using Ackerman's idioms and contexts, did not support Ackerman's (1982) findings with regard
to the reduction of contextual impact on older children. In the absence of further replicating research, it remains unclear whether or not there is a reduction in contextual influence around the age of 10 years. On the basis of his results, Gibbs (1987) does suggest however, that children do not automatically view such phrases as "sitting on pins and needles" and "spill the beans" as requiring idiomatic interpretations unless the context specifically demands it. He notes that this contrasts with the adult tendency to provide an idiomatic interpretation even when such phrases are presented in strongly literal contexts (Gibbs, 1980, 1986; Van Lancker & Canter, 1981).

Levorato (1993) provides an account of the different type of role likely to be played by context in adults and children. This distinction may also apply to young children as opposed to older children with more developed idiomatic skills. Levorato suggests that for the child who does not know a given idiom, context provides background information for an hypothesis about the idiom's meaning, it aids the memory search for information that could serve to define the figurative meaning, and it provides the material for the appropriate inferential process.... Whereas adults retrieve the idiomatic meaning first and only subsequently instantiate the interpretation of the idiom with the linguistic context (Kemper, 1986) or, even if they integrate the idiomatic meaning with the linguistic context parallel to the retrieval of the idiomatic meaning, the integration does not depend on the contextual information, children, on the other hand, analyze the figurative expression at the same time as the linguistic context and need the contextual information to grasp the idiomatic meaning (p. 105).

Interpreting Ackerman's results in terms of Levorato's discussion, it could be argued that the larger idiom repertoire and greater idiomatic experience of 10-year old children accounts for their lesser dependence on context than that found for 6- and 8-year olds. How do we account, however, for Abkarian et al.'s (1992) finding that the presence or absence of context had no significant effect on children aged 3 to 6? While not intended as a response to this question, the views of Ackerman (1982) and Cacciari and Levorato (1989, 1992) are informative in this regard. These authors suggest that young children (presumably
younger than 6 or 7) lack the inferencing and reasoning skills required to make use of context in the manner outlined by Levorato above. Thus, faced with an unfamiliar idiom, especially in a context which permits a literal interpretation, the young child is more likely to fall back on the literal meaning. In addition to inferencing skills, it is suggested that idiomatic understanding further requires the ability to recognise that words can be part of complex configurations (Cacciari & Levorato, 1989).

Indeed, Levorato and Cacciari strongly assert that figurative competence cannot be considered separately from other skills such as inferencing, activation of world knowledge, the ability to work out the communicative intent of speakers, and the activation of metalinguistic knowledge (Levorato, 1993). Instead, they consider these to be integrally-related processes and skills which are "ever more subject to a tendency ... to search for the greatest possible degree of coherence" (Levorato, 1993, 104). Levorato concedes that it is difficult to be sure which comes first: the search for coherence or the skills required for complex linguistic processing. Instead she hypothesizes that

the need to find coherence within one's own mental contents, or among the products of one's own processing, develops in parallel with the ability to carry out those analyses that allow for the creation of a coherent representation (p.107).

Expanding on an earlier theory (Levorato & Cacciari, 1992), Levorato (1993) outlines a 6-stage model in which the gradual development of figurative competence is viewed as being part of of broader linguistic and cognitive development. Beginning with "Level 0", Levorato outlines the very early "nominal realist phase". At this point, an object and its name are viewed by the child as being one and the same. The child believes that the name of an object is intrinsic to it (e.g. if the sun changed its name, it would lose its heat) and has the same properties (e.g. long objects must have long names). S/he has no concept of meaning.
Gradually, beginning at "Level 1", the child comes to recognise that the name is merely a symbol for the object and refers to a meaning. This understanding relates to the child's expanding categorization skills and to the awareness that the same label may be applied to a range of referents and any one referent may be associated with several different labels. At this point, the child still applies a literal strategy in the processing of discourse. S/he tends to conceive of meaning as the mere sum of the meanings of the component words and to be misled by the imaginative contents of figures of speech.

By the age of approximately 7 or 8, the child enters Levorato's "Level 2" or "suspended literalness phase". At this point, s/he begins to recognise the incongruity of the literal meaning in an idiomatic context. Thus s/he comes to appreciate the possibility of a discrepancy between an expression and its meaning. While this awareness is still rudimentary, it gradually enables the child to perceive that such a discrepancy is not merely an error on the speaker's part. This, in turn, motivates the child to activate inferential processes to re-establish coherence. Over time, the child's ability to use context increases, but at each developmental level, different types of contexts will facilitate the identification of non-literal meaning to a greater or lesser extent.

"Level 3" (attained at approximately 9 or 10 years) is characterized by the generalization of the arbitrary relationship between the signifier and signified. The child could be said to have acquired a kind of general rule that tells him or her not to depend too much on the surface form of a linguistic expression for meaning (Levorato, 1993, 121).

From this point on, the child begins to recognise that literal meaning is not always direct and literal. The child begins to understand that words and phrases can have multiple meanings depending on the context in which they are used.

---

1 Winner, Rosenstiel and Gardner (1976) illustrate a typical magical interpretation of metaphor in relation to the sentence, "After many years working at the jail, the prison guard had become a hard rock that could not be moved". "The king had a magic rock and he turned the guard into another rock" (p.293). Such magical interpretation by young children is not limited to non-literal language, however. Paley (1981) vividly illustrates the central role of magic in 5-year olds' interpretation of events.
expression is only one means of achieving a communicative end and that it is the speaker's intentions that are important. This development is related to the child's gradual move away from predominantly concrete thought. It is at this stage that the child's ability to understand and produce idioms, metaphors, similes, proverbs and other figurative language becomes evident.

While idioms to which the child is regularly exposed may be acquired before "Level 4", it is once this level is reached that idiom acquisition accelerates. Having acquired "a sufficient system of knowledge" (Levorato, 1993, 122), the child now relates conventionalized expressions to concepts already acquired. The acquisition of idioms now depends primarily on the extent to which they are encountered in appropriate contexts. At this point, idiomatic expressions are still understood and produced only as indivisible units, as formulas that are applied in a fixed way and cannot be modified according to circumstance.

In contrast, the child who has achieved complete figurative competence ("Level 5") is able to analyze the idiom and to reflect on the relationship between the meaning and the component words. The child can now work out meaning by making semantic inferences about the elements of the idiom and by referring to his or her background knowledge. Interpretation is no longer as reliant on context. Idioms can be understood even when they have been subjected to lexical substitution or syntactic and lexical variations. The child can also generate new idioms by means of syntactic and lexical variations on existing ones. Furthermore, figurative language may now be used for diverse purposes (e.g. an idiom may be used ironically or in an altered form for exaggeration, as in "burn a candle at 3 ends").

The strength of the hypothesis proposed by Levorato lies not only in its ability to account for the acquisition of idiomatic competence, but also in its consistency with much of the literature. Level 1 closely resembles the shift described by Chafe (1970 cited in Prinz, 1983) from "one-to-one correspondence
between communication symbols and meaning" to "semantic duality". Evidence for the emergence of semantic-duality at around 7 years has been reported in relation to other forms of dual meaning. Asch and Nerlove (1960) traced the development of double function terms like "hard", "soft" and "sweet". Whereas up to around the age of 6, children understood these terms only in relation to physical attributes, children aged between 7 and 8 years understood the psychological meanings of such terms in over half of the instances. Similarly, the emergence of an appreciation of ambiguity and double meaning in poems, jokes and riddles has been noted in children of this age (Lodge and Leach, 1975; McGhee, 1979, 1983, 1989; Prinz, 1983).

The conceptualisation of idiom competence as part of a broader language competence driven by a search for coherence is, furthermore, in line with much current thinking with respect to language comprehension (see, for example, Sperber and Wilson's, 1986, relevance theory).

Levorato's model, in contrast to the "acquisition via exposure" hypothesis, has the further advantage of being able to incorporate any number of strategies for the acquisition of individual idioms. As noted above, it is likely that different types of idioms are acquired and processed by means of different strategies or combinations of strategies.

Interestingly, however, Levorato's model suggests that only very high-frequency idioms are acquired prior to "Level 4". Yet, as illustrated above, many studies of children of 10 years or younger have demonstrated considerable idiom knowledge among these children.

A further inconsistency with previous research findings regards the point at which children become able to analyze idioms. On the basis of his studies reviewed above, Gibbs (1987, 1991) concludes that decomposable idioms are learned earlier than nondecomposable ones. This suggests that kindergarten children are already acquiring idioms through componential analysis before
they are able to acquire them as lexical units. In contrast, Levorato's model suggests that idioms are first learned as indivisible units and only later does the ability to analyze parts of expressions emerge.

Clearly, as in the field of adult processing, further research is still required to address the inconsistencies and gaps in the literature. At present, the picture that emerges is one of developmental progression, in line with wider cognitive and linguistic development, from a predominantly literal strategy to relatively full idiomatic competence by late adolescence. (As observed by Strand & Fraser, 1979, however, children of all ages may switch between literal and non-literal strategies when presented with different idioms.) At different stages in the child's acquisition of figurative competence, different degrees of influence are probably exerted by such factors as context, idiom familiarity, idiom transparency, and the likelihood of a literal interpretation. Further research is still needed to clarify developmental age norms for different types of idioms, to investigate the manner in which different idiom characteristics impact on idiom acquisition and processing, and to explore the mechanism of idiom storage and retrieval in children. Research is also required to examine the process through which layers of meaning are built up around core information about an idiom (Levorato, 1993; Wittgenstein, 1958, 1970 cited in Johnson, 1985). In addition to furthering our understanding of normally-developing children, such research is a prerequisite for increased insight into the problems of children with semantic-pragmatic difficulties.

6. ANECDOTAL ACCOUNTS OF LITERALITY IN CHILDREN WITH SEMANTIC-PRAGMATIC DIFFICULTIES

The literature on semantic-pragmatic difficulties highlights a tendency among such children to over-literality at several
grammatical levels. At the word level, over-literal interpretations may manifest in rigid concept boundaries (Hyde-Wright & Cray, 1991; McTear, 1991; Smedley, 1989). McTear (1991) cites the example of a child who said "the clock is by the wall" because he believed that the preposition "on" could only be used to refer to locations on the horizontal plane. Here the child over-restricts the use of the preposition to its most literal meaning. Hyde-Wright & Cray (1991) describe a child who insisted that football was a sport and could not therefore be added to his list of games which included tennis, cricket and rounders.

While resembling the semantic underextensions common in early language development (Aitchison, 1987; Clark, 1973; Dromi, 1987; Ingram, 1989), the over-restriction of concepts displayed by children with semantic-pragmatic difficulties is characterised by its persistence well beyond the norm and by its association with parts of speech other than nouns. Indeed, in normally-developing children, semantic underextension (e.g. using the term "dog" solely in reference to the family pet) rapidly gives way to semantic overextension (e.g. referring to all four-legged animals as "dog"), before disappearing completely by the age of 3 or 4 (Clark, 1973; Ingram, 1989).

The tendency towards over-literal interpretation in children with semantic-pragmatic difficulties is further evident in problems regarding the comprehension of indirect speech acts (Bishop & Adams, 1989; Crystal & Varley, 1993; Culloden, Hyde-Wright & Shipman, 1986; Hyde-Wright & Cray, 1991; McTear & Conti-Ramsden, 1992). Children with semantic-pragmatic difficulties typically respond to the literal, surface meaning of an indirect speech act rather than to the intended meaning of the utterance. Thus an utterance such as "Can you close the door?" may elicit the response "yes", rather than prompting the child to perform the action of closing the door. Similarly, in the following interchange described by Bishop and Adams (1989, 249), the child responds to the surface meaning and thereby misinterprets an indirect request for his opinion as constituting a command to
Failure to grasp the intention beneath the surface form of language gives rise in some children with semantic-pragmatic difficulties to great difficulty in relation to so-called figures of speech (Bishop & Rosenbloom, 1987; Crystal & Varley, 1993; Culloden et al., 1986; Hyde-Wright & Cray, 1991; Shields, 1991; Vance & Wells, 1994). The comprehension of humour, irony, sarcasm, metaphor, similes, proverbs and idioms relies on the ability to use contextual cues to infer a non-literal sense of the words uttered. All of these present problems to at least some children with semantic-pragmatic difficulties.

Hyde-Wright & Cray (1991) describe the response of a child said to have semantic-pragmatic difficulties to a therapist's use of an idiom. On hearing her comment "Good, we are all ready to begin. Off you go", the child promptly got up from his seat and began to leave the room (p. 86). Culloden et al. (1986) note the difficulty that children with semantic-pragmatic difficulties typically have in interpreting other expressions like "pull your socks up". Even when used within a context clearly biasing an idiomatic meaning, such expressions may give rise to literal interpretations.

The literature from which these reports of overliterality derive tends not to address the relationship between semantic-pragmatic difficulties, Asperger syndrome and autism. Most of the children described in this literature tended, however, to have been diagnosed with "semantic-pragmatic disorder/difficulties" rather than with Asperger syndrome or autism. As noted above, it is widely accepted that all individuals with autism have at least some semantic-pragmatic difficulties. It has been further observed that some authors have characterised Asperger syndrome in terms of lesser language difficulties than are typically found in "semantic-pragmatic disorder" (Bishop, 1989). Nevertheless,

As pointed out by Howlin (1988 cited in Gross, 1994), literal interpretation of such expressions as "it's raining cats and dogs" and "fold your arms" may cause fear in individuals with Asperger syndrome. Gross (1994) quotes a remark made by the teacher of a 7-year old with Asperger syndrome. This illustrates the effect that literal idiom interpretation may have on the way in which the literal interpreter is perceived by others:

He can be so cheeky. And yet it's as if he doesn't know he's being cheeky. The other day I was telling him off and I said "Are you with me?" - you know, expecting him to say "Yes Miss" or whatever - and he said "well I'm next to you and I'm not with anyone else so I must be with you." And it wasn't as if he was being jokey; he's like that, very serious - it's hard to describe (p.105).

Whereas such literality may still seem reasonable in a 7-year old, the persistence of literal interpretation in Asperger syndrome is clearly illustrated by Gillberg (1991b). The author cites the case of a 15-year old boy with Asperger syndrome who had a full scale WISC IQ of 132. In response to his mother's saying "she's on the way to getting better", the boy enquired, "Whereabouts is that?". He then went on to ask several further questions which indicated that he had understood "the way to getting better" to be a street address (p.142).

Even in very high-functioning individuals with Asperger syndrome, traces of literality may still be evident. Happé (1991) quotes from the autobiography of Temple Grandin ("Emergence labelled autistic", 1986). She notes the literal use of the idiom, "a pat on the back", "in the sense of actual physical contact as praise" (Happé, 1991, 213).

Anecdotes like those quoted above illustrate the types of idiom-
related difficulty experienced by at least some individuals with semantic-pragmatic difficulties. Their inclusion in the literature has been important in highlighting this area of difficulty. In order to grasp the extent of the difficulty within individuals and within the broad category of semantic-pragmatic difficulties, however, it is necessary that idiom comprehension in these clinical populations be systematically studied.

7. STUDIES OF IDIOM COMPREHENSION IN CHILDREN CONSIDERED TO HAVE SEMANTIC-PRAGMATIC DIFFICULTIES OR LANGUAGE DISORDERS

To date, investigations of idiom comprehension in children with semantic-pragmatic or other language disorders are few in number and inconclusive in results. Two unpublished papers (Johnson, 1985; Strand, 1982 cited in Johnson, 1985) explore idiom comprehension in children with general receptive and expressive language deficits. Vance and Wells (1994) examined idiom comprehension in children said to have semantic-pragmatic difficulties and children with other language disorders. Finally, Happé (1994a) included two idioms in her investigation of high-functioning autistic individuals.

Using the same pictorial multiple choice task as was previously used to assess normally-developing children (Strand and Fraser, 1979), Strand (1982 cited in Johnson, 1985) investigated idiom comprehension in 18 language-disordered children (four or five children of age 7, 9, 11 and 13, respectively. Twenty idioms were presented in isolation. For each idiom, the children chose one picture from each of two picture cards. One set of options included a drawing depicting the literal meaning and the other included a picture depicting the idiomatic meaning. The children were then asked to explain their choices and their explanations formed the basis of the analysis.
The results indicated that language disordered 7-year olds were consistently able to explain some idioms but "clearly preferred to use a literal method for interpreting idiom meanings" (qtd in Johnson, 1985, 20). Among the 9-year olds, described as a "transition group", some subjects gave more literal explanations and others produced more idiomatic responses. By 11 years of age (and still at 13), the children provided predominantly idiomatic explanations. Comparing these results with those of her normal controls (Strand & Fraser, 1979, discussed above), Strand concluded that the language-disordered children were delayed in idiom comprehension development but that their overall pattern shadowed that of normally-developing children. Certainly they did not appear to be stuck in a literal mode of interpretation.

A very similar result was reported by Johnson (1985), who compared receptively- and expressively-disordered children, aged 5, 7, 9, 11 and 14, with language-normal controls matched on chronological age and performance IQ. Borrowing Ackerman's (1982) idioms and narratives, Johnson presented each of twelve common idioms in a literal, idiomatic or neutral context and required the children to select a corresponding line drawing. Different sets of pictures were presented for the literal and idiomatic uses of the expressions. Thus, the subjects were not required to choose between literal and idiomatic meanings, but merely to demonstrate semantic knowledge of the given idioms.

The results, consistent with those of Strand (1982 cited in Johnson, 1985), revealed that the idiom comprehension of the language-disordered children was lagging behind that of the normally-developing children by two or three years. The lag was, however, entirely consonant with their general delay in language comprehension. In other words, the comprehension of idioms presented no more difficulties for these children than did the comprehension of literal language. Furthermore, by the age of 14, the discrepancy between language-disordered and normally-

---

1 The frequency of the idioms was judged by adult raters.
developing adolescents had narrowed considerably. Indeed, the five language-disordered children of this age scored 85% of their idiomatic items correct as compared with a correct score of 90% among their normally-developing peers.

A further interesting result to emerge from Johnson's study regards the effect of contextual bias on comprehension. Whereas normally-developing children consistently scored more correct selections for the literal than idiomatic use of these idioms, the effect of context was not significant for the language-disordered children. Speculating on this result, Johnson suggests that rather than being uniquely difficult forms, problems in idiom comprehension may be no different in substance than the general problems in language comprehension experienced by these students... (The effects of these problems) are apparently equivalent across both literal and idiomatic forms. If true, the distinction between "literal" and "idiomatic" forms is an inaccurate heuristic device with regard to these students (p.61).

That the idiom comprehension of language-disordered children is appropriate for their receptive language age has recently been confirmed by Vance and Wells (1994). The authors matched 18 children with specific language impairments with 6- and 7-year old mainstream children on receptive language age\(^1\). The SLI group included seven SPLD ("semantic-pragmatic language disorder") children and eleven non-SPLD children. The subjects all undertook a preliminary task designed to exclude children who could not recognise semantically anomalous input. The authors note that this recognition may be an important factor in triggering non-literal interpretations in order to make figurative language meaningful. One SLI child (whether SPLD or non-SPLD is not reported) failed this task and was eliminated from the study along with the corresponding control subject.

At the outset, the children were introduced to the task with the

---

\(^1\) Subjects were matched on three subtests of the CELF-R (Semel, Wiig & Secord, 1987): Linguistic Concepts, Sentence Structure and Oral Directions.
Sometimes we say things in a way we don't really mean. If I say "We went out on Saturday and painted the town red", it doesn't mean we had a pot of paint and painted all the buildings, it means we had lots of fun. Listen to these stories and point to the picture of what you think each really means (p.35).

The children then listened to a tape recording of ten figurative items embedded in one- or two-sentence contexts. The items ranged in type from frozen idiomatic expressions like "he's got green fingers" and "it was raining cats and dogs", to "live" metaphors, such as "she flew across the room" and "(her bedroom) was a rubbish tip". Three line drawings on a picture card accompanied each of these items. These depicted the literal meaning, the non-literal meaning and an incorrect meaning usually related to one or more nouns in the idiom. For example, the drawings for "raining cats and dogs" depicted heavy rain falling (idiomatic), cats and dogs falling through the air with raindrops (literal), and cats and dogs on the ground with no rain (distractor).

The results of the study revealed no significant difference between normally-developing and SLI children. Interestingly, however, the two groups appeared to be following different patterns of development. For normal children, non-literal comprehension was significantly correlated with receptive language scores. However, no such correlation emerged for the SLI group (i.e. including SPLDs). This suggests that this group of children are not following the normal developmental path. The authors note that this finding might be explained in terms of heterogeneity within the SLI group. Thus, some children may be following the normal developmental path, others may have developed non-literal comprehension ahead of other language structures and a third group may be delayed in non-literal comprehension relative to their comprehension of other language structures. This latter description corresponds to the characterisation of children with semantic-pragmatic difficulties. Strikingly, however, the authors found that the
difference between SPLD children and non-SPLD children was not significant. In fact, the SPLD children performed marginally better, on average, than did the non-SPLD children.

The results of this study are notable in that they refute the characterisation of children with semantic-pragmatic difficulties in terms of impaired idiom comprehension. Interestingly, Vance and Wells, themselves, point out that one SPLD child who achieved 70% accuracy on their task later responded to his mother's comment "You're full of beans today" with the words, "I most certainly haven't had any beans today" (p.39).

While it is unclear how representative this one child is of the SPLD subgroup, the discrepancy between his test and real life performance may be due in part to the effect of the task situation. Bishop and Rosenbloom (1987) observe that in children with semantic-pragmatic difficulties,

there may a marked difference between the child's ability to understand in a structured, concrete situation, such as a multiple choice test, and the ability to comprehend normal conversation, where the speaker refers to events that are not immediately deducible from the physical context (p.34).

Vance and Wells themselves conclude that the SPLD children's difficulty may lie not in an inability to interpret figurative language, as much as in their "lack of realization that a non-literal interpretation is required, or in their lack of recognition that the literal interpretation they have made is meaningless and nonsensical" (p. 39).

In addition to its clinical implications, this conclusion serves to underline the point that Vance and Wells did not actually investigate idiom comprehension in its totality. Instead, like Strand (1982 cited in Johnson, 1985) and Johnson (1985), they examined only the children's semantic knowledge of idioms. As discussed above, idiom comprehension involves both semantic knowledge and pragmatic skill. The latter includes the ability to recognise contextual incongruity and to retrieve a meaning
which is coherent with the overall discourse. In their study, Vance and Wells greatly reduced the pragmatic demands of idiom comprehension in context by explicitly indicating that it was the idiomatic meaning that was required. In this light, and in view of the fact that a pretest had screened out subjects who failed to recognise semantic anomaly, the impressive performance of the SPLD children begins to seem less surprising. Nevertheless, this result is important to the extent that it suggests that the semantic component of idiom comprehension in SP children may be largely intact.\footnote{As stressed by such authors as McTear and Conti-Ramsden (1992) and Smith and Leinonen (1992), it is important to retain the distinction between semantic and pragmatic aspects of functioning when seeking to identify and understand children's difficulties.}

With regard to this finding, it is unfortunate that the results for each individual item are not presented in the report of this study. Since the study included both idioms and metaphors, it is not possible to draw any conclusions as to the subjects' strengths and weaknesses in relation to either of these forms, individually. The authors do, however, state that one item emerged on an item analysis as having been particularly poorly interpreted (more than one standard deviation below the mean) by both the SLI and mainstream groups. This item read "James is always fooling around, at the party he was acting the goat". Furthermore, the following two items were both found to have been handled particularly well by all the children:

(1) Laura was told to tidy her bedroom, but when Mum looked in after tea, Mum said it was a rubbish tip.

(2) James is always very greedy, at the birthday party he was a pig.

The authors find themselves unable to account for the discrepancy between these items and the others on the task. It is interesting to note, however, that the items on which the subjects did very well each contained only one non-literal expression: the target item. In contrast, the item quoted above as being one on which the subjects scored particularly poorly,
contained more than one idiom. In addition to the target idiom, "acting the goat", the expression "fooling around" is also a true idiom. Indeed, a further four of the remaining nine items contain a "semi-idiom" in addition to the target item (e.g. "It was night time and all dark in the house, James turned the light on, it was a single star twinkling in the night"). Given that the "easy" items each contained only one non-literal string while the "difficult" item contained two true idioms, it may be that the unintentional inclusion of additional idiomatic expressions partially accounts for the between-items differences in this study.

It is further unfortunate that the statistical results that the authors present do not include a comparison of the mainstream controls with the SPLD and non-SPLD subgroups, individually. This limits the extent to which conclusions can be drawn from this study with regard to idiom comprehension in children with semantic-pragmatic difficulties.

While a direct comparison between the mainstream and non-SPLD children cannot thus be made, it is worth noting that the results of the non-SPLD group might have been inflated relative to those of the mainstream children. The groups were matched on receptive language scores¹. As McTear and Conti-Ramsden (1992) and Plante et al. (1993) note, however, such language matching does not ensure that the groups do not differ linguistically and cognitively. In this case, the age variable may well have played a very important role. Aged between 6;4 and 7;8 years (mean age 6;11), the mainstream children may have had less exposure to idioms than had the older SLI children (7;10-13;1, mean age 9;5). The mainstream children, in terms of current understanding, would have been further likely to be at a transitional phase between literal and idiomatic understanding. As such, their cognitive capacity for idiom comprehension may not have been as developed as that of the non-SPLD group.

¹Subjects were matched on three subtests of the CELF-R (Semel et al., 1987): Linguistic Concepts, Sentence Structure and Oral Directions.
On the basis of these results, Vance and Wells suggest that the characterization of children with semantic-pragmatic difficulties in terms of impaired idiom comprehension should be reconsidered. In time, it may well be that the description of semantic-pragmatic difficulties will need to be revised in this respect. However, given the paucity of research in this area and the several problems of Vance and Wells' study, further research would seem prudent before any firm conclusions are drawn.

Even when compared with that pertaining to children with semantic-pragmatic difficulties, the literature investigating idiom comprehension in Asperger syndrome and/or high-functioning autism is very scarce indeed. Only one study known to this author (Happe, 1994a) has investigated idiom comprehension in this group systematically. Indeed, even Happe's study included only two idioms — of which one is not identified in the report — alongside examples of eleven other forms of non-literal language use. The purpose of the study was to assess theory of mind skills in autistic individuals who had already passed first order and, in some cases, second order theory of mind tasks¹.

Happe's study examined comprehension of idioms, jokes, lies, white lies, misunderstandings, double bluffs, sarcasm, persuasion, contrary emotion, appearance/reality distinctions and instances of forgetting. As for the other types of non-literal language, each of the two idioms ("have a frog in one's throat" and another unidentified expression) was embedded in the last sentence of a short narrative, biased towards the idiomatic interpretation. After hearing the narratives, the subjects were asked about the truth of the statement and were required to justify its use, as illustrated below.

¹The ability to pass such tests has been associated with Asperger syndrome and high-functioning autism (Bowler, 1992 cited in Happe, 1994b; Ozonoff et al., 1991.). The verbal IQ range in this group of 64 to 101, mean 87.3, further supports the consideration of this group as lying at the upper end of the autistic continuum.
Emma has a cough. All through lunch she coughs and coughs and coughs. Father says, "Poor Emma, you must have a frog in your throat!"

(a) Is it true, what Father says to Emma?
(b) Why does he say that?

On the first question, most of the subjects proved able to recognise the lack of truth in the target statement for all of the story types. The second, justification question was scored on two dimensions: correct or incorrect, and mental state or physical state. Mental state responses referred to thoughts, feelings, desires and dispositions. Thus, "It's just an expression people use" is a correct mental state justification for the item, "you must have a frog in your throat". Physical state responses refer to physical appearance or physical events (eg. "She swallowed a frog"). Happé reports that incorrect mental state responses were found to constitute a discriminating measure. The able autistic subjects were found to make more errors in attributing mental states to characters than did normal and mentally handicapped controls for ten of the twelve story types. Idioms were one of only two story types (the other was contrary emotions) on which the autistic subjects made as few errors as did the control groups.

In discussing this aspect of their results, the authors state that "figures of speech may be learned without any appreciation of mental states, since they are frozen expressions" (p.141). Indeed, relative to Happé's other forms of non-literal usage, idioms may be argued to require less appreciation of mental states. Nevertheless, as evidenced in the review of idiom comprehension in children, a grasp of the speaker's/writer's intention is fundamental to the comprehension of idioms in context.

Indeed, perhaps the greatest weakness of idiom comprehension research in these clinical populations to date, has been inadequate attention to the vital pragmatic component of idiom comprehension in context.
Despite inconclusive findings regarding idiom comprehension in language-disordered populations, two standardized instruments claim to identify language disorders on tests which include an investigation of idiom comprehension. These are the Test of Language Competence - Expanded Edition (Wiig & Secord, 1989) and the Fullerton Language Test for Adolescents (Thorum, 1986). In addition, the Conley-Vernon Idioms Test (Conley, 1976) was produced specifically to assess idiom comprehension in normally-developing and hearing-impaired children.

The Test of Language Competence - Expanded Edition (henceforth TLC-E) was designed to identify children, adolescents, and young adults who have not acquired the expected levels of metalinguistic competence in semantics, syntax, and/or pragmatics" (p. 1).

The test comprises four subtests at each of two levels. Level 1, an addition to the first edition of the TLC (Wiig & Secord, 1985), is aimed at children aged 5 to 9. Level 2 is aimed at the age range 9 to 18+

The Figurative Language Subtest at Level 1 includes "idioms and metaphors selected from a repertoire of expressions commonly encountered in American life" (TLC-E Technical Manual, p.4). Of the 17 items, all but one meets the present study's criteria for idiomaticity: "give a hand", "look low" (double-function word), "go it alone", "have a ball", "go fly a kite", "get the jump on", "have a way with", "fall behind", "come apart at the seams", "be on your toes", "see red", "be cut from the same cloth", "go in circles", "make waves", "take the wind out of someone's sails", "cut corners" and "march to a different drummer".

The items included at Level 2 are all described by the authors as being metaphors, representative of four types: structural, orientational, ontological and synecdoche. Illustrative of the
overlap between metaphor and idiom, several of these items also meet the criteria of this study for idiomaticity. These include "zero in on", "new blood" and "hold all the aces".

At Level 1, children are introduced to the Figurative Subtest with the sentence "Sometimes we use different kinds of words when we talk to people". The tester then works through two trial items with the child before beginning. Each item is read within a two-sentence context. Item 11, for example, reads

This is what a girl said about her mother: "My mother saw red."

After each item, the child is asked "What did (the girl) mean when she said ('my mother saw red')?" S/he is allowed ten seconds in which to paraphrase the expression. Thereafter, the child is shown a picture card containing four black-and-white line drawings. S/he is instructed to "point to the picture that shows what (the girl) meant when (she) said (my mother saw red)". The child is allowed 15 seconds for this task.

The pictures corresponding to each expression depict the idiomatic meaning, the literal meaning, the opposite meaning to the idiomatic one and a related literal interpretation, respectively. The picture card reproduced in Figure 1 (below), for example, corresponds to the item "my mother saw red".

At Level 2, the tester introduces the subtest by saying, "I'm going to tell you about a situation and what someone said. Listen carefully to the situation and what was said." One trial item precedes the test. As at Level 1, the twelve expressions are each read within a two-sentence context. They are also shown to the student in print. The student is asked "Can you tell me what the speaker meant?" The student is allowed 15 seconds to respond. Thereafter, s/he listens to four expressions and sees them all in print. The tester instructs "Now tell me which choice could be used instead of (repeats the expression)." The student is allowed 45 seconds to respond to each item.
Interpretation responses at both levels are scored "correct" or "incorrect" in accordance with interpretation rules provided. Raw scores are translated into standard scores, comparable with the norms provided.

The TLC-E has been standardized and norm-referenced. The reported discriminant analysis indicates that language-disordered children were differentiated from language-normal children on the TLC-E. Notably, however, 26% of the sample was incorrectly classified by the Figurative Language Subtest at Level 1. Closer examination of this subtest reveals a number of important threats to the validity of the Figurative Language subtest.

The first of these relates to the inclusion of a verbal definition task at both levels. The test is designed to identify children or adolescents "who have not acquired the expected levels of metalinguistic competence in semantics, syntax, and/or pragmatics" (p. 1). In requiring these children to explain the meanings of figurative expressions, the task necessitates the use of the very skills which such children lack. Consequently, an "incorrect" score might reflect lack of comprehension or merely an inability to explain a known expression.

The expressive requirement is particularly concerning with regard to children with semantic-pragmatic difficulties. Word-finding difficulties, use of unestablished referents, inappropriate presuppositions, unnecessary reiteration and violations of Gricean maxims of quantity and relevance (Bishop & Adams, 1989) might all lead to a judgement of "incorrect", even when a child has in fact grasped the idiomatic meaning of an expression.

The nature of the pictorial multiple-choice task in Level 1 gives rise to a second major confounding factor. In order to make sense of the pictures in relation to the task, the child requires the ability to draw inferences, to interpret non-verbal language and to recognise drawing conventions. Children who have difficulty in these areas may understand the expression but fail
to understand the connection between the expression and the drawing. Certainly children with semantic-pragmatic difficulties are put at a disadvantage by the test procedure itself.

The test's reliance on this type of interpretation and inferencing is particularly problematic in relation to the many items which describe abstract states rather than concrete actions (e.g. "have a ball", "come apart at the seams", "be on your toes", "see red"). The pictures corresponding to these use facial expression, body posture and associated activities to represent their idiomatic sense. There are sometimes several stages of inference required to link the expression with the correct picture and to recognise the others as incorrect. This is clearly illustrated by Item 11 ("My mother saw red"), reproduced above (Figure 1).

Given that the pictures are all black-and-white, the child must, first of all, reject the literal interpretation by inferring "redness" from the shading on the drawing of an apple. This requires that the child recognises the convention of shading to represent colour and infers that the colour of the apple is intended to be red, rather than green or yellow. The child has further to assume that no other shaded areas in any of the other pictures are intended to be inferred as red.

In addition to this, the child must understand that the woman in the "literal" picture is meant to be looking at the picture of the apple. This requires the child to recognise and interpret the drawing convention of downcast eyes.

In order to integrate all of this information, the child needs to be attending to the whole of the picture rather than to individual parts of it. Simultaneously, s/he must recognise that the woman and the apple represent the salient aspects of the picture.

All the while, the child must also be scanning the other pictures
and making comparisons. Crucially, s/he must observe that it is only in the "idiomatic" picture that the woman looks angry. To do this, the child must interpret the woman's raised eyebrow and index finger, and forward-leaning posture as representative of anger. S/he also has to make the connection between this display of anger and the untidy bed. This requires him or her to recognise that the woman is meant to be the girl's mother and must draw on social world knowledge that mothers may get annoyed when children do not keep their bedrooms tidy. Further, s/he has to grasp the implications of the child's submissive posture and infer that the child is probably the agent of the untidiness and hence the object of the mother's anger.

Clearly, the validity of any result scored by a child with semantic-pragmatic difficulties on the pictorial multiple-choice is greatly threatened by the fact that the task itself requires skills that are characteristically lacking in such children. In this context, there is an arguably higher than usual chance that responses will guided by picture preference or will simply be guessed.

The multiple choice task introduced at Level 2 is similarly undermined by an interfering variable of some magnitude. Four options are provided for each of the twelve items. The correct match represents the same metaphorical structure and meaning as the target item. The three foils represent

(1) a metaphoric expression from the same structure category as the sample stimulus and match but with the opposite meaning,

(2) a literal statement, featuring key words or concepts from either the sample or the match, appropriate for the situation, and

(3) a non-related figurative expression, featuring words or concepts present in the sample stimulus or the match" (Wiig & Secord, TLC-E Technical Manual, 1989, 6).

These are illustrated by the following example (Item 3):
Situation: Two girls talking about a friend running for class president
Expression: "She seems to be holding all the aces."
Match: The odds certainly favor her.
Literal: She has four aces in her hand.
Opposite: The chips seem to be down for her.
Non-related: She is a real card shark.

As evidenced by this example, a correct choice on this task requires that the subject comprehend and assess the appropriacy of the target expression as well as three other figurative expressions. (Of course, guessing would also yield a correct answer by chance in 25% of instances.) Indeed, in this instance, the "situation" sentence contains yet another idiom.

Perhaps more importantly, an incorrect response on any of the twelve items is almost impossible to interpret on the basis of this task. An incorrect choice might reflect failure to understand any one of at least four idioms (or a combination) or might result from an incorrect guess. Not only is the tester left uncertain as to which expressions the child does and does not understand, but s/he is further unable to assume that a literal response reflects the child's use of a literal strategy. The structure of this task would thus appear to limit its efficacy as a basis for therapy.

The questionable generalizability of norms to British children, along with the range of serious threats to the validity of its results, make this test an unsuitable assessment measure for children with semantic-pragmatic difficulties (and arguably for children with any language disorder). Ironically, it may in fact discriminate between normal children and those with language disorders by virtue of its inclusion of inappropriate tasks. The results might well be misleading with regard to idiom comprehension, however, and would need to be treated with extreme caution.

Similarly to the TLC-E, the Fullerton Language Test for
Adolescents (henceforth FLTA) (Thorum, 1986) is designed to discriminate between language-impaired and language-normal subjects. This test comprises eight subtests all of which have been standardized on 762 normal subjects between the ages of 11 and 18. The test is administered individually.

The Idiom Subtest contains 20 items, all of which were known to at least 30% of subjects in a preliminary field test. The tester introduces the subtest by saying,

I am going to say some words to you that have a double meaning or that can be taken in more than one way. For example, do you know what it means when someone says "that person gets on my nerves"?

The tester may continue to give examples from a list of six items provided. The student is then asked to explain the meanings of the idioms which are presented in isolation.

Responses are scored "correct" or "incorrect" on the basis of whether or not they approximate the definitions provided in the manual. Each "correct" response scores one point up to a maximum of 20 points.

Mean scores are provided for 11- to 18-year olds at 6-month intervals. These reveal that there is only a three-point difference in accuracy of idiom comprehension between age 11 (mean score = 7, s.d. = 7.4) and age 18 (mean score = 10, s.d. = 6.1). The progression is furthermore not linear: 16-year olds achieve the same mean as 18-year olds and these are both lower than that reported for 17-year olds. As Stephens and Montgomery (1985) note, the absence of progression characteristic of all the subtests on the FLTA begs the question "why is this test not showing developmental differences in performance when such differences occur with intelligence tests (and are presumed to occur with language abilities)?" (p. 32). Indeed, the idiom comprehension literature reviewed above indicates continued development in this sphere. Thorum (1980) notes that a ceiling level was attained between the ages of 8 and 12 and thus the test
does not claim to be a developmental measure but only to discriminate between language-impaired and language-normal students.

In the light of this claim, Stephens and Montgomery point out the glaring omission of data regarding the proportion of each group correctly identified by the FLTA during the norming procedure. Indeed Thorum's claim for predictive (diagnostic) validity rests on a statement that t-test analyses showed a significant difference in the scores of mainstream students and students attending a variety of special education programmes.

To assist in the interpretation of scores, the FLTA outlines criteria for three levels of performance: competency level, instruction level and frustration level. Once again, Stephens and Montgomery (1985) focus on a major problem with this categorisation. According to the FLTA manual, a student achieving a score within the instruction level "is in a period of transition and has many of the skills or processes required to perform the task, but ... further instruction is needed to achieve competence" (1980, 13). The instruction level is set at scores that fall within one standard deviation of the mean for the norming population. This implies that instruction is required for students performing within the normal range. Furthermore, given an extremely wide range of scores within one standard deviation of the mean, most students would fall into this category. In fact, for those aged 11;0 to 11;5 (mean score = 7, s.d. = 7.4), a score anywhere between 0 and 14 out of 20 would qualify for instruction.

Following the FLTA's instructions, Lieberman, Heffron, West, Hutchinson and Swem (1987) found that 70% of the normal sixth grade children they tested on the idiom subtest failed to reach competency (N=30, mean age 11;9). Brasseur and Jiminez (1989) tested 71 American university students on the FLTA idiom subtest. The subjects were divided into three age groups: 18 to 21, 22 to 29, and 30 upwards. On the basis of Thorum's (1980) overall mean
of 9, subjects had to achieve between 13 and 20 to reach the competence level. In fact, only 51% (ages 18-21), 84% (22-29) and 91% (30 year up) achieved this level. These were all speech and language therapy students and were all above the age range for which the test was devised.

A further surprising result to emerge from the Brasseur and Jiminez study was that five out of the 20 idioms were incorrectly defined by at least a quarter of the subjects. In the 30+ age group, 30% of the idioms were not accurately explained by at least 25% of the subjects. This figure rose to 40% of the items among 22- to 29-year olds, and to as high as 70% in the 18-21 age group. Even Thorum, himself, found that seven items out of the 20 were understood by less than 30% of the norming sample. Only three of the 20 idioms were correctly defined by over 80% of this population.

In the light of general consensus regarding the weakness of this test (Brasseur & Jiminez, 1989; Lieberman et al., 1987; Stephens & Montgomery, 1985), any results yielded by the FLTA idiom subtest clearly need to be treated cautiously. Its suitability for use with children with semantic-pragmatic difficulties or for those with other expressive specific language disorders is further undermined by its reliance on verbal expression.

The Conley-Vernon Idiom Test (Conley, 1976) was designed specifically to assess comprehension of idioms in hearing-impaired children. One hundred idioms are included, all drawn from the list of essential idioms produced in A Dictionary of Idioms for the Deaf (Boatner & Gates, 1969 cited in Conley, 1976).

For each item, a sentence is presented with the idiom omitted. Subjects read the sentence and fill in the blank space from a choice of five responses, including the correct idiom. For example:
The students studied the poem until they knew it ________.

a) took their time  
b) by heart  
c) broke in  
d) a little  
e) mixed up  

Two parallel forms of the test have been developed, each including 50 items. Conley's subjects completed the test in two sessions with a time gap ranging from several hours to one day.

The test commences with instructions for choosing answers and using the answer sheets. Help with the sample questions may be given if required. There is no time limit. The subject's score is the number of correct items out of a total of 100.

Statistical reliability for the test was established and the author notes that the test has face validity and sampling validity. Conley concedes that concurrent validity could not be measured by relating this instrument to any other, since no other test of idioms existed at the time (1976).

In fact it is with respect to construct validity that the limitations of this test emerge. Several potentially confounding factors can be identified, many of which are illustrated by the sample item reproduced above. Firstly, it appears that two of the five possible responses could plausibly replace the blank: "by heart" or "a little". There is no indication in Conley's report of subjects having been explicitly instructed to choose an idiomatic response. In the absence of such an instruction, a student might assume that any plausible answer is acceptable. In this instance, the incorrect score may not necessarily reflect the subject's lack of idiomatic comprehension, but rather an unlucky choice from between two plausible responses.

Secondly, it is notable that only the two responses, "by heart" and "a little" are syntactically appropriate. As such, the five possible answers do not all stand an equal chance of being
Equal likelihood of selection is further reduced by the inclusion of idioms other than the target one in the list of choices. If a student fails to understand one or more of the other idioms listed, s/he might simply ignore that/those options entirely and select from among those that are understood.

The inclusion of four idiomatic expressions within these five choices gives rise to yet another potentially confounding factor. How is it possible to judge whether an incorrect-but-idiomatic response reflects misunderstanding of the target idiom or misunderstanding of the idiom selected?

A more general problem of this test relates to its length. Even in two halves, the test requires children to concentrate and maintain motivation until 50 items are completed. Certainly within the lower age range, fatigue and attention drift might well play a role in determining the outcome.

The test was designed for a specific research purpose and, as such, its inappropriacy for use with language-disordered populations is not a criticism. It is nevertheless the case that the Conley-Vernon Idiom Test would not constitute a suitable clinical assessment measure. In addition to the serious threats to its validity outlined above, the attention span and period of time required, as well as its reliance on reading skills all detract from its suitability with this group.

As evidenced by the review of the idiom comprehension literature, the tasks used to assess idiom comprehension have a significant impact on the outcome of such investigations. It is apparent that tasks requiring children to explain the meanings of idioms are more difficult than those in which they select from among pictorial or verbal interpretations. In addition, idiomatic understanding is facilitated by the inclusion of an informative, idiomatic-inducing context. Other factors such as the idiom's
syntactic flexibility, semantic transparency and degree of familiarity may also influence the child's comprehension. The effect of these factors on the comprehension of children with semantic-pragmatic difficulties is as yet unclear. Clinical and research assessments of this population should, however, be alert to the potential influence of this range of factors and should take care not to prejudice the outcome through tasks which themselves make unreasonable demands on children with semantic-pragmatic difficulties.

9. IDIOMATIC USAGE IN EDUCATIONAL SETTINGS

Given that idioms are a central feature of everyday language, a child's difficulty in interpreting these expressions is cause for concern. Indeed, it is claimed that approximately two thirds of the English language comprises idiomatic constructions (Boatner & Gates, 1969 cited in Conley, 1976). In a study of teachers' use of idiomatic language, Lazar, Warr-Leper, Nicholson and Johnson (1989) found that 11.5% of teachers' utterances contained at least one idiom. Nippold (1990 cited in Nippold, 1991) examined textbooks from three widely-used reading programmes aimed at school children. The author found that idioms were rare in books intended for kindergarten to the second grade (age 7), but that by the third grade (age 8), children were exposed to books in which 6% of the sentences contained at least one idiom. By the eighth grade (age 13), children were using books in which idioms occurred in an average of 9.7% of sentences. A study conducted by Nippold (1988 cited in Ezell & Goldstein, 1991), indicates that a grasp of idioms is indeed necessary for the achievement of literacy.

Despite these indices of the significance of idioms, Blue (1981) recommends that teachers should avoid using figurative expressions in interacting with language-disordered pupils until they have achieved proficiency in literal language use. Gross
(1994) suggests that "for children with literal understanding of language ... jokes may have to be explained, irony and idiom avoided" (p. 107). Similarly, guidelines issued at one specialist language unit advise staff against the use of "sayings" in interaction with children with semantic-pragmatic difficulties.

In considering these recommendations, we need to address two questions. Firstly, is this goal achievable? Given the extent of idiomatic usage in everyday language and the finding that many of the teachers recorded for this project did not recognise the idiomaticity of large numbers of expressions they used in the classroom, we may conclude that the total avoidance of idioms is not a realistic aim. Secondly, we need ask whether this aim is desirable. Lazar et al. (1989) argue that the absence of idiomatic language in the classroom and in therapy would have a negative effect on language-disordered children. The avoidance of these expressions would function to restrict the opportunities that such children had to learn the meanings of idioms. Instead the authors recommend that teachers be alert to verbal and non-verbal signs of confusion and respond with explanations of the relevant expressions.
CHAPTER 3: AIMS AND HYPOTHESES

The characterisation of semantic-pragmatic difficulties in terms of literality has been based predominantly on anecdotal evidence. Only one study known to this author has systematically investigated non-literal comprehension in children with semantic-pragmatic difficulties (Vance & Wells, 1994). As discussed in Chapter 2, the validity and generalizability of its results are questionable. Nevertheless, it is interesting that Vance and Wells should have failed to find evidence of idiom comprehension difficulty in children with semantic-pragmatic difficulties.

In the light of this discrepancy in the literature, further investigation is required before an impairment in idiom comprehension may be deemed characteristic of this population. At the group level, research in this field has the potential to increase understanding of the variation in skills and weaknesses within the population of children considered to have semantic-pragmatic difficulties and to contribute to the debate regarding the relationship between semantic-pragmatic difficulties, language disorders and the autistic continuum. At the individual level, insight into particular children's abilities and disabilities in relation to idioms would allow clinicians to optimise their therapeutic input.

Against the backdrop of these overarching goals, this study aimed, firstly, to investigate the extent to which idiom comprehension difficulty is characteristic of children with semantic-pragmatic difficulties. Secondly, the study aimed to examine whether ability or disability in this regard forms a sound basis for discriminating between children considered to have semantic-pragmatic difficulties, those who are developing normally and those considered to have language disorders (other than semantic-pragmatic difficulties).

To these ends, the study focused on four groups of subjects, one
of which - the SP group - divided into two subgroups: children diagnosed a priori with "semantic-pragmatic disorder/difficulties" ("SPD") and those diagnosed as having Asperger syndrome or high-functioning autism ("ASP"). In addition to the SP group, a second clinical group diagnosed with a range of language disorders other than "semantic-pragmatic disorder/difficulties" was included for comparison. Two groups of mainstream children were also investigated in order that a baseline for these idioms might be established. The younger mainstream infant school (MI) group comprised children aged 6;6 to 7;6. The mainstream junior school (MJ) group included children aged between 10;6 and 11;6.

Few formal procedures are currently available for testing idiom comprehension in children. Those that are available have been argued to have significant weaknesses, particularly when used to assess children with semantic-pragmatic difficulties. Like the formal tests, studies of normally-developing children have tended to rely on multiple choice and definition tasks. Both of these have been argued to produce questionable results when used with children with semantic-pragmatic or other language difficulties. In order that idiom comprehension might be validly investigated in these groups, the development of a more suitable methodology was considered necessary.

To this end, a play-based procedure was devised and piloted prior to this study¹ (Kerbel, Grunwell & Grundy, 1996; see Appendix A). In the present study, twelve common idioms, representing a range of idiom subtypes, were embedded in a single narrative. Each child listened to the whole story on tape before hearing it again, played sentence by sentence. On the second occasion, the child was instructed to act out the story using the play set and

¹Following completion of this study, it was discovered that a similar methodology was used by Vosniadou, Ortony, Reynolds and Wilson (1984) to investigate normally-developing children's comprehension of metaphors. This investigation required the children to act out metaphors using a play set. The play-based methodology is therefore not entirely novel. However, its use to investigate comprehension of idioms and its employment in the clinical context is innovatory. Furthermore, the play-based procedure used by Vosniadou et al. presented each metaphor in a separate narrative and alerted the children to the presence of non-literal language. In contrast, the idiom play task embedded all twelve idioms into a single narrative and the subjects remained blind as to the presence of idioms and the purpose of the play.
props provided. This play was video-taped. The child then watched the video and, after each instance of idiom, was asked to define the expression. Each play action and definition was coded in accordance with a set of codes devised for this procedure. The data were then subjected to a range of statistical analyses.

The purpose of including a definition task, despite the limitations associated with it, was twofold. Firstly, the researcher wished to compare performance on the newly-devised idiom play task with that on a more conventional task. Secondly, the researcher anticipated that the definitions produced would yield valuable supplementary information with regard to the level of idiom development within the various subject groups. In view of the questionable validity of definition task results, the results of the play task were considered to be a more accurate reflection of the children's idiom comprehension ability than were those yielded by the definition task.

On the basis of the literature reviewed in Chapter 2, it was hypothesized that

1. children considered to have semantic-pragmatic difficulties (the SP group) would be less able to demonstrate the idiomatic meaning of twelve common idioms on the play task than would normally-developing children (MJ and MI groups) and children considered to have language disorders other than semantic-pragmatic difficulties (LD group);

2. the SP group would demonstrate more inappropriate than appropriate interpretations; and

3. each of the four groups would be less able to demonstrate the idiomatic meaning of these idioms on a task which required them to use language to define idioms (the definition task) than on one which required them to demonstrate the idiomatic sense of idioms using a play set (the play task).

91
In addition to these hypotheses, the study aimed to investigate the following questions:

1. What patterns, if any, differentiate children diagnosed "SPD" ("semantic-pragmatic disorder / difficulties") from those diagnosed "ASP" (Asperger syndrome / high-functioning autism)?

2. What might account for impaired idiom comprehension and does all such disability stem from the same underlying difficulty?

3. Which idioms, if any, are consistently associated with high or low levels of comprehension, and what might account for these differences?

---

1 These diagnoses were made prior to the study by specialist SLTs, psychologists or psychiatrists. In many instances, these were the diagnoses entered on the children's Statement of Special Educational Needs.
1. SUBJECTS

A total of 75 children participated in this study. These included 15 mainstream children aged 6;6 - 7;6 (mean age 6;11, s.d. = 3.8 months), 15 mainstream children aged 10;6 - 11;6 (mean age 10;9, s.d. = 3.5 months), 15 children identified as having specific language disorders excluding primary semantic and/or pragmatic difficulties (aged 8;6 - 11;2, mean age 9;11, s.d. = 11.8 months), and 26 children identified as having semantic-pragmatic difficulties, Asperger Syndrome or high functioning autism. The latter group ranged in age from 6;6 - 11;3 (mean age 9;5, s.d. = 16.7 months).

The four remaining children (two LD and two SP children) were eliminated from the study. One LD child was found to have a hearing impairment subsequent to the session. The other LD child ceased acting two-thirds of the way through the action task and responded "don't know" or with silence to all of the questions in the definition task. The child's teacher commented on his severe general lack of motivation. The two SP children, both diagnosed autistic, failed to act at all despite repeated explanations of the task. Their teacher expressed the opinion that the task was too complex for these children. There were thus 75 children involved in the study, but only 71 completed the two tasks and were included in the analysis.

(a) MAINSTREAM SUBJECTS

The mainstream children were drawn from second-year infant classes in three different schools and from fourth-year junior classes in two schools. The teachers of each of these classes had previously been tape-recorded for the study (see below). An additional child known to the researcher was included in the
junior school group.

Prior to assessment, letters requesting parental permission for children to participate were distributed to all the children in the mainstream classes. The letters informed parents that the study was investigating children's understanding of the language they hear at school and on television. They were told that the children would be seen for two sessions. In the first session, the child would be asked to point to the pictures corresponding to a set of sentences. In the second session, the child would be required to listen to a story and would then be filmed while s/he acted it out using Playmobil people and props.

In two of the schools, parents were asked to return a permission slip if they were willing for their children to participate. In the third school, parents were asked to return a slip if they did not wish their children to participate. While a positive response was, ceteris paribus, the preferred approach to obtaining permission, the third school had previously experienced very low response rates when parents were required to return letters.

Parental permission was obtained for a total of 51 infant school children and 31 junior school children. Of these, 39 infants and 31 juniors, respectively, were identified as fulfilling all of the following criteria:

(i) the child spoke English as his/her first language
(ii) the child was aged between 6;6 and 7;6, or 10;6 and 11;6
(iii) the child had not participated in the pilot study (see below)
(iv) the child's hearing was normal
(v) the child was not receiving language therapy or special needs assistance.

From those who fulfilled all of the above criteria, five children were randomly selected from each of the three infant classes for
the first mainstream group (MI). For the second mainstream group (MJ), seven children were randomly selected from each of the two junior classes. Owing to the timing of the data collection, none of these junior school children were older than 11;2. Since it was desirable to assess children across the age group 10;6 to 11;6, a fifteenth child (aged 11;6) who attended a different school was included.

In order to screen out mainstream children with below-average language comprehension, the Test for Reception of Grammar (TROG, Bishop, 1983) was administered as a pretest to all the mainstream children. Subjects scoring less than one standard deviation below the mean were excluded from the mainstream subject groups. One child in each age group failed to attain -1 s.d. A further child in each age group was selected at random and found to meet the TROG criterion. This brought the number of subjects in each group to 15. Teachers confirmed that these children represented a range of ability levels but were all considered to fall within the normal range for their ages. The infant school (MI) group comprised nine males and six females. The junior school (MJ) group included five males and ten females. (Details of age, gender and TROG scores are included in Appendix B.)

After the pretest, the children were told what would be required of them in the second session and were asked whether or not they would like to proceed. They all expressed eagerness to continue. Following the TROG session, at least one week was allowed to lapse before the second session.
(b) CHILDREN DIAGNOSED WITH SPECIFIC LANGUAGE IMPAIRMENT / LANGUAGE DISORDERS EXCLUDING PRIMARY SEMANTIC-PRAGMATIC DIFFICULTIES (LD)

Speech and language therapists in three language units were asked to identify children in their schools who met all of the following criteria:–

(i) the child was aged between 6;6 and 11;6
(ii) the child spoke English as his/her first language
(iii) the child was diagnosed or considered by SLTs to have a specific language impairment/language disorder without primary semantic-pragmatic difficulties
(iv) the child's speech was intelligible
(v) the child's hearing was normal
(vi) the child was considered to be of average or above-average intelligence
(vii) the child did not have severe behavioral problems

Letters were distributed to the parents of all the children identified as suitable, informing them that the study aimed to investigate comprehension of classroom and television language in children with language disorders. The procedure was outlined and parents were requested to return a slip if they were willing for their children to be involved. Permission was obtained for 17 children. As noted above, two of these children were excluded from the study. The remaining 15 children included twelve males and three females. The age, gender and diagnosis of individual children are included in Appendix B, along with language test scores provided by SLTs.

Immediately prior to the session, the child's teacher or SLT told the child what s/he would be required to do and asked whether or not s/he would like to be involved. The children all agreed to participate.
The children in this group were drawn from twelve different schools. Twenty-two of the children attended specialist units for children with language disorders, moderate learning difficulties or autism. A further six children attended mainstream schools.

The children were identified by SLTs or by the Leicestershire Autism Outreach Service as meeting the following criteria:

(i) the child was aged between 6;6 and 11;6
(ii) the child spoke English as his/her first language
(iii) the child was diagnosed or considered by a SLT, clinical or educational psychologist or psychiatrist to have "semantic-pragmatic disorder/difficulties", Asperger Syndrome or high-functioning autism
(iv) the child was considered to be of average or above-average intelligence
(v) the child's hearing was normal

Letters were distributed to the parents of suitable children informing them of the study's aim to investigate comprehension of classroom and television language in children with semantic-pragmatic difficulties, Asperger syndrome or high-functioning autism. The procedure was outlined and parents returned a permission slip if they were willing for their children to participate. The researcher obtained permission to work with 28 children. As noted above, two of these children were eliminated from the study. Of the 26 remaining children, 24 were male and two were female. The age, gender and diagnosis of individual children are included in Appendix B.
Given the complexity of diagnosis and the variability of the descriptive labels provided for individual subjects, a symptom checklist was compiled and introduced into the procedure. The three-page checklist aimed to provide a symptom profile for each child in the LD and SP groups encompassing social interaction, social communication and language, social imagination, unusual responses to sensory stimuli, unusual movements and repetitive activities. The Checklist included a large number of features reported in the literature on autism and semantic-pragmatic difficulties, but was informed primarily by Wing (1991) and Andersen-Wood and Smith (in press). A copy of the Checklist along with its analysis is included in Appendix C.

For each subject in the SP and LD groups, the child's SLT or autism outreach worker was asked to complete the Checklist prior to or shortly after the session with the child.

2. MATERIALS

(a) COLLECTION OF IDIOM SAMPLE

The idioms included in the assessment were selected from those that occurred in the course of 31 hours of children's television and/or 21 hours of classroom teaching.

(i) TELEVISION SAMPLE

Over a period of four weeks, 31 hours of children's television was video-recorded. This represented a sample of all weekday children's programmes broadcast across all four terrestrial channels, with the exception of "Programmes for Schools". The
videos were then replayed and a list of all idiomatic expressions used was compiled. In view of the culture- and dialect-bound nature of idioms, it was decided that only idioms that occurred in British programmes would be considered eligible for inclusion. Approximately 1700 different idiomatic expressions were noted in the British television sample.

(ii) TEACHER SAMPLE

In order to obtain a sample of idioms used in routine classroom activity, eleven teachers were each tape-recorded for the duration of one morning's teaching. These included three teachers of mainstream, second-year infant classes in three different schools, two teachers of mainstream, fourth-year junior classes in two different schools and six language unit teachers across four different centres. The language unit teachers taught children across a range of ages. Of the eleven teachers recorded, two were male and nine were female. In view of the absence of evidence for gender-based differences in idiomatic usage, teachers' gender was not controlled for.

At the time of taping, the teachers were all blind to the specific purpose of the recording. They were told that samples of language directed at children in several contexts were being collected in order to establish a data pool for assessing comprehension. In the classroom, teachers told their pupils that a student from the university was interested in finding out what teachers said during lessons.

The teachers were instructed in the use of a tie-pin microphone attached to a small Walkman-style tape-recorder which they carried in a bag or pocket. They were asked to record one morning's routine classroom activities. The length of recordings ranged from thirty minutes to three hours.

The tapes were subsequently scanned for idiomatic expressions.
A list was drawn up of all the idioms used by each teacher. Approximately 500 different idiomatic expressions were noted in the teacher sample.

(b) SELECTION OF IDIOMATIC EXPRESSIONS

The idiom data compiled from the television and teacher samples were initially sorted according to the setting in which they occurred. Idioms were categorised into those that occurred on morning television (aimed specifically at preschool children), on afternoon television, in mainstream infant classes, mainstream junior classes or in language units. Further lists were then drawn up of those idioms which were used by at least half of the teachers in one or both settings, as well as on at least half of the morning-to-lunchtime and/or afternoon television sessions. (These data are included in Appendix D.)

The set of idioms defined by these criteria represented the most commonly-used expressions in the total language sample. Their inclusion in the materials was therefore an attractive option. However, these idioms were revealed to be unsuitable once subjected to the criteria for eligibility.

In order to ensure the validity of the procedure, it was necessary, first of all, that the idioms included had both an idiomatic and literal sense and, secondly, that both of these meanings were clearly demonstrable with props. These criteria ensured that the play task would allow for the child's literal strategy or idiomatic strategy to be demonstrated. For this reason, idioms were excluded if they could only be indirectly demonstrated (e.g. "be on top of the world", "change one's mind"), had closely related literal and idiomatic meanings which might not have been distinguishable when demonstrated (e.g. "be sick", "go to bed", "keep one's eyes open"), or if their figurative sense was difficult to demonstrate (e.g. "run out").
It was further necessary that no expressive language on the part of the child be required. As noted in Chapter 2, children with semantic-pragmatic difficulties are characterised by numerous expressive language problems which might distort the outcome of a task relying on verbal expression. Thus, all idioms which required verbal expression or cessation of speech in the play task were disqualified (e.g. "you can say that again", "shut up"). Similarly, idioms of which young children could not be assumed to have experience were excluded (e.g. "fruit machine"; "shoot up"; "the silver screen").

The application of these criteria to the total language sample yielded a list of 42 eligible idioms. These could be broken down into two further groups on the basis of the nature of their literal counterparts. The idioms in the first group had at least one indisputably literal meaning which was not to be found in either volume of the Oxford Dictionary of Current Idiomatic English (Cowie & Mackin, 1993; Cowie, Mackin & McCaig, 1985) (e.g. "give a hand", "under lock and key"). The idioms in the second group had a counterpart which fell on the border between literality and idiomaticity (as discussed in Chapter 2). These counterparts failed at least one of the three tests for idiomaticity outlined by the ODCIE (and detailed in Chapter 2) but were nevertheless listed in this dictionary. This group included such idioms as "pick up" and "take off".

(c) STORY

From the two categories of eligible idiomatic expressions, twelve idioms were selected and embedded into a 277-word story. The wording of the story was deliberately simple and unambiguous (A copy of the story is included in Appendix E.)

The story was read onto tape by a female, Southern British Standard speaker. It was considered unlikely that this accent would bias any particular children in the subject group. It is,
furthermore, an accent to which most English children are likely to have been exposed through television.

The story was read at normal speed and the reader endeavoured not to stress the idiomatic expressions unnaturally. The story was 1.5 minutes in length. The recording was subsequently duplicated so that it would not be necessary to rewind the tape during the session with the child.

(d) PLAY SET

A play set corresponding to the story was constructed from wood and hardboard and painted with non-toxic paint. The set, measuring 78 cm x 61 cm, included a house and garden, a police station fitted with a cell, the door of which could open and shut, and a vet's surgery fitted with a table. A circular road passed by each of the three buildings. In order to maximize the camcorder's view, the buildings were each constructed with at least two open sides and an open roof. The buildings were stuck to the base board with blu-tack and could not easily be moved.

Additional props, all scaled to size, included a white motor car with open roof, a green motorbike, a dog, a dog's basket, a lock and key and several bandages.

The characters of the story were represented by Playmobil people. Their realism, size and the mobility of their limbs and heads made the Playmobil characters suitable for the acting out of the story by primary school children. It was furthermore expected that most children aged six to eleven would have had some experience of playing with these or similar characters.

Photographs of the play set are included on the following pages.
FIGURE 3: DETAIL OF PLAY SET - HOUSE
Prior to the main study, a pilot study was conducted with five mainstream children aged 6;2 - 6;7 (mean age 6;6, s.d. = 1.6 months), five mainstream children aged 9;2 - 10;8 (mean age 10;0, s.d. = 8 months) and five children identified as having semantic-pragmatic difficulties (aged 6;8 - 9;3, mean age 8;1, s.d. = 14 months). In addition to testing the hypothesis that children with semantic-pragmatic difficulties would display greater difficulty in comprehending these idioms, the primary aims of the pilot study were to validate the methodology and identify a preliminary set of coding categories.

For the pilot study, 16 idiomatic expressions were embedded into a two-minute long, 400-word story. (A copy of the pilot version of the story is included in Appendix E.) Following the pilot, the story was partially modified for the main study. Four idioms were removed and some recontextualization took place. These amendments ensured that all of the idioms included were understood by at least 50% of the 6-year old, mainstream children in the pilot and that the literal and idiomatic actions corresponding to each idiom could be clearly distinguished by the researcher. The story was reduced to 277 words and 1.5 minutes in length (see Appendix E).

The pilot study confirmed the viability and appropriacy of this methodology for exploring the hypotheses and its suitability for use with mainstream children and children with semantic-pragmatic difficulties.
3. PROCEDURE

(a) SETTING

All but two children were assessed individually, in a quiet room in their own schools. The assessments all took place during the school day. One child in each of the MJ and SP groups was assessed in a quiet room at De Montfort University.

(b) INTRODUCTION TO THE TASK

Prior to the assessment, the researcher usually met the child in his or her classroom and introduced herself. Talk en route to the assessment room was designed to relax the child and to introduce the topic of television and videos.

Once in the assessment room, the child was introduced to the camera-person. The child was shown the set and asked whether s/he had played with Playmobil people before. If the child initiated play with the set and props, s/he was allowed to continue briefly before the set was rearranged to begin.

The presentation of the set and props was consistent across children (as in the photographs above). The set was placed on a table of a height which allowed the child to reach all parts of the set without standing up. The child was seated next to the researcher facing the set with the police station and house closest to him/her.

(c) PLAY TASK

As soon as the child was settled, the researcher said, "(Child's name), today we are going to make a video. The video we're going to make is a story about this little girl (indicate Sally), whose
name is Sally, and her dog (indicate Rover) whose name is Rover. This is Sally. This is Sally's mum. They live in this house. This is a vet. He's a doctor who helps animals when they're ill or injured. This is the vet's office. This is a policeman and this is his police station. This is a bike rider and his motorbike has got a piece that you can move (lift the bike and adjust the balancing lever) and that makes it stand (stand the bike as before)." (Indicate each character and prop as appropriate.)

"On this tape (indicate tape recorder), I have a story about Sally and Rover. In a moment, we'll listen to the whole story. Then we'll listen to it again and I'd like you to make all of these people do whatever the story says. First, let's listen to the whole story."

If the child played with the props during the introduction, the researcher said, "You just need to listen now." If necessary, the researcher added, "Put your hands on your knees."

The entire story was played. During this first reading of the story, the researcher noted whether or not the child scanned the play set with his/her eyes. The child's activity during the reading was recorded as "full (scanning)", "none" or "partial (scanning)" (see Appendix F).

When the story was complete, the researcher said, "Now we'll listen to the story again. This time, make the people do whatever the story says. So, if the story says, 'Sally patted Rover', then you make Sally pat Rover." The researcher acted this out unless the child immediately performed the action. Then the researcher said, "Now you try. If the story says 'Sally gave the lock and key to the policeman', then you ... " The researcher paused to let the child perform the action. If s/he did not act, the researcher added "give the lock and key to the policeman". If the child still did not perform the action, the researcher repeated "Sally gave the lock and key to the
"policeman". If the child did not act, the researcher added, "You do it". If the child still did not act, the researcher demonstrated the action.

If the child demonstrated the incorrect action, the researcher repeated "Sally gave the lock and key to the policeman". If the child did not demonstrate the correct action, the researcher moved on to the next trial.

The researcher then said, "If the story says 'The vet gave Sally's Mum a bandage', then you ... " . The researcher performed this trial as outlined for the previous trial item.

If the child had correctly demonstrated both trial items, the researcher proceeded to the story. If not, she continued with as many trials as required from the list below. If the child had failed to manipulate the props at all during the previous trials, the researcher continued with item (i) below. In all other cases, the researcher proceeded to item (ii) below.

(i) Sally patted Rover.
(ii) Sally's mum put Rover in the basket.
(iii) The vet gave the lock and key to Sally.
(iv) Sally's mum gave a bandage to the bike rider.
(v) The policeman put Sally into the cell.

The researcher then asked "Do you understand what I want you to do?" If the child indicated "no", the researcher said, "Make the people do whatever the story says" and continued with further trials.

If the child indicated "yes", the researcher said, "Now we'll listen to the story again. This time, you make the people do whatever the story says. You can use anything you need on this table. You don't need to rush because I'll stop the tape after each bit of the story. Are you ready?"
The story tape was played. The researcher paused the tape after each sentence and waited until the child removed his/her hands from the set, held the props stationary, looked at the tester, said that s/he was finished or was considered by the researcher to have completed the demonstration of the given sentence.

If the child did not do anything after the first sentence had been played, the researcher stopped the tape and prompted "You do it". If the child still failed to demonstrate this and ensuing sentences, the researcher stopped the tape and asked, "Do you understand what I want you to do?" If the child indicated "no", the researcher repeated the explanation and trials. If the child appeared not to understand the requirements of the task, the procedure was terminated at this point.

In all other instances, the researcher continued with the procedure until the story was complete. In the interests of consistency, the researcher's use of prompts was scripted. (This script is included in Appendix G.) When the story was complete, the researcher said, "Well done! That was really good! Would you like to watch your video?"

If the child had indicated "no" and seemed tired or irritable, the researcher would have ended the procedure at this point. In fact each child indicated "yes" (all of the children did) and the researcher moved on to the definition task.

(d) DEFINITION TASK

Following the play task, the television monitor was placed on the table in front of the child and the play set was removed if the child appeared to be distracted by it.

The researcher said, "While we watch your video, I'd like to ask you some questions about the story. Is that all right?" If the child answered "yes" (all of the children did), the researcher
then added, "I want to remember what you say so I've brought this tape-recorder." The researcher showed the tape recorder to the child and asked if she might clip the microphone onto the child's collar or jumper.

The video was then played back to the child. Directly after each sentence containing an idiomatic expression, the researcher paused the tape, repeated the last sentence and then asked "What do you think (the idiom) means?" For example, "The bike rider didn't see Rover and ran over him. What do you think 'ran over' means?" As in the play task, the researcher's prompts were scripted (see Appendix G). When the video ended, the researcher thanked the child for helping her.

(e) FILMING

The video recording was performed by a female technician from the Department of Human Communication at De Montfort University.

In the case of the mainstream children, the camcorder was turned on immediately before the play began. In the sessions involving SP and LD subjects, recording began as soon as the children entered the room. This allowed for subsequent review of the child's response to the introduction and to the first reading of the story.

In all cases, the camcorder was turned off at the end of the second reading of the story. No filming took place during the definition task.

During the filming of the play task, the camcorder focused primarily on the child's hands. This decision was based on the pilot study finding that keeping the child's face in the frame resulted in the loss of important detail with respect to the child's actions.
Teacher recordings were made on a Panasonic RQ-L335 (Walkman) tape-recorder, fitted with an ECM-144 tie-pin microphone and standard audio tapes. The same equipment was used to play the story to the children and to record their responses to the questions asked in the definition task.

For filming, a Panasonic M10 camcorder on a tripod with standard VHS tapes and a Panasonic Quintrix TC800G monitor were used.

A television and standard VHS video recorder were used to play back the video tapes for analysis. The audio tapes were played back for transcription on a Sony Walkman Professional Stereo Cassette Recorder WM-D6C.

(g) TRANSCRIPTION AND CODING

(i) PLAY TASK

As soon as possible following the session, the video tape was played back and transcribed. It was noted that the position of the props and the action immediately preceding a given idiom had a significant bearing on how the child's demonstration of that idiom was interpreted. It was thus decided that the transcript should include both the action immediately preceding the idiom (marked on the transcript with the symbol #) and the child's action from the moment of the idiom until the beginning of the following utterance (marked on the transcript with the symbol *).

The transcription took the form of a detailed description of the child's actions and of any verbal utterances. The researcher's utterances were included. (An example is included in Appendix H.)

After each session had been transcribed in this manner, the
responses corresponding to each idiom in turn were collated. The analysis of the data by idiom rather than by child aimed to prevent the recognition during scoring of within-child patterns of response. Any such recognition at this stage might have influenced the scoring of a particular child's subsequent responses.

A preliminary set of scoring codes had been developed on the basis of the pilot study data. In analyzing the larger data sample in the main study, two further codes were introduced in order that all of the data could be categorised.

Each subject's response to each idiom was considered as a single item. There were therefore 71 items for each of the 12 idioms. Each item received one of the eleven codes outlined below.

**IDIOMATIC**
The action clearly demonstrates the target idiomatic sense of the expression. This may involve a single action or a series of actions.

**INFERRED IDIOMATIC**
The action portrays the child's grasp of the idiomatic sense of the expression but falls short of the requirements for a coding of "idiomatic". For example, the child demonstrates the appropriate action using the incorrect character.

**LITERAL**
The action clearly demonstrates the literal sense of the expression.

**INFERRED LITERAL**
The action portrays the child's grasp of the literal sense of the expression but falls short of the full requirements for a coding of "literal".
OTHER FIGURATIVE SENSE

The action portrays a grasp of an alternative idiomatic sense of the expression or of a word within the expression. For example, the expression "give a hand" has two idiomatic senses (help and clap) and one literal sense (donating or passing one's hand to somebody). In the sentence, "Sally gave her mum a hand to carry Rover to the white car", the target sense is that of "helping". An action depicting helping will therefore be coded as "idiomatic". An action portraying the child attempting to remove a hand from a character's body will be coded as "literal". An action portraying a character clapping will be coded as "other figurative sense".

AMBIGUOUS - IDIOMATIC/LITERAL

(1) The action can be interpreted as reflecting either the "(inferred) idiomatic" or "(inferred) literal" sense of the expression, or
(2) The action includes a portrayal of both the "(inferred) idiomatic" sense and the "(inferred) literal" sense.

AMBIGUOUS - LITERAL/OTHER FIGURATIVE SENSE

(1) The action can be interpreted as reflecting either the "(inferred) literal" sense or the "other figurative sense" of the expression, or
(2) The action includes a portrayal of both the "(inferred) literal" sense and the "other figurative sense" of the expression.

AMBIGUOUS - IDIOMATIC/OTHER FIGURATIVE SENSE

(1) The action can be interpreted as reflecting either the "(inferred) idiomatic" sense or the "other figurative sense", or
(2) The action includes a portrayal of both the "(inferred) idiomatic" sense and the "other figurative sense".
APPROPRIATE NO ACTION
Prior to the expression being played, the props are already appropriately positioned to portray the idiomatic sense of the expression. The child does not act on the props at all.

INAPPROPRIATE NO ACTION
Prior to the expression being played, the props are not appropriately positioned to portray the idiomatic sense of the expression. The child does not act on the props at all.

FUZZY
The action appears to be unrelated or only questionably related to the "(inferred) idiomatic" sense, the "other figurative sense" or the "(inferred) literal" sense of the expression.

More detailed guidelines for the coding of each idiom included criteria for an acceptable idiomatic demonstration and an acceptable literal demonstration of the individual expressions. Where appropriate, the guidelines for individual expressions also included principles by which given codes should be distinguished. The complete guidelines for the coding of the play task actions are included in Appendix I. An example of coded data is included in Appendix H.

(ii) DEFINITION TASK
As soon as possible after the session, the children's definitions were transcribed from the audio tapes. As in the case of the play task, the data for all the children were then arranged by idiom.

Each child's definition of each expression was considered to be a single item. As in the play task, there were 71 items for each idiom. These each received one of the 13 possible codes described below.
IDIOMATIC
The child defines the target idiomatic sense of the expression.

INFERRED IDIOMATIC
The definition allows one to infer that the child has understood the expression in its idiomatic sense, but fails to meet the criteria for a coding of "idiomatic".

LITERAL
The child defines the literal sense of the expression.

INFERRED LITERAL
The definition allows one to infer that the child has understood the literal sense of the expression, but fails to meet the criteria for a coding of "idiomatic". This inference may be based on the child's analogy or emphasis.

OTHER FIGURATIVE SENSE
The definition relates to an alternative idiomatic sense of the expression or of a word within it.

AMBIGUOUS - IDIOMATIC/LITERAL
(1) The definition can be interpreted as reflecting either the "(inferred) idiomatic" or "(inferred) literal" sense of the expression, or
(2) The response includes a definition of both the "(inferred) idiomatic" sense and the "(inferred) literal" sense of the expression.

AMBIGUOUS - LITERAL/OTHER FIGURATIVE SENSE
(1) The definition is compatible with either the "(inferred) literal" sense or the "other figurative sense" of the expression, or
(2) The response includes a definition of both the "(inferred) literal" sense and the "other figurative sense" of the expression.
AMBIGUOUS - IDIOMATIC/OTHER FIGURATIVE SENSE
(1) The definition is compatible with either the target "idiomatic" sense or the "other figurative sense" of the expression, or
(2) The response includes a definition of both the target "idiomatic" sense and the "other figurative sense" of the expression.

REPEAT
The child repeats the exact words of the idiom and/or other words in the sentence quoted by the researcher in the question. The child may change the morphology of the words as long as this does not change their meaning eg. the child may change the tense of a verb, change a verb to its participle form or change a singular noun to its plural form.

For example, in response to "The vet moved them up so there was enough space for Rover. What do you think 'moved up' means?", any of the following would be coded "repeat": "moved up", "moving up", move up", "enough space", "spaces".

NARRATIVE
(1) The child paraphrases part/all of the sentence quoted by the researcher in the question, or
(2) The child retells another part of the story, or
(3) The child proposes a motivation for or explanation of a character's action in any part of the story.

REPEAT-NARRATIVE
The child's response includes features of both "repeat" and "narrative" as defined above. For example, in response to "The vet moved them up so there was enough space for Rover. What do you think moved up means?", the response "move things up" would be coded "repeat-narrative".

118
FUZZY
The child's meaning is very unclear and appears to be unrelated or only questionably related to the "idiomatic" sense, the "literal" sense or the "other figurative sense", and does not qualify for a code of "repeat", "narrative" or "repeat-narrative".

DON'T KNOW
The child responds only with "I don't know", "I don't remember" or silence.

As for the play task, more detailed guidelines defined the necessary criteria for a code of "idiomatic" and "literal", respectively for each of the twelve expressions. The guidelines also provided principles by which to distinguish codes for specific expressions, where appropriate. The complete guidelines for the coding of the definitions are included in Appendix I. An example of coded data is included in Appendix H.

(h) RELIABILITY AND VALIDITY

(i) RELIABILITY OF THE CODING PROCEDURE

Four types of reliability were considered in relation to the coding of data in this study: test-retest reliability, split-half reliability, equivalence reliability and inter-rater reliability. Test-retest reliability might have been established by the recoding of the data by the researcher after a suitable time interval. Unfortunately, time constraints prevented the researcher from introducing a sufficient time gap between the data collection and coding to allow for the identity of group membership to fade in her memory.

Split-half reliability was considered an inappropriate measure
for this data since a considerable degree of between-idioms variation was expected. Equivalence reliability, involving comparison across two parallel forms of the same test, was not suitable for the same reason.

Inter-rater reliability was considered the only appropriate measure of the reliability of the coding procedure. It was particularly desirable that the data be coded blind by a coder unfamiliar with both the aims of the study and the clinical population investigated. To this end, an independent coder - an academic psychologist - was introduced. The procedure was explained and a copy of the guidelines for coding (included in Appendix I) was provided. The second-coder coded a random sample comprising 20% of the total data sample proportionally selected across each subject group. The second-coder thus coded 15 items for each of the twelve expressions (three items for the MI, MJ and LD groups, respectively, and six items for the SP group). The items for each idiom were presented in random order with all identifying data omitted.

Comparison of the first and second codings yielded 82.2% agreement on the play task and 78.9% agreement on the definition task. The number and proportion of items on which the two coders agreed are presented by idiom in Table 1 and by group in Table 2.

Subsequent to the second coding, the items on which the coders differed were all negotiated to consensus. Discussion of these items revealed that a large proportion of disagreement resulted either from ambiguities and lack of clarity in the transcript or from the second coder having misread the guidelines. Of the 32 play task disagreements, eleven (34.4%) arose from lack of clarity or ambiguity in the transcript and eight (25%) resulted from a misreading of the guidelines on the part of the second coder. Of the 38 definition task disagreements, ten (31.3%) arose from the second coder's misreading of the guidelines. The levels of inter-coder
agreement excluding items on which disagreement arose from transcript ambiguity or misreadings are presented in Tables 1 and 2 below.

**TABLE 1: BETWEEN-IDIOMS ANALYSIS OF INTER-CODER AGREEMENT**

*(The second coder coded 15 items for each idiom in each task.)*

<table>
<thead>
<tr>
<th>IDIOM</th>
<th>WHOLE SAMPLE</th>
<th></th>
<th>SAMPLE EXCLUDING DISAGREEMENTS ARISING FROM TRANSCRIPT AMBIGUITY &amp; MISREADING OF GUIDELINES</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PLAY TASK</td>
<td>DEFINITION TASK</td>
<td>PLAY TASK</td>
<td>DEFINITION TASK</td>
</tr>
<tr>
<td>RAN OVER</td>
<td>13</td>
<td>13</td>
<td>15</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>86.7%</td>
<td>86.7%</td>
<td>100%</td>
<td>93.3%</td>
</tr>
<tr>
<td>TOOK OFF</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>86.7%</td>
<td>86.7%</td>
<td>86.7%</td>
<td>86.7%</td>
</tr>
<tr>
<td>GAVE A HAND</td>
<td>14</td>
<td>15</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>93.3%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>MOVED UP</td>
<td>12</td>
<td>11</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>80%</td>
<td>73.3%</td>
<td>80%</td>
<td>73.3%</td>
</tr>
<tr>
<td>GOT ON TO</td>
<td>13</td>
<td>10</td>
<td>14</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>86.7%</td>
<td>66.7%</td>
<td>93.3%</td>
<td>66.7%</td>
</tr>
<tr>
<td>ON THE SPOT</td>
<td>14</td>
<td>11</td>
<td>15</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>93.3%</td>
<td>73.3%</td>
<td>100%</td>
<td>86.7%</td>
</tr>
<tr>
<td>TURNED OVER</td>
<td>11</td>
<td>14</td>
<td>14</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>73.3%</td>
<td>93.3%</td>
<td>93.3%</td>
<td>100%</td>
</tr>
<tr>
<td>UNDER LOCK &amp; KEY</td>
<td>15</td>
<td>11</td>
<td>15</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>100%</td>
<td>73.3%</td>
<td>100%</td>
<td>80%</td>
</tr>
<tr>
<td>PICK UP</td>
<td>5</td>
<td>14</td>
<td>11</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>33.3%</td>
<td>93.3%</td>
<td>73.3%</td>
<td>93.3%</td>
</tr>
<tr>
<td>KEPT AN EYE ON</td>
<td>13</td>
<td>13</td>
<td>14</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>86.7%</td>
<td>86.7%</td>
<td>93.3%</td>
<td>100%</td>
</tr>
<tr>
<td>DROPPED IN</td>
<td>11</td>
<td>12</td>
<td>15</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>73.3%</td>
<td>80%</td>
<td>100%</td>
<td>80%</td>
</tr>
<tr>
<td>WENT ROUND</td>
<td>14</td>
<td>5</td>
<td>14</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>93.3%</td>
<td>33.3%</td>
<td>93.3%</td>
<td>53.3%</td>
</tr>
<tr>
<td>TOTAL (OUT OF 180)</td>
<td>148</td>
<td>142</td>
<td>167</td>
<td>152</td>
</tr>
<tr>
<td></td>
<td>82.2%</td>
<td>78.9%</td>
<td>92.8%</td>
<td>84.4%</td>
</tr>
</tbody>
</table>
The concordance rates presented here were considered to represent an acceptable level of inter-rater reliability. Given that the researcher (the first coder) was unable to code the data blind, it is particularly noteworthy that the proportion of items on which the coders disagreed was not significantly greater for any one group. The author and second coder were easily able to reach consensus on the remaining items. This result indicates that the absence of blind coding in the first round did not introduce any significant bias into the coding process.

As evidenced in Table 1, the idioms "got on to" and "went round" were found to have attracted a lower level of inter-coder agreement on the definition task than did the other idioms. In many instances, the disagreement regarding these idioms reflected a difference of degree rather than of type. For example, one coder may have coded an item "idiomatic", whereas the other coded it "inferred idiomatic". The relatively low levels of agreement

<table>
<thead>
<tr>
<th>GROUP</th>
<th>WHOLE SAMPLE</th>
<th>SAMPLE EXCLUDING DISAGREEMENTS ARISING FROM TRANSCRIPT AMBIGUITY AND MISREADING OF GUIDELINES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PLAY TASK</td>
<td>DEFINITION TASK</td>
</tr>
<tr>
<td></td>
<td>MI</td>
<td>30  83.3%</td>
</tr>
<tr>
<td>(36 items)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MJ</td>
<td>33  91.7%</td>
</tr>
<tr>
<td>(36 items)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>LD</td>
<td>28  77.8%</td>
</tr>
<tr>
<td>(36 items)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SP</td>
<td>57  79.2%</td>
</tr>
<tr>
<td>(72 items)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TOTAL</td>
<td>148  82.2%</td>
</tr>
<tr>
<td>(180 items)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
for these two idioms does, however, reflect general coding
difficulties with regard to the definitions of these two idioms
(see between-idioms discussion in Chapter 6).

Where inter-coder differences occurred, the code used in the
analysis was that agreed upon by both coders in subsequent
discussion.

(ii) RELIABILITY OF THE TASKS AS A MEASURE OF IDIOM
COMPREHENSION

The four forms of reliability outlined above were also considered
in relation to the test as a whole. Once again, three of these
were not considered suitable for this data. Test-retest
reliability cannot validly be measured for learned knowledge such
as idioms (Mouton & Marais, 1988). Split-half reliability was
unsuitable for use in relation to a test of idiom comprehension
since between-idioms variation is reasonably expected. For the
same reason, it was not possible to measure equivalence
reliability with a parallel form of the test.

Given the academic context in which this study was conducted, it
was necessary that all the testing be conducted solely by the
researcher. It was therefore not possible to compare data
collected by different researchers.

(iii) VALIDITY OF THE TEST

Consideration of validity focused on content validity, construct
validity and criterion-related validity. Content validity
requires that the "content of this measure (be) representative
of the content or the universe of content of the property being
measured" (Kerlinger, 1986, 417). As discussed in Chapter 2, the
target expressions used in this study all met a widely-accepted
definition of idioms. Their idiomatic status was further
validated by their inclusion in the Oxford Dictionary of Current Idiomatic English (Cowie & Mackin, 1993; Cowie, Mackin & McCaig, 1985). Having been extracted from samples of teacher and children's television language, the idioms can validly be claimed to be ones to which children are exposed. Although a range of idiom types has been represented in this procedure, no claims are made regarding the representativeness of this set of idioms in relation to the universe of idioms.

Construct validity "refers to the degree to which a test measures the theoretical construct it is intended to measure" (Anastasi, 1976 cited in McCauley & Swisher, 1984). In the case of comprehension, it not possible to measure the construct directly but only to observe the behaviour reflecting comprehension or lack of comprehension. Like all tests of comprehension, the tasks used in this study are likely to tap skills in some areas other than idiom comprehension (eg. symbolic play or attention). As noted above, the definition task was employed in the knowledge that its expressive and metalinguistic demands were likely to affect the outcome of the task. In comparison with the methodologies of previous tests and studies, the play task is, however, considered to involve fewer potentially interfering variables. As such, the data that it yields are assumed to be a more accurate reflection of idiom comprehension than is that produced in the definition task.

As discussed in Chapter 2, the validity of the currently-available formal tests of idiom comprehension is largely questionable. As such, it was not possible to test the criterion-related validity of this procedure by comparing the results with those obtained by the same subjects on a recognised test of idiom comprehension. Wiig and Secord (1989) assert, however, that

an alternative to using another standardized test as a criterion measure is to use an a priori classification of students by a related variable (such as) the school district's classification of students as language-learning disabled... or normally achieving" (TLC-E Technical Manual, p. 50).
Given that the LD and SP subjects in this study were thus classified prior to the study by outside clinicians, the fact that both the play task and the definition task discriminated them with a high degree of accuracy (as discussed below) can thus be considered as evidence of the criterion-related validity of these procedures. That the MJ and MI children were also discriminated from each other and from the two clinical populations further supports this conclusion.

(i) ANALYSIS OF DATA

The codes outlined above for each task are referred to in this report as "narrow codes". While providing useful detail about idiom comprehension, the narrow code results are difficult to compare with the findings of previous studies which have scored subject responses in terms of broader categories (see Chapter 2). The difference in narrow codes for the play and definition tasks also makes it difficult to compare performance across tasks. Given that the narrow codes do in fact group together into four broad codes, it was decided that both narrow and broad levels of analysis should be conducted. The relationship between the narrow and broad codes is presented in Table 3 below.

As evidenced in Table 3, the code "other figurative sense" (OFS) is considered to be both a broad and narrow level code.
### Table 3: Relationship Between Broad and Narrow Codes

<table>
<thead>
<tr>
<th>Broad Code</th>
<th>Narrow Codes: Play Task</th>
<th>Narrow Codes: Definition Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;Appropriate&quot;</td>
<td>Idiomatic</td>
<td>Idiomatic</td>
</tr>
<tr>
<td></td>
<td>Inferred idiomatic</td>
<td>Inferred idiomatic</td>
</tr>
<tr>
<td></td>
<td>Appropriate no action</td>
<td></td>
</tr>
<tr>
<td>&quot;Inappropriate&quot;</td>
<td>Literal</td>
<td>Literal</td>
</tr>
<tr>
<td></td>
<td>Inferred literal</td>
<td>Inferred literal</td>
</tr>
<tr>
<td></td>
<td>Fuzzy</td>
<td>Fuzzy</td>
</tr>
<tr>
<td></td>
<td>Inappropriate no action</td>
<td>Repeat</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Repeat-narrative</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Don't know</td>
</tr>
<tr>
<td>Ambiguous</td>
<td>Ambiguous - idiomatic/literal</td>
<td>Ambiguous - idiomatic/literal</td>
</tr>
<tr>
<td></td>
<td>Ambiguous - idiomatic/OFS</td>
<td>Ambiguous - idiomatic/OFS</td>
</tr>
<tr>
<td></td>
<td>Ambiguous - literal/OFS</td>
<td>Ambiguous - literal/OFS</td>
</tr>
<tr>
<td>Other figurative sense (OFS)</td>
<td>Other figurative sense (OFS)</td>
<td>Other figurative sense (OFS)</td>
</tr>
</tbody>
</table>

It should be noted that the broad code terms "appropriate" and "inappropriate" are used here in a more specific sense than is generally the case in the semantic-pragmatic difficulties literature (see for example, Leinonen & Smith, 1994; Bishop & Adams, 1989; Prutting & Kirchner, 1983). In this report, an "appropriate" item (in inverted commas) refers to an item coded "idiomatic", "inferred idiomatic" or "appropriate no action". Similarly, an "inappropriate" item is one that has been assigned any of the corresponding narrow codes.

Once coded, the data for each of the two tasks were subjected to a series of statistical analyses. These were administered on both the broad and narrow codes. Between-groups and between-subgroups analyses were conducted using the Mann-Whitney Test and Kth neighbour discriminant analysis. For between-tasks analyses,
the Wilcoxon Signed Rank Test was employed. Spearman's Rank Correlation was used to examine the relationship between task performance and age, TROG scores and number of play task trials, respectively. It was also used to compare idiom rankings across tasks. Idiom rankings across groups were compared with Kendall's Concordance Coefficient. Finally, Point Biserial Correlation was employed to investigate the relationship between play task performance and individual features on the Checklist.

(j) DETAILED HYPOTHESES

The method of obtaining and analyzing data outlined above allowed for the three main hypotheses presented in Chapter 3 to be explored through the testing of the more detailed hypotheses listed below. (The subgroups referred to in association with the SP group are the "SPD" and "ASP" subgroups introduced in Chapter 2.)

BETWEEN-GROUPS HYPOTHESES

(i) The MJ group will score more "appropriate" items and "idiomatic" items on both tasks than will the MI group.
(ii) The MJ group will score fewer "inappropriate" items on the definition task than will the MI group.
(iii) The MJ group will score more "appropriate" items and more "idiomatic" items on both tasks than will the LD group.
(iv) The MJ group will score fewer "inappropriate" items on the definition task than will the LD group.
(v) The MI group will score more "appropriate" items on the definition task than will the LD group.
(vi) The MI group will score fewer "inappropriate" items on the definition task than will the LD group.
(vii) The SP group (and subgroups) will score fewer "appropriate" items and "idiomatic" items on both tasks than will the MI or MJ group.
(viii) The SP group (and subgroups) will score more "inappropriate" items and "literal" items on both tasks than will the MI or MJ group.

(ix) The SP group (and subgroups) will score fewer "appropriate" items and "idiomatic" items on the play task than will the LD group.

(x) The SP group (and subgroups) will score more "inappropriate" items and "literal" items on the play task than will the LD group.

BETWEEN-TASKS HYPOTHESES

(xi) The number of "appropriate" items scored by any group on the play task will be greater than the number scored by that group on the definition task.

(xii) The number of "inappropriate" items scored by any group on the play task will be smaller than the number scored by that group on the definition task.

(xiii) The MI and MJ groups will both score more "appropriate" items than "inappropriate" items on both tasks.

(xiv) The LD group will score more "appropriate" items than "inappropriate" items on the play task.

(xv) The SP group (and subgroups) will score more "inappropriate" items than "appropriate" items on both tasks.
This chapter presents the results of the play task and of the definition task. Between-groups results of the play task are presented first, followed by between-groups results of the definition task. The between-tasks analysis is then presented. This is followed by within-group results and then by an analysis of the results for the subgroups "SPD" and "ASP". A between-idioms analysis is then presented. Finally, several miscellaneous analyses are included.

The results are presented for two levels of analysis. The broad analysis is concerned with the codes "appropriate", "inappropriate", "ambiguous" and "other figurative sense". With the exception of "other figurative sense", each of these broad codes breaks down into a series of narrow codes as described in Chapter 4.
1. PLAY TASK: BETWEEN-GROUPS ANALYSIS

(a) BROAD ANALYSIS

The broad level results for the four groups (MJ, MI, LD and SP) are summarised in Table 4 below and are represented graphically in Figure 6. Since only one subject scored "other figurative sense" on the play task, this code is not represented on Figure 6.

TABLE 4: PLAY TASK - BETWEEN-GROUPS BROAD CODES RESULTS

<table>
<thead>
<tr>
<th>GROUP</th>
<th>APPROPRIATE</th>
<th>INAPPROPRIATE</th>
<th>AMBIGUOUS</th>
<th>OTHER FIG'ITIVE SENSE</th>
<th>TOTAL NO. ITEMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MJ</td>
<td>147</td>
<td>15</td>
<td>18</td>
<td>180</td>
<td>180</td>
</tr>
<tr>
<td></td>
<td>81.7%</td>
<td>8.3%</td>
<td>10%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>x = 9.80</td>
<td>x = 1</td>
<td>x = 1.20</td>
<td>s. d. = 1.32</td>
<td></td>
</tr>
<tr>
<td>MI</td>
<td>137</td>
<td>17</td>
<td>25</td>
<td>180</td>
<td>180</td>
</tr>
<tr>
<td></td>
<td>76.1%</td>
<td>9.4%</td>
<td>13.9%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>x = 9.13</td>
<td>x = 1.13</td>
<td>x = 1.67</td>
<td>s. d. = 1.11</td>
<td></td>
</tr>
<tr>
<td></td>
<td>s. d. = 1.77</td>
<td>s. d. = 1.19</td>
<td>s. d. = 0.07</td>
<td>s. d. = 0.26</td>
<td></td>
</tr>
<tr>
<td>LD</td>
<td>134</td>
<td>18</td>
<td>28</td>
<td>180</td>
<td>180</td>
</tr>
<tr>
<td></td>
<td>74.4%</td>
<td>10%</td>
<td>15.6%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>x = 8.93</td>
<td>x = 1.20</td>
<td>x = 1.87</td>
<td>s. d. = 0.99</td>
<td></td>
</tr>
<tr>
<td></td>
<td>s. d. = 1.39</td>
<td>s. d. = 1.01</td>
<td>s. d. = 0.09</td>
<td>s. d. = 0.09</td>
<td></td>
</tr>
<tr>
<td>SP</td>
<td>190</td>
<td>78</td>
<td>44</td>
<td>312</td>
<td>312</td>
</tr>
<tr>
<td></td>
<td>60.9%</td>
<td>25%</td>
<td>14.1%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>x = 7.31</td>
<td>x = 3</td>
<td>x = 1.69</td>
<td>s. d. = 1.32</td>
<td></td>
</tr>
<tr>
<td></td>
<td>s. d. = 2.33</td>
<td>s. d. = 2.02</td>
<td>s. d. = 0.82</td>
<td>s. d. = 0.12</td>
<td></td>
</tr>
<tr>
<td>TOTAL*</td>
<td>608</td>
<td>128</td>
<td>115</td>
<td>852</td>
<td>852</td>
</tr>
<tr>
<td></td>
<td>71.4%</td>
<td>15%</td>
<td>13.5%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>x = 8.56</td>
<td>x = 1.80</td>
<td>x = 1.62</td>
<td>s. d. = 1.14</td>
<td></td>
</tr>
<tr>
<td></td>
<td>s. d. = 2.08</td>
<td>s. d. = 1.73</td>
<td>s. d. = 0.12</td>
<td>s. d. = 0.12</td>
<td></td>
</tr>
</tbody>
</table>

APPROPRIATE includes "idiomatic", "inferred idiomatic", "appropriate no action"
INAPPROPRIATE includes "literal", "inferred literal", "fuzzy", "inappropriate no action"

Percentages reflect the proportion of the total number of items for the given group (i.e. total number of items including those coded "appropriate", "inappropriate", "ambiguous" and "other figurative sense"): 180 for MJ, MI and LD, respectively, and 312 for SP.

* The percentages in this row reflect the proportion of the total number of items for all the groups summed i.e. 852.
FIGURE 6

PLAY TASK - GROUP RESULTS (BROAD CODES)

GROUP

- APPROPRIATE
- INAPPROPRIATE
- AMBIGUOUS

PERCENTAGE OF ITEMS

100 90 80 70 60 50 40 30 20 10 0

MU  MI  LD  SP
The differences between groups on the broad codes were tested with a series of Mann-Whitney Tests. The results are presented in Table 5 below.

<table>
<thead>
<tr>
<th>GROUPS COMPARED</th>
<th>APPROPRIATE</th>
<th>INAPPROPRIATE</th>
<th>AMBIGUOUS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MJ - MI</td>
<td>U = 95</td>
<td>U = 108.5</td>
<td>U = 87</td>
</tr>
<tr>
<td></td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td>MJ - LD</td>
<td>U = 76</td>
<td>U = 100</td>
<td>U = 75</td>
</tr>
<tr>
<td></td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td>MJ - SP</td>
<td>U = 75</td>
<td>U = 76.5</td>
<td>U = 159.5</td>
</tr>
<tr>
<td></td>
<td>p = .0005</td>
<td>p = .0006</td>
<td>n.s.</td>
</tr>
<tr>
<td>MI - LD</td>
<td>U = 96.5</td>
<td>U = 104</td>
<td>U = 102</td>
</tr>
<tr>
<td></td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td>MI - SP</td>
<td>U = 105</td>
<td>U = 84</td>
<td>U = 189</td>
</tr>
<tr>
<td></td>
<td>p = .0070</td>
<td>p = .0012</td>
<td>n.s.</td>
</tr>
<tr>
<td>LD - SP</td>
<td>U = 108.5</td>
<td>U = 87</td>
<td>U = 170.5</td>
</tr>
<tr>
<td></td>
<td>p = .0091</td>
<td>p = .0015</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

These results indicate that the SP group scored significantly fewer "appropriate" items and significantly more "inappropriate" items than did the MJ, MI and LD groups, respectively. The MJ and MI groups, the MJ and LD groups, and the MI and LD groups, respectively, did not differ significantly in the frequency with which they scored either "appropriate" or "inappropriate". (While not reaching significance, the MJ-LD difference on "appropriate" scores did come close to it (p = 0.0606) with the MJ group scoring more "appropriate" items than did the LD group.)

No group differed significantly from any other group in the number of items scored "ambiguous".

132
(b) NARROW ANALYSIS

Group results on each of the narrow codes making up the broad code "appropriate" are summarised in Table 6 below. (These codes are henceforth referred to collectively as the "'appropriate' narrow codes").

TABLE 6: PLAY TASK - BREAKDOWN OF "APPROPRIATE" ITEMS INTO "APPROPRIATE" NARROW CODES

<table>
<thead>
<tr>
<th>GROUP</th>
<th>IDIOMATIC</th>
<th>INFERRED IDIOMATIC</th>
<th>APPROPRIATE NO ACTION</th>
<th>TOTAL NO. APPROPRIATE ITEMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MJ</td>
<td>131</td>
<td>14</td>
<td>2</td>
<td>147</td>
</tr>
<tr>
<td></td>
<td>72.6%</td>
<td>7.6%</td>
<td>1.1%</td>
<td>81.7%</td>
</tr>
<tr>
<td></td>
<td>x = 8.73</td>
<td>x = 0.93</td>
<td>x = 0.13</td>
<td>x = 9.80</td>
</tr>
<tr>
<td></td>
<td>s. d. = 1.22</td>
<td>s. d. = 1.16</td>
<td>s. d. = 0.52</td>
<td>s. d. = 1.32</td>
</tr>
<tr>
<td>MI</td>
<td>115</td>
<td>20</td>
<td>2</td>
<td>137</td>
</tr>
<tr>
<td></td>
<td>63.9%</td>
<td>11.1%</td>
<td>1.1%</td>
<td>76.1%</td>
</tr>
<tr>
<td></td>
<td>x = 7.67</td>
<td>x = 1.33</td>
<td>x = 0.13</td>
<td>x = 9.13</td>
</tr>
<tr>
<td></td>
<td>s. d. = 1.92</td>
<td>s. d. = 1.05</td>
<td>s. d. = 0.35</td>
<td>s. d. = 1.77</td>
</tr>
<tr>
<td>LD</td>
<td>110</td>
<td>21</td>
<td>3</td>
<td>134</td>
</tr>
<tr>
<td></td>
<td>61.1%</td>
<td>11.7%</td>
<td>1.7%</td>
<td>74.4%</td>
</tr>
<tr>
<td></td>
<td>x = 7.33</td>
<td>x = 1.40</td>
<td>x = 0.20</td>
<td>x = 8.93</td>
</tr>
<tr>
<td></td>
<td>s. d. = 1.99</td>
<td>s. d. = 1.06</td>
<td>s. d. = 0.41</td>
<td>s. d. = 1.39</td>
</tr>
<tr>
<td>SP</td>
<td>157</td>
<td>32</td>
<td>1</td>
<td>190</td>
</tr>
<tr>
<td></td>
<td>50.3%</td>
<td>10.3%</td>
<td>0.3%</td>
<td>60.9%</td>
</tr>
<tr>
<td></td>
<td>x = 6.04</td>
<td>x = 1.23</td>
<td>x = 0.04</td>
<td>x = 7.31</td>
</tr>
<tr>
<td></td>
<td>s. d. = 2.27</td>
<td>s. d. = 1.21</td>
<td>s. d. = 0.20</td>
<td>s. d. = 2.33</td>
</tr>
<tr>
<td>TOTAL*</td>
<td>513</td>
<td>87</td>
<td>8</td>
<td>608</td>
</tr>
<tr>
<td></td>
<td>60.2%</td>
<td>10.2%</td>
<td>0.9%</td>
<td>71.4%</td>
</tr>
<tr>
<td></td>
<td>x = 7.23</td>
<td>x = 1.23</td>
<td>x = 0.11</td>
<td>x = 8.56</td>
</tr>
<tr>
<td></td>
<td>s. d. = 2.17</td>
<td>s. d. = 1.12</td>
<td>s. d. = 0.36</td>
<td>s. d. = 2.08</td>
</tr>
</tbody>
</table>

Percentages reflect the proportion of the total number of items for the given group (i.e. total number of items including those coded "appropriate", "inappropriate", "ambiguous" and "other figurative sense"): 180 for MJ, MI and LD, respectively, and 312 for SP.

* The percentages in this row reflect the proportion of the total number of items for all the groups summed i.e. 852.
Between-groups differences on the "appropriate" narrow codes were tested with a series of Mann-Whitney Tests. Table 7 presents the results for the "idiomatic" and "inferred idiomatic" codes. Differences between groups for the code "appropriate no action" were not tested since it was clear from the raw scores that these would not reach significance.

**TABLE 7: PLAY-TASK - BETWEEN-GROUPS DIFFERENCES ON "APPROPRIATE" NARROW CODES (Mann-Whitney Test)**

<table>
<thead>
<tr>
<th>GROUPS COMPARED</th>
<th>IDIOMATIC</th>
<th>INFERRED IDIOMATIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>MJ - MI</td>
<td>U = 80.5</td>
<td>U = 84.5</td>
</tr>
<tr>
<td></td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td>MJ - LD</td>
<td>U = 63.5</td>
<td>U = 81</td>
</tr>
<tr>
<td></td>
<td>p = .0186</td>
<td>n.s.</td>
</tr>
<tr>
<td>MJ - SP</td>
<td>U = 68</td>
<td>U = 166</td>
</tr>
<tr>
<td></td>
<td>p = .0003</td>
<td>n.s.</td>
</tr>
<tr>
<td>MI - LD</td>
<td>U = 98</td>
<td>U = 109.5</td>
</tr>
<tr>
<td></td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td>MI - SP</td>
<td>U = 115</td>
<td>U = 179</td>
</tr>
<tr>
<td></td>
<td>p = .0145</td>
<td>n.s.</td>
</tr>
<tr>
<td>LD - SP</td>
<td>U = 128</td>
<td>U = 172.5</td>
</tr>
<tr>
<td></td>
<td>p = .0337</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

The results indicate that the SP group scored significantly fewer "idiomatic" items than did any of the other groups. There was no significant difference in the number of items scored "inferred idiomatic" by the SP group and any of the other groups.

The LD group was found to score significantly fewer "idiomatic" items than did the MJ group. The difference between LD and MI on "idiomatic" was not significant. The LD group did not differ significantly from the MJ or MI groups on "inferred idiomatic".

The MJ and MI groups did not differ significantly on "idiomatic" or "inferred idiomatic" codes.
Group results on each of the narrow codes making up the broad code "inappropriate" are presented in Table 8 below. (These codes are henceforth referred to collectively as the "inappropriate narrow codes").

### TABLE 8: PLAY TASK - BREAKDOWN OF "INAPPROPRIATE" ITEMS INTO "INAPPROPRIATE" NARROW CODES

<table>
<thead>
<tr>
<th>GROUP</th>
<th>LITERAL</th>
<th>INFERRED LITERAL</th>
<th>INAPPROPRIATE NO ACTION</th>
<th>FUZZY</th>
<th>TOTAL NO. INAPPROPR. ITEMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MJ</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>6</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>3.3%</td>
<td>1.1%</td>
<td>0.6%</td>
<td>3.3%</td>
<td>8.3%</td>
</tr>
<tr>
<td></td>
<td>x = 0.40</td>
<td>x = 0.13</td>
<td>x = 0.07</td>
<td>x = 0.40</td>
<td>1.1%</td>
</tr>
<tr>
<td></td>
<td>s. d. = 0.51</td>
<td>s. d. = 0.35</td>
<td>s. d. = 0.26</td>
<td>s. d. = 0.74</td>
<td>s. d. = 1</td>
</tr>
<tr>
<td>MI</td>
<td>7</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>3.9%</td>
<td>1.1%</td>
<td>2.2%</td>
<td>2.2%</td>
<td>9.4%</td>
</tr>
<tr>
<td></td>
<td>x = 0.47</td>
<td>x = 0.13</td>
<td>x = 0.27</td>
<td>x = 0.27</td>
<td>1.13</td>
</tr>
<tr>
<td></td>
<td>s. d. = 0.64</td>
<td>s. d. = 0.35</td>
<td>s. d. = 0.46</td>
<td>s. d. = 0.59</td>
<td>s. d. = 1.19</td>
</tr>
<tr>
<td>LD</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>13</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>1.7%</td>
<td>0.6%</td>
<td>0.6%</td>
<td>7.2%</td>
<td>10%</td>
</tr>
<tr>
<td></td>
<td>x = 0.20</td>
<td>x = 0.07</td>
<td>x = 0.07</td>
<td>x = 0.87</td>
<td>1.20</td>
</tr>
<tr>
<td></td>
<td>s. d. = 0.41</td>
<td>s. d. = 0.26</td>
<td>s. d. = 0.26</td>
<td>s. d. = 0.74</td>
<td>s. d. = 1.01</td>
</tr>
<tr>
<td>SP</td>
<td>15</td>
<td>9</td>
<td>4</td>
<td>50</td>
<td>78</td>
</tr>
<tr>
<td></td>
<td>4.8%</td>
<td>2.9%</td>
<td>1.3%</td>
<td>16%</td>
<td>25%</td>
</tr>
<tr>
<td></td>
<td>x = 0.58</td>
<td>x = 0.35</td>
<td>x = 0.15</td>
<td>x = 1.92</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>s. d. = 0.76</td>
<td>s. d. = 0.56</td>
<td>s. d. = 0.37</td>
<td>s. d. = 1.62</td>
<td>s. d. = 2.02</td>
</tr>
<tr>
<td>TOTAL*</td>
<td>31</td>
<td>14</td>
<td>10</td>
<td>73</td>
<td>128</td>
</tr>
<tr>
<td></td>
<td>3.6%</td>
<td>1.6%</td>
<td>1.2%</td>
<td>8.6%</td>
<td>15%</td>
</tr>
<tr>
<td></td>
<td>x = 0.44</td>
<td>x = 0.20</td>
<td>x = 0.14</td>
<td>x = 1.03</td>
<td>1.80</td>
</tr>
<tr>
<td></td>
<td>s. d. = 0.63</td>
<td>s. d. = 0.44</td>
<td>s. d. = 0.35</td>
<td>s. d. = 1.32</td>
<td>s. d. = 1.73</td>
</tr>
</tbody>
</table>

Percentages reflect the proportion of the total number of items for the given group (i.e. including items coded "appropriate", "inappropriate", "ambiguous" and "other figurative sense"): 180 for MJ, MI and LD, respectively, and 312 for SP.

* The percentages in this row reflect the proportion of the total number of items for all the groups summed i.e. 852.
Between-groups differences on the "inappropriate" narrow codes were tested with a series of Mann-Whitney Tests. The results are presented in Table 9 below.

**TABLE 9: PLAY TASK - BETWEEN-GROUPS DIFFERENCES ON "INAPPROPRIATE" NARROW CODES (Mann-Whitney Test)**

<table>
<thead>
<tr>
<th>GROUPS COMPARED</th>
<th>LITERAL</th>
<th>INFERRED LITERAL</th>
<th>INAPPROPRIATE NO ACTION</th>
<th>FUZZY</th>
</tr>
</thead>
<tbody>
<tr>
<td>MJ - MI</td>
<td>$U = 109.5$ n.s.</td>
<td>$U = 112.5$ n.s.</td>
<td>$U = 90$ n.s.</td>
<td>$U = 104$ n.s.</td>
</tr>
<tr>
<td>MJ - LD</td>
<td>$U = 90$ n.s.</td>
<td>$U = 105$ n.s.</td>
<td>$U = 112.5$ n.s.</td>
<td>$U = 71.5$ n.s.</td>
</tr>
<tr>
<td>MJ - SP</td>
<td>$U = 178.5$ n.s.</td>
<td>$U = 160$ n.s.</td>
<td>$U = 178$ n.s.</td>
<td>$U = 62.5$ p = .0002</td>
</tr>
<tr>
<td>MI - LD</td>
<td>$U = 88.5$ n.s.</td>
<td>$U = 105$ n.s.</td>
<td>$U = 90$ n.s.</td>
<td>$U = 60.5$ p = .0152</td>
</tr>
<tr>
<td>MI - SP</td>
<td>$U = 184$ n.s.</td>
<td>$U = 160$ n.s.</td>
<td>$U = 173$ n.s.</td>
<td>$U = 48.5$ p = .0000</td>
</tr>
<tr>
<td>LD - SP</td>
<td>$U = 145.5$ n.s.</td>
<td>$U = 147.5$ n.s.</td>
<td>$U = 178$ n.s.</td>
<td>$U = 106.5$ p = .0124</td>
</tr>
</tbody>
</table>

These results indicate that the SP group did not differ significantly from the MJ, MI or LD groups in the number of items scored "literal", "inferred literal" or "inappropriate no action". The SP group did score significantly more "fuzzy" codes than did each of the other three groups.

The LD group did not differ significantly from either the MJ or MI group on the codes "literal", "inferred literal" or "inappropriate no action". The LD group scored more "fuzzy" items than both mainstream groups. This difference reached significance between the LD and MI groups but not between the LD and MJ groups.

136
The difference between the MJ and MI groups was not significant for "literal", "inferred literal", "inappropriate no action" or "fuzzy" scores.

The "ambiguous" items scored on the play task were all coded "ambiguous-idiomatic/literal" on the narrow codes. No items were coded "ambiguous-idiomatic/other figurative sense" or "ambiguous-literal/other figurative sense". The results pertaining to "ambiguous-idiomatic/literal" are thus the same as are reported for the broad code "ambiguous" in Tables 4 and 5 above.

Only one item was coded "other figurative sense" on the play task. This was scored by a subject in the MI group.
(c) DISCRIMINANT ANALYSIS

A Kth neighbour discriminant analysis of the play task data accurately predicted the group membership of 64 out of the 71 subjects (90.1%). The results of the discriminant analysis by group are presented in Table 10 below.

<table>
<thead>
<tr>
<th>GROUP</th>
<th>NO. CORRECTLY ASSIGNED</th>
<th>NO. INCORRECTLY ASSIGNED</th>
<th>ANALYSIS OF SUBJECTS INCORRECTLY ASSIGNED</th>
</tr>
</thead>
<tbody>
<tr>
<td>MJ</td>
<td>15 (100%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MI</td>
<td>13 (86.7%)</td>
<td>2 (13.3%)</td>
<td>MI5 GROUPED WITH MJ</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MI7 GROUPED WITH MJ</td>
</tr>
<tr>
<td>LD</td>
<td>14 (93.3%)</td>
<td>1 (6.7%)</td>
<td>LD2 GROUPED WITH MJ</td>
</tr>
<tr>
<td>SP</td>
<td>22 (84.6%)</td>
<td>4 (15.4%)</td>
<td>SP10 GROUPED WITH LD</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SP16 GROUPED WITH MJ</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SP17 GROUPED WITH MI</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SP25 GROUPED WITH MJ</td>
</tr>
<tr>
<td>TOTAL</td>
<td>64 (90.1%)</td>
<td>7 (9.9%)</td>
<td></td>
</tr>
</tbody>
</table>
2. DEFINITION TASK: BETWEEN-GROUPS ANALYSIS

(a) BROAD ANALYSIS

The broad level results for the four groups (MJ, MI, LD and SP) are summarised in Table 11 below and are represented graphically in Figure 7.

<table>
<thead>
<tr>
<th>GROUP</th>
<th>APPROPRIATE</th>
<th>INAPPROPRIATE</th>
<th>AMBIGUOUS</th>
<th>OTHER FIG'TIVE SENSE</th>
<th>TOTAL NO. ITEMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MJ</td>
<td>136</td>
<td>27</td>
<td>14</td>
<td>3</td>
<td>180</td>
</tr>
<tr>
<td></td>
<td>75.6%</td>
<td>15%</td>
<td>7.8%</td>
<td>1.7%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>x = 9.07</td>
<td>x = 1.80</td>
<td>x = 0.93</td>
<td>s. d. = 1.10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>s. d. = 1.75</td>
<td>s. d. = 1.57</td>
<td>s. d. = 0.41</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MI</td>
<td>113</td>
<td>56</td>
<td>9</td>
<td>2</td>
<td>180</td>
</tr>
<tr>
<td></td>
<td>62.8%</td>
<td>31.1%</td>
<td>5%</td>
<td>1.1%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>x = 7.53</td>
<td>x = 3.73</td>
<td>x = 0.60</td>
<td>s. d. = 0.91</td>
<td></td>
</tr>
<tr>
<td></td>
<td>s. d. = 2.07</td>
<td>s. d. = 2.05</td>
<td>s. d. = 0.35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LD</td>
<td>84</td>
<td>81</td>
<td>12</td>
<td>3</td>
<td>180</td>
</tr>
<tr>
<td></td>
<td>46.7%</td>
<td>45%</td>
<td>6.7%</td>
<td>1.7%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>x = 5.60</td>
<td>x = 5.40</td>
<td>x = 0.80</td>
<td>s. d. = 1.01</td>
<td></td>
</tr>
<tr>
<td></td>
<td>s. d. = 2.47</td>
<td>s. d. = 2.75</td>
<td>s. d. = 0.41</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SP</td>
<td>150</td>
<td>132</td>
<td>24</td>
<td>5</td>
<td>311*</td>
</tr>
<tr>
<td></td>
<td>48.2%</td>
<td>42.4%</td>
<td>7.7%</td>
<td>1.6%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>x = 5.77</td>
<td>x = 5.08</td>
<td>x = 0.92</td>
<td>s. d. = 0.49</td>
<td></td>
</tr>
<tr>
<td></td>
<td>s. d. = 2.67</td>
<td>s. d. = 2.83</td>
<td>s. d. = 0.19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL†</td>
<td>483</td>
<td>296</td>
<td>59</td>
<td>13</td>
<td>851†</td>
</tr>
<tr>
<td></td>
<td>56.8%</td>
<td>34.8%</td>
<td>6.9%</td>
<td>1.5%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>x = 6.80</td>
<td>x = 4.17</td>
<td>x = 0.83</td>
<td>s. d. = 0.43</td>
<td></td>
</tr>
<tr>
<td></td>
<td>s. d. = 2.68</td>
<td>s. d. = 2.73</td>
<td>s. d. = 0.99</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

APPROPRIATE includes "idiomatic" and "inferred idiomatic". INAPPROPRIATE includes "literal", "inferred literal", "fuzzy", "repeat", "narrative", "repeat-narrative", and "don't know". AMBIGUOUS includes "ambiguous-idiomatic/literal", "ambiguous-idiomatic/other figurative sense" and "ambiguous-literal/other figurative sense"

Percentages reflect the proportion of the total number of items for the given group (i.e. including items coded "appropriate", "inappropriate", "ambiguous" and "other figurative sense"): 180 for MJ, MI and LD, respectively, and 311 for SP.

* One item omitted for one child in the SP group

† The percentages in this row reflect the proportion of the total number of items for all the groups summed i.e. 851.
FIGURE 7

DEFINITION TASK - GROUP RESULTS  (BROAD CODES)

PERCENTAGE OF ITEMS

GROUP

- APPROPRIATE
- INAPPROPRIATE
- OTHER FIGURATIVE SENSE
- AMBIGUOUS

FIGURE 7
The differences between groups on the broad codes were tested with a series of Mann-Whitney Tests. The results are presented in Table 12 below. Between-groups differences for "other figurative sense" were not tested as it was clear from the raw scores that these would not reach significance.

**TABLE 12: DEFINITION TASK - BETWEEN-GROUPS DIFFERENCES ON BROAD CODES (Mann-Whitney Test)**

<table>
<thead>
<tr>
<th>GROUPS COMPARED</th>
<th>APPROPRIATE</th>
<th>INAPPROPRIATE</th>
<th>AMBIGUOUS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MJ - MI</strong></td>
<td>U = 66</td>
<td>U = 52</td>
<td>U = 93.5</td>
</tr>
<tr>
<td></td>
<td>p = .0253</td>
<td>p = .0055</td>
<td>n.s.</td>
</tr>
<tr>
<td><strong>MJ - LD</strong></td>
<td>U = 27</td>
<td>U = 26.5</td>
<td>U = 106.5</td>
</tr>
<tr>
<td></td>
<td>p = .0002</td>
<td>p = .0002</td>
<td>n.s.</td>
</tr>
<tr>
<td><strong>MJ - SP</strong></td>
<td>U = 62.5</td>
<td>U = 66.5</td>
<td>U = 190</td>
</tr>
<tr>
<td></td>
<td>p = .0002</td>
<td>p = .0003</td>
<td>n.s.</td>
</tr>
<tr>
<td><strong>MI - LD</strong></td>
<td>U = 66</td>
<td>U = 72.5</td>
<td>U = 98.5</td>
</tr>
<tr>
<td></td>
<td>p = .0253</td>
<td>p = .0468</td>
<td>n.s.</td>
</tr>
<tr>
<td><strong>MI - SP</strong></td>
<td>U = 120</td>
<td>U = 144.5</td>
<td>U = 152.5</td>
</tr>
<tr>
<td></td>
<td>p = .0200</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td><strong>LD - SP</strong></td>
<td>U = 192.5</td>
<td>U = 180</td>
<td>U = 175</td>
</tr>
<tr>
<td></td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

These results indicate that the SP group scored significantly fewer "appropriate" items and significantly more "inappropriate" items than did the MJ group. The difference between the SP and MI group was significant for "appropriate" (SP group scored fewer) but did not reach significance for the number of "inappropriate" items. The number of ambiguous items was not significantly different for the SP and MJ, and SP and MI groups, respectively.

The SP group did not differ significantly from the LD group in the number of "appropriate", "inappropriate" or "ambiguous" items scored.
The LD group scored significantly fewer "appropriate" and significantly more "inappropriate" items than either the MJ or MI groups. The LD group did not differ significantly from either the MJ or MI group on "ambiguous".

The MI group scored significantly fewer "appropriate" and significantly more "inappropriate" items than did the MJ group. They did not differ significantly in the number of "ambiguous" items scored.
(b) NARROW ANALYSIS

Group results on each of the "appropriate" narrow codes are summarised in Table 13 below.

**TABLE 13: DEFINITION TASK – BREAKDOWN OF "APPROPRIATE" ITEMS INTO "APPROPRIATE" NARROW CODES**

<table>
<thead>
<tr>
<th>GROUP</th>
<th>IDIOMATIC</th>
<th>INFERRED IDIOMATIC</th>
<th>TOTAL NO. APPROPRIATE ITEMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MJ</td>
<td>101</td>
<td>35</td>
<td>136</td>
</tr>
<tr>
<td></td>
<td>56.1%</td>
<td>19.4%</td>
<td>75.6%</td>
</tr>
<tr>
<td></td>
<td>x = 6.73</td>
<td>x = 2.33</td>
<td>x = 9.07</td>
</tr>
<tr>
<td></td>
<td>s. d. = 1.58</td>
<td>s. d. = 1.18</td>
<td>s. d. = 1.75</td>
</tr>
<tr>
<td>MI</td>
<td>82</td>
<td>31</td>
<td>113</td>
</tr>
<tr>
<td></td>
<td>45.6%</td>
<td>17.2%</td>
<td>62.8%</td>
</tr>
<tr>
<td></td>
<td>x = 5.47</td>
<td>x = 2.07</td>
<td>x = 7.53</td>
</tr>
<tr>
<td></td>
<td>s. d. = 2.07</td>
<td>s. d. = 1.39</td>
<td>s. d. = 2.07</td>
</tr>
<tr>
<td>LD</td>
<td>58</td>
<td>26</td>
<td>84</td>
</tr>
<tr>
<td></td>
<td>32.2%</td>
<td>14.4%</td>
<td>46.7%</td>
</tr>
<tr>
<td></td>
<td>x = 3.87</td>
<td>x = 1.73</td>
<td>x = 5.60</td>
</tr>
<tr>
<td></td>
<td>s. d. = 1.73</td>
<td>s. d. = 1.16</td>
<td>s. d. = 2.47</td>
</tr>
<tr>
<td>SP</td>
<td>94</td>
<td>56</td>
<td>150</td>
</tr>
<tr>
<td></td>
<td>30.2%</td>
<td>18%</td>
<td>48.2%</td>
</tr>
<tr>
<td></td>
<td>x = 3.62</td>
<td>x = 2.15</td>
<td>x = 5.77</td>
</tr>
<tr>
<td></td>
<td>s. d. = 2.73</td>
<td>s. d. = 1.38</td>
<td>s. d. = 2.67</td>
</tr>
<tr>
<td>TOTAL*</td>
<td>335</td>
<td>148</td>
<td>483</td>
</tr>
<tr>
<td></td>
<td>39.4%</td>
<td>17.4%</td>
<td>56.8%</td>
</tr>
<tr>
<td></td>
<td>x = 4.72</td>
<td>x = 2.09</td>
<td>x = 6.80</td>
</tr>
<tr>
<td></td>
<td>s. d. = 2.49</td>
<td>s. d. = 1.29</td>
<td>s. d. = 2.68</td>
</tr>
</tbody>
</table>

Percentages reflect the proportion of the total number of items for the given group (i.e. including items coded "appropriate", "inappropriate", "ambiguous" and "other figurative sense"): 180 for MJ, MI and LD, respectively, and 311 for SP.

* The percentages in this row reflect the proportion of the total number of items for all the groups summed i.e. 851.
Between-groups differences on the "appropriate" narrow codes were tested with a series of Mann-Whitney Tests. The results are presented in Table 14 below.

TABLE 14: DEFINITION TASK – BETWEEN-GROUPS DIFFERENCES ON "APPROPRIATE" NARROW CODES (Mann-Whitney Test)

<table>
<thead>
<tr>
<th>GROUPS COMPARED</th>
<th>IDIOMATIC</th>
<th>INFERRED IDIOMATIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>MJ - MI</td>
<td>U = 70</td>
<td>U = 91.5</td>
</tr>
<tr>
<td></td>
<td>p = .0360</td>
<td>n.s.</td>
</tr>
<tr>
<td>MJ - LD</td>
<td>U = 19.5</td>
<td>U = 80</td>
</tr>
<tr>
<td></td>
<td>p = .0001</td>
<td>n.s.</td>
</tr>
<tr>
<td>MJ - SP</td>
<td>U = 62</td>
<td>U = 173</td>
</tr>
<tr>
<td></td>
<td>p = .0002</td>
<td>n.s.</td>
</tr>
<tr>
<td>MI - LD</td>
<td>U = 65.5</td>
<td>U = 104</td>
</tr>
<tr>
<td></td>
<td>p = .0483</td>
<td>n.s.</td>
</tr>
<tr>
<td>MI - SP</td>
<td>U = 108.5</td>
<td>U = 187.5</td>
</tr>
<tr>
<td></td>
<td>p = .0092</td>
<td>n.s.</td>
</tr>
<tr>
<td>LD - SP</td>
<td>U = 168.5</td>
<td>U = 169</td>
</tr>
<tr>
<td></td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

The results indicate that the SP group scored significantly fewer "idiomatic" items than did the MJ and MI groups, respectively. The number of "inferred idiomatic" items was not significantly different for the SP and MJ groups or for the SP and MI groups. The SP group did not differ significantly from the LD group on either "idiomatic" or "inferred idiomatic" codes.

The LD group scored significantly fewer "idiomatic" items than did the MJ and MI groups, respectively. The LD group did not differ significantly from the MJ or MI groups in the number of "inferred idiomatic" items scored.

The MJ group scored significantly more "idiomatic" items than did the MI group. The number of "inferred idiomatic" items was not significantly different.
Group results on each of the "inappropriate" narrow codes are summarised in Table 15 below.

**TABLE 15: DEFINITION TASK - BREAKDOWN OF "INAPPROPRIATE" ITEMS INTO "INAPPROPRIATE" NARROW CODES**

<table>
<thead>
<tr>
<th>GROUP</th>
<th>LITERAL</th>
<th>INFERENCE LITERAL</th>
<th>FUZZY</th>
<th>REPEAT</th>
<th>NARRATIVE</th>
<th>REPEAT-NARRATIVE</th>
<th>DON'T KNOW</th>
<th>TOTAL NO. IN-APPROPR. ITEMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MJ</td>
<td>3</td>
<td>1.7%</td>
<td>2</td>
<td>1.1%</td>
<td>4.2%</td>
<td>2.2%</td>
<td>1</td>
<td>0.6%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>x = 0.20</td>
<td></td>
<td>x = 0.13</td>
<td>x = 0.27</td>
<td>x = 0.07</td>
<td>x = 0.07</td>
<td>x = 0.47</td>
</tr>
<tr>
<td></td>
<td></td>
<td>sd = 0.41</td>
<td></td>
<td>sd = 0.35</td>
<td>sd = 0.46</td>
<td>sd = 0.26</td>
<td>sd = 0.26</td>
<td>sd = 0.74</td>
</tr>
<tr>
<td>MI</td>
<td>7</td>
<td>3.9%</td>
<td>9</td>
<td>5%</td>
<td>4.2%</td>
<td>2.2%</td>
<td>1</td>
<td>1.1%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>x = 0.47</td>
<td></td>
<td>x = 0.60</td>
<td>x = 0.27</td>
<td>x = 0.20</td>
<td>x = 0.20</td>
<td>x = 0.13</td>
</tr>
<tr>
<td></td>
<td></td>
<td>sd = 0.64</td>
<td></td>
<td>sd = 0.91</td>
<td>sd = 0.46</td>
<td>sd = 0.56</td>
<td>sd = 0.35</td>
<td>sd = 1.22</td>
</tr>
<tr>
<td>LD</td>
<td>8</td>
<td>4.4%</td>
<td>7</td>
<td>3.9%</td>
<td>12.2%</td>
<td>1.7%</td>
<td>3</td>
<td>1.7%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>x = 0.53</td>
<td></td>
<td>x = 0.47</td>
<td>x = 1.47</td>
<td>x = 0.20</td>
<td>x = 0.20</td>
<td>x = 1.13</td>
</tr>
<tr>
<td></td>
<td></td>
<td>sd = 0.83</td>
<td></td>
<td>sd = 0.64</td>
<td>sd = 1.25</td>
<td>sd = 0.41</td>
<td>sd = 0.41</td>
<td>sd = 1.51</td>
</tr>
<tr>
<td>SP</td>
<td>12</td>
<td>3.9%</td>
<td>29</td>
<td>9.3%</td>
<td>10.9%</td>
<td>1%</td>
<td>3</td>
<td>1%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>x = 0.46</td>
<td></td>
<td>x = 1.12</td>
<td>x = 1.31</td>
<td>x = 0.12</td>
<td>x = 0.12</td>
<td>x = 1.08</td>
</tr>
<tr>
<td></td>
<td></td>
<td>sd = 0.71</td>
<td></td>
<td>sd = 1.40</td>
<td>sd = 1.44</td>
<td>sd = 0.43</td>
<td>sd = 0.43</td>
<td>sd = 1.16</td>
</tr>
<tr>
<td>TOTAL</td>
<td>30</td>
<td>3.5%</td>
<td>47</td>
<td>5.5%</td>
<td>7.5%</td>
<td>1.2%</td>
<td>10</td>
<td>1.2%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>x = 0.42</td>
<td></td>
<td>x = 0.66</td>
<td>x = 0.90</td>
<td>x = 0.14</td>
<td>x = 0.14</td>
<td>x = 0.76</td>
</tr>
<tr>
<td></td>
<td></td>
<td>sd = 0.67</td>
<td></td>
<td>sd = 1.06</td>
<td>sd = 1.20</td>
<td>sd = 0.42</td>
<td>sd = 1.11</td>
<td>sd = 1.13</td>
</tr>
</tbody>
</table>

Percentages reflect the proportion of the total number of items for the given group (i.e. including items coded "appropriate", "inappropriate", "ambiguous" and "other figurative sense"): 180 for MJ, MI and LD, respectively, and 311 for SP.

* The percentages in this row reflect the proportion of the total number of items for all the groups summed i.e. 851.
Between-groups differences on the "inappropriate" narrow codes were tested with a series of Mann-Whitney Tests. The results are presented in Table 16 below.

### TABLE 16: DEFINITION TASK - BETWEEN-GROUPS DIFFERENCES ON "INAPPROPRIATE" NARROW CODES (Mann-Whitney Test)

<table>
<thead>
<tr>
<th>GROUPS COMPARED</th>
<th>LITERAL</th>
<th>INFERRED LITERAL</th>
<th>FUZZY</th>
<th>REPEAT</th>
<th>NARRATIVE</th>
<th>REPEAT-NARRATIVE</th>
<th>DON'T KNOW</th>
</tr>
</thead>
<tbody>
<tr>
<td>MJ - MI</td>
<td>U = 88.5 n.s.</td>
<td>U = 90.5 n.s.</td>
<td>U = 112.5 n.s.</td>
<td>U = 104.5 n.s.</td>
<td>U = 88 n.s.</td>
<td>U = 54.5 p = .0120</td>
<td>U = 82.5 p = .0350</td>
</tr>
<tr>
<td>MJ - LD</td>
<td>U = 93 n.s.</td>
<td>U = 81.5 n.s.</td>
<td>U = 46 p = .0028</td>
<td>U = 97.5 n.s.</td>
<td>U = 85 n.s.</td>
<td>U = 85.5 p = .0351</td>
<td></td>
</tr>
<tr>
<td>MJ - SP</td>
<td>U = 157.5 n.s.</td>
<td>U = 101.5 p = .0043</td>
<td>U = 103.5 p = .0073</td>
<td>U = 192.5 n.s.</td>
<td>U = 134.5 n.s.</td>
<td>U = 181.5 n.s.</td>
<td>U = 172.5 n.s.</td>
</tr>
<tr>
<td>MI - LD</td>
<td>U = 111.5 n.s.</td>
<td>U = 109 n.s.</td>
<td>U = 46 p = .0028</td>
<td>U = 106.5 n.s.</td>
<td>U = 63.5 p = .0149</td>
<td>U = 85.5 n.s.</td>
<td>U = 110.5 n.s.</td>
</tr>
<tr>
<td>MI - SP</td>
<td>U = 190.5 n.s.</td>
<td>U = 153.5 n.s.</td>
<td>U = 103.5 p = .0073</td>
<td>U = 184 n.s.</td>
<td>U = 94 p = .0023</td>
<td>U = 107.5 p = .0132</td>
<td>U = 166 n.s.</td>
</tr>
<tr>
<td>LD - SP</td>
<td>U = 194 n.s.</td>
<td>U = 145.5 n.s.</td>
<td>U = 172 n.s.</td>
<td>U = 172.5 n.s.</td>
<td>U = 186.5 n.s.</td>
<td>U = 163.5 n.s.</td>
<td>U = 164.5 n.s.</td>
</tr>
</tbody>
</table>

In comparison with the MJ group, the SP group scored significantly more "inferred literal" and "fuzzy" items. The SP and MJ group differences were not significant for the codes "literal", "repeat", "narrative", "repeat-narrative" or "don't know".

Compared with the MI group, the SP group scored significantly more "fuzzy" and "narrative" codes, and significantly fewer "repeat-narrative" codes. The number of "literal", "inferred literal", "repeat" and "don't know" codes scored by the SP and MI groups did not differ significantly.

The SP group did not differ significantly from the LD group on
any of the "inappropriate" narrow codes.

The LD group scored significantly more "fuzzy" and "don't know" codes than did the MJ group but did not otherwise differ significantly from the MJ group on the "inappropriate" narrow codes.

The LD group also scored significantly more "fuzzy" and "narrative" codes than did the MI group. The LD and MI groups did not differ significantly on "literal", "inferred literal", "repeat", "repeat-narrative" or "don't know" codes.

The MJ and MI groups did not differ significantly in the number of "literal", "inferred literal", "fuzzy", "repeat" or "narrative" codes they scored. The MI group did, however, score significantly more "repeat-narrative" and "don't know" codes than did the MJ group.
Group results on each of the narrow codes making up the broad code "ambiguous" are summarised in Table 17 below. (These codes are henceforth referred to collectively as the "'ambiguous' narrow codes".)

**TABLE 17 : DEFINITION TASK - BREAKDOWN OF "AMBIGUOUS" ITEMS INTO "AMBIGUOUS" NARROW CODES**

<table>
<thead>
<tr>
<th>GROUP</th>
<th>AMBIGUOUS-ID/LIT</th>
<th>AMBIGUOUS-ID/OFS</th>
<th>AMBIGUOUS-LIT/OFS</th>
<th>TOTAL NO. AMBIGUOUS ITEMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MJ</td>
<td>14</td>
<td>14</td>
<td></td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>7.8%</td>
<td></td>
<td></td>
<td>7.8%</td>
</tr>
<tr>
<td></td>
<td>(x = 0.93)</td>
<td></td>
<td></td>
<td>(x = 0.93)</td>
</tr>
<tr>
<td></td>
<td>s.d. = 1.10</td>
<td></td>
<td></td>
<td>s.d. = 1.10</td>
</tr>
<tr>
<td>MI</td>
<td>8</td>
<td>1</td>
<td></td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>4.4%</td>
<td>0.6%</td>
<td></td>
<td>5%</td>
</tr>
<tr>
<td></td>
<td>(x = 0.53)</td>
<td>(x = 0.07)</td>
<td></td>
<td>(x = 0.60)</td>
</tr>
<tr>
<td></td>
<td>s.d. = 0.92</td>
<td>s.d. = 0.26</td>
<td></td>
<td>s.d. = 0.91</td>
</tr>
<tr>
<td>LD</td>
<td>10</td>
<td>2</td>
<td></td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>5.6%</td>
<td>1.1%</td>
<td></td>
<td>6.7%</td>
</tr>
<tr>
<td></td>
<td>(x = 0.67)</td>
<td>(x = 0.13)</td>
<td></td>
<td>(x = 0.80)</td>
</tr>
<tr>
<td></td>
<td>s.d. = 0.90</td>
<td>s.d. = 0.35</td>
<td></td>
<td>s.d. = 1.01</td>
</tr>
<tr>
<td>SP</td>
<td>22</td>
<td>1</td>
<td></td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>7.1%</td>
<td>0.3%</td>
<td></td>
<td>7.7%</td>
</tr>
<tr>
<td></td>
<td>(x = 0.85)</td>
<td>(x = 0.04)</td>
<td></td>
<td>(x = 0.92)</td>
</tr>
<tr>
<td></td>
<td>s.d. = 0.93</td>
<td>s.d. = 0.20</td>
<td></td>
<td>s.d. = 0.98</td>
</tr>
<tr>
<td>TOTAL*</td>
<td>54</td>
<td>4</td>
<td></td>
<td>59</td>
</tr>
<tr>
<td></td>
<td>6.4%</td>
<td>0.5%</td>
<td></td>
<td>6.9%</td>
</tr>
<tr>
<td></td>
<td>(x = 0.76)</td>
<td>(x = 0.06)</td>
<td></td>
<td>(x = 0.83)</td>
</tr>
<tr>
<td></td>
<td>s.d. = 0.95</td>
<td>s.d. = 0.23</td>
<td></td>
<td>s.d. = 0.99</td>
</tr>
</tbody>
</table>

Percentages reflect the proportion of the total number of items for the given group (i.e. including items coded "appropriate", "inappropriate", "ambiguous" and "other figurative sense"): 180 for MJ, MI and LD, respectively, and 311 for SP.

* The percentages in this row reflect the proportion of the total number of items for all the groups summed i.e. 851.
The between-groups differences on the "ambiguous" narrow codes were tested with a series of Mann-Whitney tests. The results are presented in Table 18 below. The differences on "ambiguous-idiomatic/OFS" and "ambiguous-lit/OFS" were not tested since it was clear from the raw scores that these would not reach significance.

TABLE 18: DEFINITION TASK - BETWEEN-GROUPS DIFFERENCES ON "AMBIGUOUS" NARROW CODES (Mann-Whitney Test)

<table>
<thead>
<tr>
<th>GROUPS COMPARED</th>
<th>AMBIGUOUS - ID/LIT</th>
</tr>
</thead>
<tbody>
<tr>
<td>MJ - MI</td>
<td>U = 88 n.s.</td>
</tr>
<tr>
<td>MJ - LD</td>
<td>U = 99 n.s.</td>
</tr>
<tr>
<td>MJ - SP</td>
<td>U = 194 n.s.</td>
</tr>
<tr>
<td>MI - LD</td>
<td>U = 99.5 n.s.</td>
</tr>
<tr>
<td>MI - SP</td>
<td>U = 146.5 n.s.</td>
</tr>
<tr>
<td>LD - SP</td>
<td>U = 168.5 n.s.</td>
</tr>
</tbody>
</table>

The results indicate that the number of items scored "ambiguous-id/lit" by the SP group did not differ significantly from that scored by any of the other groups. The difference in the number of items scored "ambiguous-id/lit" did not reach significance for any other pair of groups, either.

As noted above, the broad code "other figurative sense" corresponds to the narrow code of the same name. The number of items coded "other figurative sense" on the definition task for each group are presented in Table 11 above. No between-groups statistics were calculated for this code since it was clear from the raw scores that these differences would not reach significance.
A $K$th neighbour discriminant analysis of the definition task data accurately predicted the group membership of 68 out of the 71 subjects (95.8%). The results of the discriminant analysis by group are presented in Table 19 below.

### Table 19: Definition Task - Discriminant Analysis by Group ($K$th neighbour discriminant analysis)

<table>
<thead>
<tr>
<th>GROUP</th>
<th>NO. CORRECTLY ASSIGNED</th>
<th>NO. INCORRECTLY ASSIGNED</th>
<th>ANALYSIS OF SUBJECTS INCORRECTLY ASSIGNED</th>
</tr>
</thead>
<tbody>
<tr>
<td>MJ</td>
<td>15 (100%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MI</td>
<td>15 (100%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LD</td>
<td>15 (100%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SP</td>
<td>23 (88.5%)</td>
<td>3 (11.5%)</td>
<td>SP4 grouped with MJ</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SP14 grouped with MI</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SP25 grouped with MI</td>
</tr>
<tr>
<td>TOTAL</td>
<td>68 (95.8%)</td>
<td>3 (4.2%)</td>
<td></td>
</tr>
</tbody>
</table>
3. BETWEEN-TASKS ANALYSIS

In order to compare the results across the two tasks, a series of Wilcoxon Signed Rank Tests was performed. Table 20 below summarises the broad code results for each group and the differences between the play task and definition task scores for each of the broad codes.

### TABLE 20: BETWEEN-TASKS DIFFERENCES - BROAD CODES (Wilcoxon Signed Rank Test)

<table>
<thead>
<tr>
<th>BROAD CODE</th>
<th>FEATURE</th>
<th>MJ</th>
<th>MI</th>
<th>LD</th>
<th>SP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>81.7%</td>
<td>76.1%</td>
<td>74.4%</td>
<td>60.9%</td>
</tr>
<tr>
<td>APPROPRIATE</td>
<td>PLAY TASK</td>
<td>75.6%</td>
<td>62.8%</td>
<td>46.7%</td>
<td>48.2%</td>
</tr>
<tr>
<td></td>
<td>DEFINITION TASK</td>
<td>W = 18</td>
<td>W = 22.5</td>
<td>W = 0</td>
<td>W = 39</td>
</tr>
<tr>
<td></td>
<td></td>
<td>n.s.</td>
<td>p = .0299</td>
<td>p = .0005</td>
<td>p = .0023</td>
</tr>
<tr>
<td></td>
<td>DIFFERENCE</td>
<td>U = 18</td>
<td>U = 22.5</td>
<td>U = 0</td>
<td>U = 39</td>
</tr>
<tr>
<td></td>
<td></td>
<td>n.s.</td>
<td>p = .0299</td>
<td>p = .0005</td>
<td>p = .0023</td>
</tr>
<tr>
<td>INAPPROPRIATE</td>
<td>PLAY TASK</td>
<td>8.3%</td>
<td>9.4%</td>
<td>10%</td>
<td>25%</td>
</tr>
<tr>
<td></td>
<td>DEFINITION TASK</td>
<td>15%</td>
<td>31.1%</td>
<td>45%</td>
<td>42.4%</td>
</tr>
<tr>
<td></td>
<td>DIFFERENCE</td>
<td>W = 8.5</td>
<td>W = 2.5</td>
<td>W = 0</td>
<td>W = 14.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>p = .0486</td>
<td>p = .0009</td>
<td>p = .0004</td>
<td>p = .0001</td>
</tr>
<tr>
<td>AMBIGUOUS</td>
<td>PLAY TASK</td>
<td>10%</td>
<td>13.9%</td>
<td>15.6%</td>
<td>14.1%</td>
</tr>
<tr>
<td></td>
<td>DEFINITION TASK</td>
<td>7.8%</td>
<td>5%</td>
<td>6.7%</td>
<td>7.7%</td>
</tr>
<tr>
<td></td>
<td>DIFFERENCE</td>
<td>W = 23.5</td>
<td>W = 5</td>
<td>W = 6</td>
<td>W = 32</td>
</tr>
<tr>
<td></td>
<td></td>
<td>n.s.</td>
<td>p = .0128</td>
<td>p = .0164</td>
<td>p = .0352</td>
</tr>
<tr>
<td>OTHER FIGURATIVE SENSE</td>
<td>PLAY TASK</td>
<td>0.6%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DEFINITION TASK</td>
<td>1.7%</td>
<td>1.1%</td>
<td>1.7%</td>
<td>1.6%</td>
</tr>
<tr>
<td></td>
<td>DIFFERENCE</td>
<td>W = 0</td>
<td>W = 0</td>
<td>W = 0</td>
<td>W = 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

151
As is evident in Table 20, significantly more items were scored "appropriate" on the play task than on the definition task by the MI, LD and SP groups, respectively. The MJ group scored more "appropriate" items on the play task than on the definition task but this difference failed to reach significance.

Significantly fewer items were scored "inappropriate" on the play task than on the definition task by each of the four groups.

Significantly more items were scored "ambiguous" on the play task than on the definition task by the MI, LD and SP groups, respectively. The MJ group scored more "ambiguous" items on the play task than on the definition task but this difference was not significant.

The number of "other figurative sense" items did not differ significantly between tasks for any of the four groups.
Between-tasks differences were also calculated for the narrow codes "idiomatic", "inferred idiomatic", "literal", "inferred literal" and "fuzzy". The results of this analysis are presented in Table 21 below.

**TABLE 21: BETWEEN-TASKS DIFFERENCES - NARROW CODES (Wilcoxon Signed Rank Test)**

<table>
<thead>
<tr>
<th>NARROW CODE</th>
<th>FEATURE</th>
<th>Mj</th>
<th>Mi</th>
<th>LD</th>
<th>SP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>IDIOMATIC</strong></td>
<td>PLAY TASK</td>
<td>72.8%</td>
<td>63.9%</td>
<td>61.1%</td>
<td>50.3%</td>
</tr>
<tr>
<td></td>
<td>DEFINITION TASK</td>
<td>56.1%</td>
<td>45.6%</td>
<td>32.2%</td>
<td>30.2%</td>
</tr>
<tr>
<td></td>
<td>DIFFERENCE</td>
<td>W = 9, p = .0063</td>
<td>W = 7.5, p = .0135</td>
<td>W = 0, p = .0007</td>
<td>W = 6, p = .0001</td>
</tr>
<tr>
<td><strong>INFERRED IDIOMATIC</strong></td>
<td>PLAY TASK</td>
<td>7.8%</td>
<td>11.1%</td>
<td>11.7%</td>
<td>10.3%</td>
</tr>
<tr>
<td></td>
<td>DEFINITION TASK</td>
<td>19.4%</td>
<td>17.2%</td>
<td>14.4%</td>
<td>18%</td>
</tr>
<tr>
<td></td>
<td>DIFFERENCE</td>
<td>W = 12, p = .0110</td>
<td>W = 25.5, n.s.</td>
<td>W = 41.5, n.s.</td>
<td>W = 61.5, p = .0348</td>
</tr>
<tr>
<td><strong>LITERAL</strong></td>
<td>PLAY TASK</td>
<td>3.3%</td>
<td>3.9%</td>
<td>1.7%</td>
<td>4.8%</td>
</tr>
<tr>
<td></td>
<td>DEFINITION TASK</td>
<td>1.7%</td>
<td>3.9%</td>
<td>4.4%</td>
<td>3.9%</td>
</tr>
<tr>
<td><strong>INFERRED LITERAL</strong></td>
<td>PLAY TASK</td>
<td>1.1%</td>
<td>1.1%</td>
<td>0.6%</td>
<td>2.9%</td>
</tr>
<tr>
<td></td>
<td>DEFINITION TASK</td>
<td>1.1%</td>
<td>5%</td>
<td>3.9%</td>
<td>9.3%</td>
</tr>
<tr>
<td></td>
<td>DIFFERENCE</td>
<td>W = 5, n.s.</td>
<td>W = 3, n.s.</td>
<td>W = 3.5, n.s.</td>
<td>W = 12, p = .0192</td>
</tr>
<tr>
<td><strong>FUZZY</strong></td>
<td>PLAY TASK</td>
<td>3.3%</td>
<td>2.2%</td>
<td>7.2%</td>
<td>16%</td>
</tr>
<tr>
<td></td>
<td>DEFINITION TASK</td>
<td>2.2%</td>
<td>2.2%</td>
<td>12.2%</td>
<td>10.9%</td>
</tr>
<tr>
<td></td>
<td>DIFFERENCE</td>
<td>W = 14, n.s.</td>
<td>W = 10.5, n.s.</td>
<td>W = 4, p = .0499</td>
<td>W = 32.5, n.s.</td>
</tr>
</tbody>
</table>
The results of this analysis indicate that each of the four groups scored significantly more "idiomatic" items on the play task than they did on the definition task.

The number of items scored "inferred idiomatic" by the MJ and SP groups was significantly greater on the definition task than on the play task. This difference did not reach significance for either the MI or LD groups.

No significant differences emerged in the number of items scored "literal" on the two tasks.

Significantly more "inferred literal" items were scored by the SP group on the definition task than on the play task. This difference was not significant for the MJ, MI or LD groups.

The number of "fuzzy" items scored by the LD group on the definition task was significantly greater than that scored on the play task. This difference did not reach significance for any of the other three groups.
The difference between the number of items scored "appropriate" and "inappropriate" on each task was calculated with a series of Wilcoxon Signed Rank Tests. The results are presented in Table 22 below.

<table>
<thead>
<tr>
<th>GROUP</th>
<th>PLAY TASK</th>
<th>DEFINITION TASK</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>APPROPRIATE</td>
<td>INAPPROPRIATE</td>
</tr>
<tr>
<td>MJ</td>
<td>81.7%</td>
<td>8.3%</td>
</tr>
<tr>
<td>MI</td>
<td>76.1%</td>
<td>9.4%</td>
</tr>
<tr>
<td>LD</td>
<td>74.4%</td>
<td>10%</td>
</tr>
<tr>
<td>SP</td>
<td>60.9%</td>
<td>25%</td>
</tr>
</tbody>
</table>

The results indicate that the MJ group and the MI group scored significantly more "appropriate" than "inappropriate" items on both tasks. The LD group and the SP group both scored significantly more "appropriate" than "inappropriate" items on the play task. On the definition task, the difference between "appropriate" and "inappropriate" items was not significant for either the LD or SP group.
4. WITHIN-GROUP RESULTS

The narrow code results for each individual subject are summarised in this section.

(a) PLAY TASK

Within-group play task results are summarised in Tables 23, 24, 25 and 26, below. These reflect the results obtained by MJ, MI, LD and SP subjects, respectively.
TABLE 23: PLAY TASK - MJ GROUP RESULTS

<table>
<thead>
<tr>
<th>SUBJECT</th>
<th>APPROPRIATE</th>
<th>INAPPROPRIATE</th>
<th>AMBIGUOUS</th>
<th>OFS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ID</td>
<td>INF</td>
<td>ANA</td>
<td>LIT</td>
</tr>
<tr>
<td>MJ1</td>
<td>11</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>MJ2</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>MJ3</td>
<td>8</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>MJ4</td>
<td>8</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>MJ5</td>
<td>9</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MJ6</td>
<td>10</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MJ7</td>
<td>8</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MJ8</td>
<td>9</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>MJ9</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MJ10</td>
<td>10</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MJ11</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MJ12</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MJ13</td>
<td>9</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MJ14</td>
<td>8</td>
<td></td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>MJ15</td>
<td>8</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MEAN</td>
<td>8.73</td>
<td>0.93</td>
<td>0.13</td>
<td>0.40</td>
</tr>
<tr>
<td>S.D.</td>
<td>1.22</td>
<td>1.16</td>
<td>0.52</td>
<td>0.51</td>
</tr>
</tbody>
</table>

ID = "idiomatic"; INF ID = "inferred idiomatic"; ANA = "appropriate no action"; LIT = "literal"; INF LIT = "inferred literal"; FUZ = "fuzzy"; INA = "inappropriate no action"; AMB-ID/LIT = "ambiguous-idiomatic/literal"; OFS = "other figurative sense".

* No subject scored either of the codes "ambiguous-idiomatic/OFS" or "ambiguous-literal/OFS" on this task.
### TABLE 24: PLAY TASK - MI GROUP RESULTS

<table>
<thead>
<tr>
<th>SUBJ</th>
<th>ID</th>
<th>INF</th>
<th>ANA</th>
<th>LIT</th>
<th>INF</th>
<th>FUZ</th>
<th>INA</th>
<th>AMB-ID/LIT</th>
<th>OFS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MI1</td>
<td>9</td>
<td>1</td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MI2</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MI3</td>
<td>10</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MI4</td>
<td>8</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MI5</td>
<td>7</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MI6</td>
<td>8</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MI7</td>
<td>9</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MI8</td>
<td>5</td>
<td></td>
<td></td>
<td>2</td>
<td>1</td>
<td></td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MI9</td>
<td>9</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MI10</td>
<td>7</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>MI11</td>
<td>6</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>MI12</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MI13</td>
<td>4</td>
<td>2</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>MI14</td>
<td>9</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>MI15</td>
<td>5</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>MEAN</td>
<td>7.67</td>
<td>1.33</td>
<td>0.13</td>
<td>0.47</td>
<td>0.13</td>
<td>0.27</td>
<td>0.27</td>
<td>1.67</td>
<td>0.07</td>
</tr>
<tr>
<td>S.D.</td>
<td>1.92</td>
<td>1.05</td>
<td>0.35</td>
<td>0.64</td>
<td>0.35</td>
<td>0.59</td>
<td>0.46</td>
<td>1.11</td>
<td>0.26</td>
</tr>
</tbody>
</table>

ID = "idiomatic"; INF ID = "inferred idiomatic"; ANA = "appropriate no action"; LIT = "literal"; INF LIT = "inferred literal"; FUZ = "fuzzy"; INA = "inappropriate no action"; AMB-ID/LIT = "ambiguous-idiomatic/literal"; OFS = "other figurative sense".

* No subject scored either of the codes "ambiguous-idiomatic/OFS" or "ambiguous-literal/OFS" on this task.
| SUBJ | APPROPRIATE | | INAPPROPRIATE | | AMBIGUOUS* | | OFS |
|------|-------------|-------------|-------------|-------------|-------------|-------------|
|      | ID | INF ID | ANA | LIT | INF LIT | FUZ | INA | AMB-ID/LIT | OFS |
| LD1  | 9  | 1     |     | 1   |         |     |     | 1           | 1   |
| LD2  | 8  | 2     |     |     |         |     |     | 2           |     |
| LD3  | 8  | 1     |     | 1   |         |     |     |             | 1   |
| LD4  | 7  | 2     |     | 1   |         |     |     | 2           |     |
| LD5  | 5  | 3     |     | 2   | 1       |     |     | 1           |     |
| LD6  | 4  | 3     |     | 2   |         |     |     | 3           |     |
| LD7  | 5  | 1     | 1   |     |         |     |     |             | 4   |
| LD8  | 8  |       |     | 1   |         | 1   |     | 2           |     |
| LD9  | 11 |       |     |     |         |     |     |             | 1   |
| LD10 | 8  | 1     | 1   |     |         |     |     |             | 1   |
| LD11 | 10 |       |     |     |         |     |     |             | 1   |
| LD12 | 9  | 2     |     |     |         |     |     |             | 1   |
| LD13 | 6  | 1     | 1   |     |         |     |     |             | 3   |
| LD14 | 6  | 3     |     |     |         |     |     |             | 3   |
| LD15 | 6  | 1     |     | 1   |         |     | 2   |             | 2   |
| MEAN | 7.33 | 1.40 | 0.20 | 0.20 | 0.07 | 0.87 | 0.07 | 1.87 |
| S.D. | 1.99 | 1.06 | 0.41 | 0.41 | 0.26 | 0.74 | 0.26 | 0.99 |

ID = "idiomatic"; INF ID = "inferred idiomatic"; ANA = "appropriate no action"; LIT = "literal"; INF LIT = "inferred literal"; FUZ = "fuzzy"; INA = "inappropriate no action"; AMB-ID/LIT = "ambiguous-idiomatic/literal"; OFS = "other figurative sense".

* No subject scored either of the codes "ambiguous-idiomatic/OFS" or "ambiguous-literal/OFS" on this task.
<table>
<thead>
<tr>
<th>SUBJ#</th>
<th>APPROPRIATE</th>
<th>INAPPROPRIATE</th>
<th>AMBIGUOUS†</th>
<th>OFS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ID</td>
<td>INF ID</td>
<td>ANA</td>
<td>LIT</td>
</tr>
<tr>
<td>SP1</td>
<td>8</td>
<td>1</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>SP2</td>
<td>5</td>
<td>1</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>SP4</td>
<td>5</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>SP5</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SP6</td>
<td>10</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SP7</td>
<td>4</td>
<td>1</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>SP8</td>
<td>5</td>
<td>2</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>SP9</td>
<td>3</td>
<td>2</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>SP10</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SP11</td>
<td>6</td>
<td>2</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>SP12</td>
<td>6</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>SP13</td>
<td>7</td>
<td>1</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>SP14</td>
<td>8</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>SP15</td>
<td>7</td>
<td>3</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>SP16</td>
<td>10</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SP17</td>
<td>10</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SP18</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>SP3</td>
<td>5</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>SP19</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SP20</td>
<td>5</td>
<td>1</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>SP21</td>
<td>3</td>
<td>3</td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

continued overleaf
TABLE 26 continued

<table>
<thead>
<tr>
<th>SUBJ*</th>
<th>APPROPRIATE</th>
<th>INAPPROPRIATE</th>
<th>AMBIGUOUS†</th>
<th>OFS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ID</td>
<td>INF</td>
<td>ANA</td>
<td>LIT</td>
</tr>
<tr>
<td>SP22</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SP23</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>SP24</td>
<td>3</td>
<td></td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>SP25</td>
<td>9</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SP26</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>MEAN</td>
<td>6.04</td>
<td>1.23</td>
<td>0.04</td>
<td>0.58</td>
</tr>
<tr>
<td>S.D.</td>
<td>2.27</td>
<td>1.21</td>
<td>0.20</td>
<td>0.76</td>
</tr>
</tbody>
</table>

ID = "idiomatic"; INF ID = "inferred idiomatic"; ANA = "appropriate no action"; LIT = "literal"; INF LIT = "inferred literal"; FUZ = "fuzzy"; INA = "inappropriate no action"; AMB-ID/LIT = "ambiguous-idiomatic/literal"; OFS = "other figurative sense".

* The subgroup "SPD" comprises SP1-SP2 and SP4-SP18. The subgroup "ASP" comprises SP3 and SP19-SP26.
† No subject scored either of the codes "ambiguous-idiomatic/OFS" or "ambiguous-literal/OFS" on this task.

In order to examine the relationship between age and performance, Spearman's Rank Correlation was calculated for each group. The results are presented in Table 27 below.

TABLE 27: PLAY TASK - CORRELATION BETWEEN AGE AND BROAD CODES PERFORMANCE (Spearman’s Rank Correlation)

<table>
<thead>
<tr>
<th>CODE</th>
<th>MJ</th>
<th>MI</th>
<th>LD</th>
<th>SP</th>
</tr>
</thead>
<tbody>
<tr>
<td>APPROPRIATE</td>
<td>rho = 0.038</td>
<td>rho = -0.159</td>
<td>rho = 0.138</td>
<td>rho = 0.402</td>
</tr>
<tr>
<td></td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
<td>p = 0.0210</td>
</tr>
<tr>
<td>INAPPROPRIATE</td>
<td>rho = -0.002</td>
<td>rho = 0.405</td>
<td>rho = -0.128</td>
<td>rho = -0.483</td>
</tr>
<tr>
<td></td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
<td>p = 0.0063</td>
</tr>
<tr>
<td>AMBIGUOUS</td>
<td>rho = 0.014</td>
<td>rho = -0.016</td>
<td>rho = -0.075</td>
<td>rho = -0.155</td>
</tr>
<tr>
<td></td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
</tbody>
</table>
As evidenced by these correlation statistics, significant age-performance correlations on the broad codes "appropriate" and "inappropriate" were obtained only for the SP group. A positive correlation was recorded between SP age and "appropriate" performance. A negative correlation was recorded between SP age and "inappropriate" performance.

Further Spearman's Rank Correlations were calculated in order to examine the relationship between TROG (Bishop, 1983) scores and broad code performance for the two mainstream groups. (See Appendix B for individual TROG scores.) These results are presented in Table 28 below.

**TABLE 28: PLAY TASK - CORRELATION BETWEEN TROG SCORES AND BROAD CODES PERFORMANCE**
*(Spearman's Rank Correlation)*

<table>
<thead>
<tr>
<th>CODE</th>
<th>MJ</th>
<th>MI</th>
</tr>
</thead>
<tbody>
<tr>
<td>APPROPRIATE</td>
<td>rho = -0.102 n.s.</td>
<td>rho = 0.255 n.s.</td>
</tr>
<tr>
<td>INAPPROPRIATE</td>
<td>rho = 0.232 n.s.</td>
<td>rho = -0.382 n.s.</td>
</tr>
<tr>
<td>AMBIGUOUS</td>
<td>rho = -0.074 n.s.</td>
<td>rho = -0.236 n.s.</td>
</tr>
</tbody>
</table>

No significant correlations were found between TROG scores and broad code results on the play task.
(b) DEFINITION TASK

The within-group results for the MJ, MI, LD and SP groups are presented in Tables 29, 30, 31 and 32, respectively.

<table>
<thead>
<tr>
<th>SUBJ</th>
<th>APPROPRIATE</th>
<th>INAPPROPRIATE</th>
<th>AMBIGUOUS</th>
<th>OFS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MJ1</td>
<td>7</td>
<td>1 1 1 1 1 1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>MJ2</td>
<td>7</td>
<td>2 2 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MJ3</td>
<td>10</td>
<td>1</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>MJ4</td>
<td>7</td>
<td>3 1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>MJ5</td>
<td>6</td>
<td>2 1 1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>MJ6</td>
<td>6</td>
<td>1 1 1 1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>MJ7</td>
<td>8</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MJ8</td>
<td>6</td>
<td>3 1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>MJ9</td>
<td>6</td>
<td>2</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>MJ10</td>
<td>6</td>
<td>5 1</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>MJ11</td>
<td>9</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MJ12</td>
<td>7</td>
<td>3 1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>MJ13</td>
<td>6</td>
<td>2 1 1</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>MJ14</td>
<td>3</td>
<td>3 2 2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>MJ15</td>
<td>7</td>
<td>2 1 2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>MEAN</td>
<td>6.73</td>
<td>2.33 0.20 0.13 0.27 0.07 0.47 0.67</td>
<td>0.93 0.20</td>
<td></td>
</tr>
<tr>
<td>s.d.</td>
<td>1.58</td>
<td>1.18 0.41 0.35 0.46 0.26 0.74 0.82</td>
<td>1.10 0.41</td>
<td></td>
</tr>
</tbody>
</table>
### TABLE 30: DEFINITION TASK - MI GROUP RESULTS

| SUBJ | ID | INF | LIT | INF | LIT | FUZ | REP | NARR | REP-NARR | DON'T KNOW | AMB-ID/LIT | AMB-ID/OFS | AMB-LIT/OFS | OFS |
|------|----|-----|-----|-----|-----|-----|-----|------|---------|----------|------------|------------|------------|------------|-----|
| MI1  | 6  | 2   |     |     |     | 1   | 1   | 2    |          |           |            |            |            |            |     |
| MI2  | 9  | 1   |     |     |     |     | 1   |      |          |           |            |            |            |     |
| MI3  | 7  | 3   |     |     |     |     | 1   |      |          | 1         |            |            |            |     |
| MI4  | 7  | 1   |     |     |     | 1   |     |      |          | 3         |            |            |            |     |
| MI5  | 7  |     |     |     |     |     | 2   |      |          |           |            |            |            | 3   |
| MI6  | 9  | 2   |     |     |     |     |     |      |          | 1         |            |            |            |     |
| MI7  | 4  | 4   | 1   |     |     | 1   |     |      |          | 2         |            |            |            |     |
| MI8  | 4  | 5   | 1   |     |     |     | 1   |      |          |           |            |            |            |     |
| MI9  | 3  | 1   |     |     |     |     | 2   | 4    | 1       | 1         |            |            |            |     |
| MI10 | 4  | 4   | 1   |     |     |     |     | 1    | 1       |           |            |            |            |     |
| MI11 | 5  | 2   | 2   |     |     |     | 1   | 1    | 1       |           |            |            |            | 1   |
| MI12 | 2  | 2   | 1   | 1    |     |     |     | 1    | 4       | 1         |            |            |            |     |
| MI13 | 6  | 1   |     | 3    |     |     |     | 2    |          |           |            |            |            |     |
| MI14 | 4  | 1   | 1   | 1    | 1    |     |     | 2    |          |           |            |            |            |     |
| MI15 | 5  | 2   | 2   | 1    | 1    |     |     | 2    |          |           |            |            |            |     |
| MEAN | 5.47 | 2.07 | 0.47 | 0.60 | 0.27 | 0.20 | 0.13 | 1.73 | 0.33 | 0.53 | 0.07 | 0.13 |
| S.D. | 2.07 | 1.39 | 0.64 | 0.91 | 0.46 | 0.56 | 0.35 | 1.22 | 0.62 | 0.92 | 0.26 | 0.35 |

ID = "idiomatic"; INF ID = "inferred idiomatic"; LIT = "literal"; INF LIT = "inferred literal"; FUZ = "fuzzy"; REP = "repeat"; NARR = "narrative"; REP-NARR = "repeat-narrative"; AMB-ID/LIT = "ambiguous-idiomatic/literal"; AMB-ID/OFS = "ambiguous-idiomatic/other figurative sense"; AMB-LIT/OFS = "ambiguous-literal/other figurative sense"; OFS = "other figurative sense".

164
### TABLE 31: DEFINITION TASK – LD GROUP RESULTS

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>LD1</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LD2</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LD3</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LD4</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LD5</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LD6</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>LD7</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LD8</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LD9</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>LD10</td>
<td>5</td>
<td>4</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>LD11</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LD12</td>
<td>5</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>LD13</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>MEAN</td>
<td>3.87</td>
<td>1.73</td>
<td>0.53</td>
<td>0.47</td>
<td>1.47</td>
<td>0.20</td>
<td>1.13</td>
<td>1.20</td>
<td>0.40</td>
<td>0.67</td>
<td>0.13</td>
<td>0.20</td>
<td>0.90</td>
<td>0.35</td>
</tr>
<tr>
<td>S.D.</td>
<td>1.73</td>
<td>1.16</td>
<td>0.83</td>
<td>0.64</td>
<td>1.25</td>
<td>0.41</td>
<td>1.51</td>
<td>1.21</td>
<td>0.74</td>
<td>0.90</td>
<td>0.35</td>
<td>0.41</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ID = "idiomatic"; INF ID = "inferred idiomatic"; LIT = "literal"; INF LIT = "inferred literal"; FUZ = "fuzzy"; REP = "repeat"; NARR = "narrative"; REP-NARR = "repeat-narrative"; AMB-ID/LIT = "ambiguous-idiomatic/literal"; AMB-ID/OFS = "ambiguous-idiomatic/other figurative sense"; AMB-LIT/OFS = "ambiguous-literal/other figurative sense"; OFS = "other figurative sense".

165
### TABLE 32: DEFINITION TASK - SP GROUP RESULTS

<table>
<thead>
<tr>
<th>SUBJECT</th>
<th>APPROPRIATE</th>
<th>INAPPROPRIATE</th>
<th>AMBIGUOUS</th>
<th>OFS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ID</td>
<td>INF ID</td>
<td>LIT</td>
<td>INF LIT</td>
</tr>
<tr>
<td>SP1</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>SP2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>SP4</td>
<td>5</td>
<td>4</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>SP5</td>
<td>8</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SP6</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>SP7</td>
<td>1</td>
<td></td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>SP8</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>SP9</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>SP10</td>
<td>4</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SP11</td>
<td>7</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SP12</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>SP13</td>
<td>4</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SP14</td>
<td>3</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SP15</td>
<td>5</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SP16</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SP17</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SP18</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>SP3</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>SP19</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>SP20</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>SP21</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

continued overleaf
In order to examine the relationship between age and performance, Spearman's Rank Correlation was calculated for each group. The results are presented in Table 33 below.

**TABLE 33: DEFINITION TASK - CORRELATION BETWEEN AGE AND BROAD CODES PERFORMANCE**

(Spearman's Rank Correlation)

<table>
<thead>
<tr>
<th>CODE</th>
<th>MJ</th>
<th>MI</th>
<th>LD</th>
<th>SP</th>
</tr>
</thead>
<tbody>
<tr>
<td>APPROPRIATE</td>
<td>rho = 0.329 n.s.</td>
<td>rho = 0.367 n.s.</td>
<td>rho = -0.152 n.s.</td>
<td>rho = 0.562 p = .0014</td>
</tr>
<tr>
<td>INAPPROPRIATE</td>
<td>rho = -0.126 n.s.</td>
<td>rho = 0.156 n.s.</td>
<td>rho = 0.019 n.s.</td>
<td>rho = -0.584 p = .0009</td>
</tr>
<tr>
<td>AMBIGUOUS</td>
<td>rho = -0.347 n.s.</td>
<td>rho = -0.300 n.s.</td>
<td>rho = 0.006 n.s.</td>
<td>rho = 0.208 n.s.</td>
</tr>
</tbody>
</table>
As evidenced by these results, the age-performance correlations were significant only for the SP group on "appropriate" (positive correlation) and "inappropriate" (negative correlation) broad codes.

Spearman's Rank Correlation coefficients for the TROG (Bishop, 1983) - performance relationship were calculated for the two mainstream groups. (See Appendix B for individual TROG scores.) The results are presented in Table 34 below.

<table>
<thead>
<tr>
<th>CODE</th>
<th>MJ</th>
<th>MI</th>
</tr>
</thead>
<tbody>
<tr>
<td>APPROPRIATE</td>
<td>rho = 0.199</td>
<td>rho = 0.466</td>
</tr>
<tr>
<td></td>
<td>n.s.</td>
<td>p = .0401</td>
</tr>
<tr>
<td>INAPPROPRIATE</td>
<td>rho = -0.116</td>
<td>rho = -0.597</td>
</tr>
<tr>
<td></td>
<td>n.s.</td>
<td>p = .0094</td>
</tr>
<tr>
<td>AMBIGUOUS</td>
<td>rho = -0.095</td>
<td>rho = -0.087</td>
</tr>
<tr>
<td></td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

A significant positive correlation was revealed between the MI TROG scores and performance on the "appropriate" codes. The negative correlation between MI TROG scores and "inappropriate" codes was also significant. No other significant correlations emerged on the definition task.
In order to assess the meaningfulness of the diagnostic distinction between "SPD" and "ASP", the results of these subgroups were examined.

(a) PLAY TASK

The breakdown of play task broad codes for the SP group and subgroups is presented in Table 35 below. The differences between the subgroups (as tested on the Mann-Whitney Test) are also presented.

<table>
<thead>
<tr>
<th>GROUP</th>
<th>APPROPRIATE</th>
<th>INAPPROPRIATE</th>
<th>AMBIGUOUS</th>
<th>OTHER FIGURATIVE SENSE</th>
<th>TOTAL NO. ITEMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;SPD&quot;</td>
<td>134</td>
<td>41</td>
<td>29</td>
<td>14.2%</td>
<td>204</td>
</tr>
<tr>
<td></td>
<td>65.7%</td>
<td>20.1%</td>
<td>14.2%</td>
<td>65.7%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>x = 7.88</td>
<td>x = 2.41</td>
<td>x = 1.71</td>
<td>x = 2.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td>s.d. = 2.09</td>
<td>s.d. = 1.54</td>
<td>s.d. = 1.26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&quot;ASP&quot;</td>
<td>56</td>
<td>37</td>
<td>15</td>
<td>13.9%</td>
<td>108</td>
</tr>
<tr>
<td></td>
<td>51.9%</td>
<td>34.3%</td>
<td>13.9%</td>
<td>51.9%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>x = 6.22</td>
<td>x = 4.11</td>
<td>x = 1.67</td>
<td>x = 2.49</td>
<td></td>
</tr>
<tr>
<td></td>
<td>s.d. = 2.49</td>
<td>s.d. = 2.42</td>
<td>s.d. = 1.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL SP</td>
<td>190</td>
<td>78</td>
<td>44</td>
<td>14.1%</td>
<td>312</td>
</tr>
<tr>
<td></td>
<td>60.9%</td>
<td>25%</td>
<td>14.1%</td>
<td>60.9%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>x = 7.31</td>
<td>x = 3</td>
<td>x = 1.69</td>
<td>x = 2.33</td>
<td></td>
</tr>
<tr>
<td></td>
<td>s.d. = 2.33</td>
<td>s.d. = 2.02</td>
<td>s.d. = 1.32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&quot;SPD&quot;-&quot;ASP&quot; DIFFERENCE</td>
<td>U = 41 n.s.</td>
<td>U = 42.5 n.s.</td>
<td>U = 72.5 n.s.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Percentages reflect the proportion of the total number of items for the given group including items coded "appropriate", "inappropriate", "ambiguous" and "other figurative sense" i.e. 204 for the "SPD" group, 108 for the "ASP" group, and 312 for the whole SP group.
The results of the Mann-Whitney Test reveal that the "ASP" subgroup scored fewer "appropriate" items and more "inappropriate" items on the play task than did the "SPD" subgroup. These differences approached significance but did not reach it ("appropriate" p = 0.0529; "inappropriate" p = 0.0630). The two subgroups did not differ significantly in the number of "ambiguous" items scored on the play task.

"SPD" and "ASP" subgroup results on each of the "appropriate" narrow codes and the results of the Mann-Whitney Tests applied to these data are presented in Table 36 below.

<table>
<thead>
<tr>
<th>GROUP</th>
<th>IDIOMATIC</th>
<th>INFERRED IDIOMATIC</th>
<th>APPROPRIATE NO ACTION</th>
<th>TOTAL NO. APPROPRIATE ITEMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;SPD&quot;</td>
<td>114</td>
<td>19</td>
<td>1</td>
<td>134</td>
</tr>
<tr>
<td></td>
<td>55.9%</td>
<td>9.3%</td>
<td>0.5%</td>
<td>65.7%</td>
</tr>
<tr>
<td>x=6.71</td>
<td>x=1.12</td>
<td>x=0.99</td>
<td>s. d. = 2.23</td>
<td>s. d. = 0.99</td>
</tr>
<tr>
<td>s. d. = 2.23</td>
<td>s. d. = 0.99</td>
<td>x=0.06</td>
<td>s. d. = 0.24</td>
<td></td>
</tr>
<tr>
<td>&quot;ASP&quot;</td>
<td>43</td>
<td>13</td>
<td>1</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>39.8%</td>
<td>12%</td>
<td>51.9%</td>
<td>51.9%</td>
</tr>
<tr>
<td>x=4.78</td>
<td>x=1.44</td>
<td>x=1.59</td>
<td>s. d. = 1.86</td>
<td>s. d. = 1.59</td>
</tr>
<tr>
<td>s. d. = 1.86</td>
<td>s. d. = 1.59</td>
<td>x=6.22</td>
<td>s. d. = 2.49</td>
<td></td>
</tr>
<tr>
<td>TOTAL SP</td>
<td>157</td>
<td>32</td>
<td>1</td>
<td>190</td>
</tr>
<tr>
<td></td>
<td>50.3%</td>
<td>10.3%</td>
<td>0.3%</td>
<td>60.9%</td>
</tr>
<tr>
<td>x=6.04</td>
<td>x=1.23</td>
<td>x=0.04</td>
<td>s. d. = 2.27</td>
<td>s. d. = 2.33</td>
</tr>
<tr>
<td>s. d. = 2.27</td>
<td>s. d. = 1.21</td>
<td>s. d. = 0.20</td>
<td>s. d. = 2.33</td>
<td></td>
</tr>
<tr>
<td>&quot;SPD&quot;-&quot;ASP&quot; DIFFERENCE</td>
<td>U = 36.5</td>
<td>U = 72</td>
<td>U = 72</td>
<td>U = 41</td>
</tr>
<tr>
<td>p = .0294</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

Percentages reflect the proportion of the total number of items for the given group including items coded "appropriate", "inappropriate", "ambiguous" and "other figurative sense" i.e. 204 for the "SPD" group, 108 for the "ASP" group, and 312 for the whole SP group.

The results indicate that the "ASP" subgroup scored significantly fewer "idiomatic" items on the play task than did the "SPD" subgroup. The differences in "inferred idiomatic" and "appropriate no action" items were not significant.
The frequencies and differences (as tested with the Mann-Whitney Test) in the "inappropriate" narrow codes for the "SPD" and "ASP" subgroups are presented in Table 37 below.

<table>
<thead>
<tr>
<th>GROUP</th>
<th>LITERAL</th>
<th>INFERRERED LITERAL</th>
<th>INAPPROPRIATE NO ACTION</th>
<th>FUZZY</th>
<th>TOTAL NO. IN-APPROPRIATE ITEMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;SPD&quot;</td>
<td>10</td>
<td>5</td>
<td>2</td>
<td>24</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>4.9%</td>
<td>2.5%</td>
<td>1%</td>
<td>11.8%</td>
<td>20.1%</td>
</tr>
<tr>
<td></td>
<td>x = 0.59</td>
<td>x = 0.29</td>
<td>x = 0.12</td>
<td>x = 1.42</td>
<td>x = 0.87</td>
</tr>
<tr>
<td></td>
<td>s. d. = 0.71</td>
<td>s. d. = 0.47</td>
<td>s. d. = 0.33</td>
<td>s. d. = 0.87</td>
<td>s. d. = 1.54</td>
</tr>
<tr>
<td>&quot;ASP&quot;</td>
<td>5</td>
<td>4</td>
<td>2</td>
<td>26</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>4.6%</td>
<td>3.7%</td>
<td>1.9%</td>
<td>24.1%</td>
<td>34.3%</td>
</tr>
<tr>
<td></td>
<td>x = 0.56</td>
<td>x = 0.44</td>
<td>x = 0.22</td>
<td>x = 2.89</td>
<td>x = 4.11</td>
</tr>
<tr>
<td></td>
<td>s. d. = 0.88</td>
<td>s. d. = 0.73</td>
<td>s. d. = 0.44</td>
<td>s. d. = 2.26</td>
<td>s. d. = 2.42</td>
</tr>
<tr>
<td>TOTAL SP</td>
<td>15</td>
<td>9</td>
<td>4</td>
<td>50</td>
<td>78</td>
</tr>
<tr>
<td></td>
<td>4.8%</td>
<td>2.9%</td>
<td>1.3%</td>
<td>16X</td>
<td>25%</td>
</tr>
<tr>
<td></td>
<td>x = 0.58</td>
<td>x = 0.35</td>
<td>x = 0.15</td>
<td>x = 1.92</td>
<td>x = 3</td>
</tr>
<tr>
<td></td>
<td>s. d. = 0.76</td>
<td>s. d. = 0.56</td>
<td>s. d. = 0.37</td>
<td>s. d. = 1.62</td>
<td>s. d. = 2.02</td>
</tr>
<tr>
<td>&quot;SPD&quot;-&quot;ASP&quot;</td>
<td>U = 71</td>
<td>U = 71</td>
<td>U = 68.5</td>
<td>U = 39</td>
<td>U = 42.5</td>
</tr>
<tr>
<td>DIFFERENCE</td>
<td>n. s.</td>
<td>n. s.</td>
<td>n. s.</td>
<td>p = .0363</td>
<td>n. s.</td>
</tr>
</tbody>
</table>

Percentages reflect the proportion of the total number of items for the given group including items coded "appropriate", "inappropriate", "ambiguous" and "other figurative sense" i.e. 204 for the "SPD" group, 108 for the "ASP" group, and 312 for the whole SP group.

The results indicate that the "ASP" subgroup scored significantly more "fuzzy" items on the play task than did the "SPD" subgroup. The differences in the number of "literal", "inferred literal" and "inappropriate no action" items did not reach significance.

All of the "ambiguous" codes scored on the play task were "ambiguous-idiomatic/literal". The results for the latter narrow code are thus the same as for the broad code "ambiguous" (see Table 35).
Figure 8 below is a graphic representation of the play task broad code results for the "SPD" and "ASP" subgroups and for the other three groups. These frequencies are presented in Tables 4 and 35 above. Since only one subject scored "other figurative sense" on the play task, this code is not represented in Figure 8.

![Figure 8](image)

The "SPD" and "ASP" subgroup results were compared with those of the MJ, MI and LD groups, respectively, on a series of Mann-Whitney Tests. The broad code differences on the play task are presented in Table 38 below.
The results indicate that the "SPD" subgroup scored significantly fewer "appropriate" items than did the MJ and MI groups, respectively. The "SPD" subgroup also scored fewer "appropriate" items than did the LD group, but this difference did not reach significance.

The number of "inappropriate" items scored by the "SPD" subgroup on the play task was significantly greater than that scored by the MJ, MI and LD groups, respectively.

The number of "ambiguous" items scored by the "SPD" subgroup was not significantly different from that scored by the MJ, MI and LD groups, respectively.

The "ASP" subgroup scored significantly fewer "appropriate" items and significantly more "inappropriate" items than did each of the MJ, MI and LD groups. They did not differ significantly from any of these three groups in the number of "ambiguous" items scored.
The differences between the "SPD"/"ASP" subgroups and each of the other three groups on the "appropriate" narrow codes were tested with a series of Mann-Whitney Tests. The results are presented in Table 39 below. The frequencies for the MJ, MI and LD groups for these codes are presented in Table 6. The frequencies for the "SPD" and "ASP" subgroups are presented in Table 36.

<table>
<thead>
<tr>
<th>GROUPS COMPARED</th>
<th>IDIOMATIC</th>
<th>INFERRED IDIOMATIC</th>
<th>APPROPRIATE NO ACTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;SPD&quot; - MJ</td>
<td>U = 58.5</td>
<td>U = 109</td>
<td>U = 126</td>
</tr>
<tr>
<td></td>
<td>p = .0041</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td>&quot;ASP&quot; - MJ</td>
<td>U = 9.5</td>
<td>U = 57</td>
<td>U = 63</td>
</tr>
<tr>
<td></td>
<td>p = .0002</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td>&quot;SPD&quot; - MI</td>
<td>U = 95.5</td>
<td>U = 111.5</td>
<td>U = 118</td>
</tr>
<tr>
<td></td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td>&quot;ASP&quot; - MI</td>
<td>U = 19.5</td>
<td>U = 67.5</td>
<td>U = 58.5</td>
</tr>
<tr>
<td></td>
<td>p = .0019</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td>&quot;SPD&quot; - LD</td>
<td>U = 106</td>
<td>U = 108</td>
<td>U = 109.5</td>
</tr>
<tr>
<td></td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td>&quot;ASP&quot; - LD</td>
<td>U = 22</td>
<td>U = 64.5</td>
<td>U = 54</td>
</tr>
<tr>
<td></td>
<td>p = .0031</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

As evidenced by these results, the "SPD" subgroup scored significantly fewer "idiomatic" items than did the MJ group. They did not differ significantly on this code from either the MI or LD group. The number of "SPD" items coded "inferred idiomatic" or "appropriate no action" did not differ significantly from those of the MJ, MI or LD group, respectively.

The "ASP" subgroup scored significantly fewer "idiomatic" items on the play task than did the MJ, MI and LD groups, respectively. They did not differ significantly from any of these groups on the codes "inferred idiomatic" and "appropriate no action".
The differences on "inappropriate" narrow codes between the "SPD"/"ASP" subgroups and the MJ, MI and LD groups, respectively, are presented in Table 40 below. The "inappropriate" narrow code frequency data for MJ, MI and LD groups is presented in Table 8. The data for the "SPD" and "ASP" subgroups is presented in Table 37.

<table>
<thead>
<tr>
<th>GROUPS COMPARED</th>
<th>LITERAL</th>
<th>INFERRED LITERAL</th>
<th>INAPPROPRIATE NO ACTION</th>
<th>FUZZY</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;SPD&quot; - MJ</td>
<td>U = 112.5 n.s.</td>
<td>U = 107 n.s.</td>
<td>U = 121 n.s.</td>
<td>U = 48 p = .0015</td>
</tr>
<tr>
<td>&quot;ASP&quot; - MJ</td>
<td>U = 66 n.s.</td>
<td>U = 53 n.s.</td>
<td>U = 57 n.s.</td>
<td>U = 14.5 p = .0007</td>
</tr>
<tr>
<td>&quot;SPD&quot; - MI</td>
<td>U = 116.5 n.s.</td>
<td>U = 107 n.s.</td>
<td>U = 108.5 n.s.</td>
<td>U = 36.5 p = .0002</td>
</tr>
<tr>
<td>&quot;ASP&quot; - MI</td>
<td>U = 67.5 n.s.</td>
<td>U = 53 n.s.</td>
<td>U = 64.5 n.s.</td>
<td>U = 12 p = .0003</td>
</tr>
<tr>
<td>&quot;SPD&quot; - LD</td>
<td>U = 90 p = .0448</td>
<td>U = 98.5 n.s.</td>
<td>U = 121 n.s.</td>
<td>U = 84.5 n.s.</td>
</tr>
<tr>
<td>&quot;ASP&quot; - LD</td>
<td>U = 55.5 n.s.</td>
<td>U = 49 n.s.</td>
<td>U = 57 n.s.</td>
<td>U = 22 p = .0051</td>
</tr>
</tbody>
</table>

These results indicate that there were no significant differences in the number of "literal", "inferred literal" or "inappropriate no action" items scored by the "SPD" subgroup as opposed to either the MJ or the MI group. Significantly more "fuzzy" items were scored by the "SPD" subgroup than by the MJ or MI group, however.

The "SPD" subgroup scored significantly more "literal" items than did the LD group, but the two groups did not differ significantly on "inferred literal", "inappropriate no action" or "fuzzy" scores.
The "ASP" subgroup scored significantly more "fuzzy" items than did the MJ, MI and LD groups, respectively. They did not differ significantly from any of these groups on "literal", "inferred literal" or "inappropriate no action" codes, however.

A Kth neighbour discriminant analysis assigned a total of 64 (90.1%) of the subjects accurately. An analysis of these results is presented in Table 41 below.

<table>
<thead>
<tr>
<th>GROUP</th>
<th>NO. CORRECTLY ASSIGNED</th>
<th>NO. INCORRECTLY ASSIGNED</th>
<th>ANALYSIS OF SUBJECTS INCORRECTLY ASSIGNED</th>
</tr>
</thead>
<tbody>
<tr>
<td>MJ</td>
<td>15 (100%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MI</td>
<td>13 (86.7%)</td>
<td>2 (13.3%)</td>
<td>MI5 GROUPED WITH MJ</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MI7 GROUPED WITH MJ</td>
</tr>
<tr>
<td>LD</td>
<td>14 (93.3%)</td>
<td>1 (6.7%)</td>
<td>LD2 GROUPED WITH MJ</td>
</tr>
<tr>
<td>&quot;SPD&quot;</td>
<td>14 (82.4%)</td>
<td>3 (17.7%)</td>
<td>SP10 GROUPED WITH LD SP16 GROUPED WITH MJ</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SP17 GROUPED WITH MI SP25 GROUPED WITH MJ</td>
</tr>
<tr>
<td>&quot;ASP&quot;</td>
<td>8 (88.9%)</td>
<td>1 (11.1%)</td>
<td>SP25 GROUPED WITH MJ</td>
</tr>
<tr>
<td>TOTAL</td>
<td>64 (90.1%)</td>
<td>7 (9.9%)</td>
<td></td>
</tr>
</tbody>
</table>
(b) DEFINITION TASK

The frequency with which each of broad codes was scored by the "SPD" and "ASP" subgroups on the definition task and the differences between the subgroups (calculated on the Mann-Whitney Test) are presented in Table 42 below.

TABLE 42: DEFINITION TASK - "SPD"/"ASP" SUBGROUPS: FREQUENCIES AND DIFFERENCES ON BROAD CODES

<table>
<thead>
<tr>
<th>GROUP</th>
<th>APPROPRIATE</th>
<th>INAPPROPRIATE</th>
<th>AMBIGUOUS</th>
<th>OTHER FIG’TIVE SENSE</th>
<th>TOTAL NO. ITEMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;SPD&quot;</td>
<td>112</td>
<td>70</td>
<td>21</td>
<td></td>
<td>203*</td>
</tr>
<tr>
<td></td>
<td>55.2%</td>
<td>34.5%</td>
<td>10.4%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>x = 6.59</td>
<td>x = 4.12</td>
<td>x = 1.24</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>s. d. = 2.53</td>
<td>s. d. = 2.62</td>
<td>s. d. = 1.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&quot;ASP&quot;</td>
<td>38</td>
<td>62</td>
<td>3</td>
<td></td>
<td>108</td>
</tr>
<tr>
<td></td>
<td>35.2%</td>
<td>57.4%</td>
<td>2.8%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>x = 4.22</td>
<td>x = 6.89</td>
<td>x = 0.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>s. d. = 2.33</td>
<td>s. d. = 2.37</td>
<td>s. d. = 0.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL SP</td>
<td>150</td>
<td>132</td>
<td>24</td>
<td></td>
<td>311*</td>
</tr>
<tr>
<td></td>
<td>48.2%</td>
<td>42.4%</td>
<td>7.7%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>x = 5.77</td>
<td>x = 5.08</td>
<td>x = 0.92</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>s. d. = 2.67</td>
<td>s. d. = 2.83</td>
<td>s. d. = 0.98</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&quot;SPD&quot;-&quot;ASP&quot; DIFFERENCE</td>
<td>U = 35.5</td>
<td>U = 32</td>
<td>U = 34.5</td>
<td>U = 65</td>
<td>p = .0254</td>
</tr>
</tbody>
</table>

Percentages reflect the proportion of the total number of items for the given group including items coded "appropriate", "inappropriate", "ambiguous" and "other figurative sense" i.e. 203 for the "SPD" group, 108 for the "ASP" group, and 311 for the whole SP group.

* one item omitted in this group on the definition task

As is evident from these results, the "SPD" subgroup scored significantly more "appropriate" items and significantly fewer "inappropriate" items on the definition task than did the "ASP" subgroup. The "SPD" subgroup also scored significantly more "ambiguous" and significantly fewer "OFS" items than did the "ASP" subgroup.
The "SPD"/"ASP" subgroup frequencies and differences (as tested on the Mann-Whitney Test) for the "appropriate" narrow codes are presented in Table 43 below.

### TABLE 43: DEFINITION TASK - "SPD"/"ASP" SUBGROUPS: FREQUENCIES AND DIFFERENCES ON "APPROPRIATE" NARROW CODES

<table>
<thead>
<tr>
<th>GROUP</th>
<th>IDIOMATIC</th>
<th>INFERRRED IDIOMATIC</th>
<th>TOTAL NO. APPROPRIATE ITEMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;SPD&quot;</td>
<td>73</td>
<td>39</td>
<td>112</td>
</tr>
<tr>
<td></td>
<td>36%</td>
<td>19.2%</td>
<td>55.2%</td>
</tr>
<tr>
<td></td>
<td>x = 4.30</td>
<td>x = 2.30</td>
<td>x = 6.59</td>
</tr>
<tr>
<td></td>
<td>s. d. = 2.82</td>
<td>s. d. = 1.61</td>
<td>s. d. = 2.53</td>
</tr>
<tr>
<td>&quot;ASP&quot;</td>
<td>21</td>
<td>17</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>19.4%</td>
<td>15.7%</td>
<td>35.2%</td>
</tr>
<tr>
<td></td>
<td>x = 2.33</td>
<td>x = 1.89</td>
<td>x = 4.22</td>
</tr>
<tr>
<td></td>
<td>s. d. = 2.12</td>
<td>s. d. = 0.78</td>
<td>s. d. = 2.33</td>
</tr>
<tr>
<td>TOTAL SP</td>
<td>94</td>
<td>56</td>
<td>150</td>
</tr>
<tr>
<td></td>
<td>30.2%</td>
<td>18%</td>
<td>48.2%</td>
</tr>
<tr>
<td></td>
<td>x = 3.62</td>
<td>x = 2.15</td>
<td>x = 5.77</td>
</tr>
<tr>
<td></td>
<td>s. d. = 2.73</td>
<td>s. d. = 1.38</td>
<td>s. d. = 2.67</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>&quot;SPD&quot;-&quot;ASP&quot; DIFFERENCE</th>
<th>U = 42.5</th>
<th>U = 68.5</th>
<th>U = 35.5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n.s.</td>
<td>n.s.</td>
<td>p = .0254</td>
</tr>
</tbody>
</table>

Percentages reflect the proportion of the total number of items for the given group including items coded "appropriate", "inappropriate", "ambiguous" and "other figurative sense" i.e. 203 for the "SPD" group, 108 for the "ASP" group, and 311 for the whole SP group.

These results indicate that the number of "SPD" items scored "idiomatic" or "inferred idiom" did not differ significantly from that of the "ASP" subgroup. The total number of "appropriate" items was, however, significantly higher in the "SPD" than "ASP" subgroup.
The frequencies and differences (as tested on the Mann-Whitney Test) between the "SPD" and "ASP" subgroups on the "inappropriate" narrow codes are presented in Table 44 below.

<table>
<thead>
<tr>
<th>GROUP</th>
<th>LITERAL</th>
<th>INFERRED LITERAL</th>
<th>FUZZY</th>
<th>REPEAT</th>
<th>NARRATIVE</th>
<th>REPEAT-NARRATIVE</th>
<th>DON'T KNOW</th>
<th>TOTAL NO. INAPPROPR. ITEMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;SPD&quot;</td>
<td>6 3%</td>
<td>13 6.4%</td>
<td>10 4.9%</td>
<td>3 1.5%</td>
<td>23 11.3%</td>
<td>13 6.4%</td>
<td>2 1%</td>
<td>70 34.5%</td>
</tr>
<tr>
<td></td>
<td>x = 0.35</td>
<td>x = 0.77</td>
<td>x = 0.59</td>
<td>x = 0.18</td>
<td>x = 1.35</td>
<td>x = 0.77</td>
<td>x = 0.12</td>
<td>x = 4.12</td>
</tr>
<tr>
<td></td>
<td>sd = 0.49</td>
<td>sd = 1.30</td>
<td>sd = 0.62</td>
<td>sd = 0.53</td>
<td>sd = 1.17</td>
<td>sd = 1.20</td>
<td>sd = 0.49</td>
<td>sd = 2.62</td>
</tr>
<tr>
<td>&quot;ASP&quot;</td>
<td>6 5.6%</td>
<td>16 14.8%</td>
<td>24 22.2%</td>
<td>5 4.6%</td>
<td>9 8.3%</td>
<td>9 1.9%</td>
<td>2 1%</td>
<td>62 57.4%</td>
</tr>
<tr>
<td></td>
<td>x = 0.67</td>
<td>x = 1.78</td>
<td>x = 2.67</td>
<td>x = 0.56</td>
<td>x = 1.00</td>
<td>x = 0.22</td>
<td>x = 0.44</td>
<td>x = 2.37</td>
</tr>
<tr>
<td></td>
<td>sd = 1.00</td>
<td>sd = 1.39</td>
<td>sd = 1.58</td>
<td>sd = 1.01</td>
<td>sd = 0.71</td>
<td>sd = 0.44</td>
<td>sd = 0.44</td>
<td>sd = 2.37</td>
</tr>
<tr>
<td>TOTAL SP</td>
<td>12 3.9%</td>
<td>29 9.3%</td>
<td>34 10.9%</td>
<td>3 1%</td>
<td>28 9%</td>
<td>22 7.1%</td>
<td>4 1.3%</td>
<td>132 42.4%</td>
</tr>
<tr>
<td></td>
<td>x = 0.46</td>
<td>x = 1.12</td>
<td>x = 1.31</td>
<td>x = 0.12</td>
<td>x = 1.08</td>
<td>x = 0.85</td>
<td>x = 0.15</td>
<td>x = 5.08</td>
</tr>
<tr>
<td></td>
<td>sd = 0.71</td>
<td>sd = 1.40</td>
<td>sd = 1.44</td>
<td>sd = 0.43</td>
<td>sd = 1.16</td>
<td>sd = 1.05</td>
<td>sd = 0.46</td>
<td>sd = 2.83</td>
</tr>
<tr>
<td>DIFFERENCE</td>
<td>U = 66.5 n.s.</td>
<td>U = 38 p = 0.0280</td>
<td>U = 19 n.s.</td>
<td>U = 67.5 n.s.</td>
<td>U = 42 n.s.</td>
<td>U = 54 n.s.</td>
<td>U = 65 n.s.</td>
<td>U = 32 p = 0.0155</td>
</tr>
</tbody>
</table>

Percentages reflect the proportion of the total number of items for the given group including items coded "appropriate", "inappropriate", "ambiguous" and "other figurative sense" i.e. 203 for the "SPD" group, 108 for the "ASP" group, and 311 for the whole SP group.

The "SPD" subgroup was found to have scored significantly fewer "inferred literal" and "fuzzy" items on the definition task than did the "ASP" subgroup. The number of "SPD" items scored "narrative" was considerably greater than that scored by the "ASP" subgroup, but this difference narrowly missed significance (p = 0.05). The difference between the subgroups on "literal", "repeat", "repeat-narrative" and "don't know" items, respectively, did not reach significance.

179
The "SPD"/"ASP" frequencies and differences (calculated on the Mann-Whitney Test) for the "ambiguous" narrow codes are presented in Table 45 below.

### Table 45: Definition Task - "SPD"/"ASP" Subgroups: Frequencies and Differences on "Ambiguous" Narrow Codes

<table>
<thead>
<tr>
<th>GROUP</th>
<th>Ambiguous - Idiomatic/Literal</th>
<th>Ambiguous - Idiomatic/OFS</th>
<th>Ambiguous - Literal/OFS</th>
<th>Total No. Ambiguous Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;SPD&quot;</td>
<td>19 9.4% x = 1.12 s.d. = 0.99</td>
<td>1 0.5% x = 0.06 s.d. = 0.24</td>
<td>1 0.5% x = 0.06 s.d. = 0.24</td>
<td>21 10.4% x = 1.24 s.d. = 1.03</td>
</tr>
<tr>
<td>&quot;ASP&quot;</td>
<td>3 2.8% x = 0.33 s.d. = 0.50</td>
<td></td>
<td></td>
<td>3 2.8% x = 0.33 s.d. = 0.50</td>
</tr>
<tr>
<td>TOTAL SP</td>
<td>22 7.1% x = 0.85 s.d. = 0.93</td>
<td>1 0.3% 0.04 s.d. = 0.20</td>
<td>1 0.3% x = 0.04 s.d. = 0.20</td>
<td>24 7.7% x = 0.92 s.d. = 0.98</td>
</tr>
<tr>
<td>&quot;SPD&quot;-&quot;ASP&quot; DIFFERENCE</td>
<td>U = 33.5 n.s.</td>
<td>U = 72 n.s.</td>
<td>U = 72 n.s.</td>
<td>U = 34.5 p = .0158</td>
</tr>
</tbody>
</table>

Percentages reflect the proportion of the total number of items for the given group including items coded "appropriate", "inappropriate", "ambiguous" and "other figurative sense" i.e. 203 for the "SPD" group, 108 for the "ASP" group, and 311 for the whole SP group.

These results indicate that the "SPD" and "ASP" subgroups did not differ significantly on any of the "ambiguous" narrow codes on the definition task.
Figure 9 below is a graphic representation of the definition task broad code results for the "SPD" and "ASP" subgroups and for the other three groups. These frequencies are presented in Tables 11 and 42 above.

The broad code differences (tested on the Mann-Whitney Test) between the "SPD"/"ASP" subgroups and the MJ, MI and LD groups, respectively, are presented in Table 46 below.
As indicated by these results, the "SPD" subgroup scored significantly fewer "appropriate" items and significantly more "inappropriate" items than did the MJ group. The "SPD" subgroup also scored significantly fewer "OFS" items than did the MJ group. The two groups did not differ significantly in the number of "ambiguous" items scored.

The "SPD" and MI groups did not differ significantly in the number of "appropriate", "inappropriate" or "OFS" items scored on this task. The "SPD" subgroup did, however, score significantly more "ambiguous" items than did the MI group.

The "SPD" group scored significantly fewer "OFS" items than did the LD group, but the two groups did not differ significantly in the number of "appropriate", "inappropriate" or "ambiguous" items scored.

The "ASP" subgroup scored significantly fewer "appropriate" items and significantly more "inappropriate" items than did the MJ and
MI groups, respectively. No significant differences were found for the number of "ambiguous" or "OFS" items scored by the "ASP" subgroup as opposed to the MJ or MI groups.

The "ASP" subgroup did not differ significantly from the LD group on any of the broad codes.

The "appropriate" narrow code differences between the "SPD"/"ASP" subgroups and the MJ, MI and LD groups, respectively, are presented in Table 47 below. The frequency data for the MJ, MI and LD groups are presented in Table 13. The frequency data for the "SPD" and "ASP" groups are presented in Table 43.

<table>
<thead>
<tr>
<th>GROUPS COMPARED</th>
<th>IDIOMATIC</th>
<th>INFERRED IDIOMATIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;SPD&quot; - MJ</td>
<td>U = 56</td>
<td>U = 123</td>
</tr>
<tr>
<td></td>
<td>p = .0032</td>
<td>n.s.</td>
</tr>
<tr>
<td>&quot;ASP&quot; - MJ</td>
<td>U = 6</td>
<td>U = 50</td>
</tr>
<tr>
<td></td>
<td>p = .0001</td>
<td>n.s.</td>
</tr>
<tr>
<td>&quot;SPD&quot; - MI</td>
<td>U = 88</td>
<td>U = 120</td>
</tr>
<tr>
<td></td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td>&quot;ASP&quot; - MI</td>
<td>U = 20.5</td>
<td>U = 67.5</td>
</tr>
<tr>
<td></td>
<td>p = .0024</td>
<td>n.s.</td>
</tr>
<tr>
<td>&quot;SPD&quot; - LD</td>
<td>U = 124.5</td>
<td>U = 106</td>
</tr>
<tr>
<td></td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td>&quot;ASP&quot; - LD</td>
<td>U = 38</td>
<td>U = 63</td>
</tr>
<tr>
<td></td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

The results indicate that the "SPD" group scored significantly fewer "idiomatic" items on the definition task than did the MJ group. The two groups did not differ significantly in the number of "inferred idiomatic" items scored.
The "SPD" subgroup did not differ significantly from either the MI or LD group with respect to the number of "idiomatic" or "inferred idiomatic" items scored.

The "ASP" subgroup scored significantly fewer "idiomatic" items than did either the MJ or MI group. The number of "ASP" items coded "inferred idiomatic" did not differ significantly from that of the MJ or MI groups.

Neither the difference in "idiomatic" nor that in "inferred idiomatic" items scored by the "ASP" and LD groups reached significance.

The "inappropriate" narrow code differences between the "SPD"/"ASP" subgroups and the MJ, MI and LD groups, respectively, are presented in Table 48 below. The frequency data for the MJ, MI and LD groups are presented in Table 15. The frequency data for the "SPD" and "ASP" subgroups are presented in Table 44.

<table>
<thead>
<tr>
<th>GROUPS COMPARED</th>
<th>LITERAL</th>
<th>INFERRED LITERAL</th>
<th>FUZZY</th>
<th>REPEAT</th>
<th>NARRATIVE</th>
<th>REPEAT - NARRATIVE</th>
<th>DON'T KNOW</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;SPD&quot; - MJ</td>
<td>U = 108</td>
<td>n.s.</td>
<td>U = 92</td>
<td>n.s.</td>
<td>U = 120.5</td>
<td>U = 67.5</td>
<td>U = 123.5</td>
</tr>
<tr>
<td>&quot;ASP&quot; - MJ</td>
<td>U = 49.5</td>
<td>n.s.</td>
<td>U = 12.5</td>
<td>p = .0002</td>
<td>U = 11.5</td>
<td>U = 63</td>
<td>U = 67</td>
</tr>
<tr>
<td>&quot;SPD&quot; - MI</td>
<td>U = 118.5</td>
<td>n.s.</td>
<td>U = 124</td>
<td>n.s.</td>
<td>U = 125.5</td>
<td>U = 41</td>
<td>U = 64</td>
</tr>
<tr>
<td>&quot;ASP&quot; - MI</td>
<td>U = 63</td>
<td>n.s.</td>
<td>U = 29.5</td>
<td>p = .0163</td>
<td>U = 11.5</td>
<td>U = 53</td>
<td>U = 43.5</td>
</tr>
<tr>
<td>&quot;SPD&quot; - LD</td>
<td>U = 121</td>
<td>n.s.</td>
<td>U = 120</td>
<td>n.s.</td>
<td>U = 118.5</td>
<td>U = 103.5</td>
<td>U = 99</td>
</tr>
<tr>
<td>&quot;ASP&quot; - LD</td>
<td>U = 62</td>
<td>n.s.</td>
<td>U = 25.5</td>
<td>p = .0074</td>
<td>U = 37</td>
<td>U = 52</td>
<td>U = 64.5</td>
</tr>
</tbody>
</table>

TABLE 48: DEFINITION TASK - DIFFERENCES BETWEEN "SPD"/"ASP" SUBGROUPS AND MJ, MI AND LD GROUPS ON "INAPPROPRIATE" NARROW CODES (Mann-Whitney Test)
The results indicate that the "SPD" subgroup scored significantly more "narrative" items than did the MJ group. These two groups did not, however, differ significantly on any other of the "inappropriate" narrow codes.

The "SPD" subgroup also scored significantly more "narrative" and significantly fewer "repeat-narrative" items than did the MI group. None of the other "inappropriate" narrow code differences between these two groups reached significance.

In relation to the LD group, the "SPD" subgroup scored significantly fewer "fuzzy" items on the definition task. The other "inappropriate" narrow codes did not yield any significant differences between the "SPD" and LD groups.

The "ASP" subgroup scored significantly more "inferred literal" and "fuzzy" items than did either the MJ or MI group. The "ASP" subgroup did not differ significantly from the MJ or MI groups on any other "inappropriate" narrow codes.

The "ASP" subgroup scored significantly more "inferred literal" items than did the LD group. These two groups did not differ significantly on any other "inappropriate" narrow codes.
The differences on the code "ambiguous-idiomatic/literal" between the "SPD"/"ASP" subgroups and the MJ, MI and LD groups, respectively, are presented in Table 49 below. The frequency data for the MJ, MI and LD groups are presented in Table 17. The frequency data for the "SPD" and "ASP" subgroups are presented in Table 45.

The results indicate that the "SPD" group scored significantly more "ambiguous-idiomatic/literal" items on the definition task than did the MI group. The number of "ambiguous-idiomatic/literal" items scored by the "SPD" group did not differ significantly from that scored by the MJ or LD groups.

The "ASP" group did not differ significantly on "ambiguous-idiomatic/literal" from the MJ, MI or LD group, respectively.

As noted above, the "OFS" broad code does not break down into narrow codes.
A Kth neighbour discriminant analysis was performed on the definition task data with the SP group broken down into "SPD" and "ASP" subgroups. The results revealed that 68 (95.8%) of the subjects were correctly assigned to these five groups. The results of this analysis are presented in Table 50 below.

<table>
<thead>
<tr>
<th>GROUP</th>
<th>NO. CORRECTLY ASSIGNED</th>
<th>NO. INCORRECTLY ASSIGNED</th>
<th>ANALYSIS OF SUBJECTS INCORRECTLY ASSIGNED</th>
</tr>
</thead>
<tbody>
<tr>
<td>MJ</td>
<td>15 (100%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MI</td>
<td>15 (100%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LD</td>
<td>15 (100%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&quot;SPD&quot;</td>
<td>15 (88.2%)</td>
<td>2 (11.8%)</td>
<td>SP4 GROUPED WITH MJ SP14 GROUPED WITH MI</td>
</tr>
<tr>
<td>&quot;ASP&quot;</td>
<td>8 (88.9%)</td>
<td>1 (11.1%)</td>
<td>SP25 GROUPED WITH MI</td>
</tr>
<tr>
<td>TOTAL</td>
<td>68 (95.8%)</td>
<td>3 (4.2%)</td>
<td></td>
</tr>
</tbody>
</table>
(c) BETWEEN-TASKS DIFFERENCES

Between-tasks differences for the subgroups "SPD" and "ASP" on the broad codes are presented in Table 51 below.

<table>
<thead>
<tr>
<th>BROAD CODE</th>
<th>FEATURE</th>
<th>&quot;SPD&quot;</th>
<th>&quot;ASP&quot;</th>
<th>TOTAL SP</th>
</tr>
</thead>
<tbody>
<tr>
<td>APPROPRIATE</td>
<td>PLAY TASK</td>
<td>65.7%</td>
<td>51.9%</td>
<td>60.9%</td>
</tr>
<tr>
<td></td>
<td>DEFINITION TASK</td>
<td>55.2%</td>
<td>35.2%</td>
<td>48.2%</td>
</tr>
<tr>
<td></td>
<td>DIFFERENCE</td>
<td>W = 19</td>
<td>W = 8.5</td>
<td>W = 39</td>
</tr>
<tr>
<td></td>
<td></td>
<td>p = .0178</td>
<td>n.s.</td>
<td>p = .0023</td>
</tr>
<tr>
<td>INAPPROPRIATE</td>
<td>PLAY TASK</td>
<td>20.1%</td>
<td>34.3%</td>
<td>25%</td>
</tr>
<tr>
<td></td>
<td>DEFINITION TASK</td>
<td>34.5%</td>
<td>57.4%</td>
<td>42.4%</td>
</tr>
<tr>
<td></td>
<td>DIFFERENCE</td>
<td>W = 3.5</td>
<td>W = 2</td>
<td>W = 14.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>p = .0007</td>
<td>p = .0126</td>
<td>p = .0001</td>
</tr>
<tr>
<td>AMBIGUOUS</td>
<td>PLAY TASK</td>
<td>14.2%</td>
<td>13.9%</td>
<td>14.1%</td>
</tr>
<tr>
<td></td>
<td>DEFINITION TASK</td>
<td>10.4%</td>
<td>2.8%</td>
<td>7.7%</td>
</tr>
<tr>
<td></td>
<td>DIFFERENCE</td>
<td>W = 22</td>
<td>W = 0</td>
<td>W = 32</td>
</tr>
<tr>
<td></td>
<td></td>
<td>n.s.</td>
<td>p = .0277</td>
<td>p = .0352</td>
</tr>
<tr>
<td>OTHER FIGURATIVE SENSE</td>
<td>PLAY TASK</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DEFINITION TASK</td>
<td></td>
<td>4.6%</td>
<td>1.6%</td>
</tr>
<tr>
<td></td>
<td>DIFFERENCE</td>
<td>W = 0</td>
<td>W = 0</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

It is evident from these results that the "SPD" subgroup scored significantly more "appropriate" items and significantly fewer
"inappropriate" items on the play task than on the definition task. No significant between-tasks differences were found for this subgroup on the codes "ambiguous" or "other figurative sense".

The "ASP" subgroup scored significantly fewer "inappropriate" items on the play task than on the definition task. They also scored significantly more "ambiguous" items on the play task than on the definition task. Between-tasks differences on "appropriate" and "other figurative sense" were not significant for the "ASP" subgroup.
Table 52, below, presents the between-tasks differences for the "SPD" and "ASP" subgroups on the narrow codes "idiomatic", "inferred idiomatic", "literal", "inferred literal" and "fuzzy".

<table>
<thead>
<tr>
<th>NARROW CODE</th>
<th>FEATURE</th>
<th>&quot;SPD&quot;</th>
<th>&quot;ASP&quot;</th>
<th>TOTAL SP</th>
</tr>
</thead>
<tbody>
<tr>
<td>IDIOMATIC</td>
<td>PLAY TASK</td>
<td>55.9%</td>
<td>39.8%</td>
<td>50.3%</td>
</tr>
<tr>
<td></td>
<td>DEFINITION TASK</td>
<td>36%</td>
<td>19.4%</td>
<td>30.2%</td>
</tr>
<tr>
<td></td>
<td>DIFFERENCE</td>
<td>W = 6</td>
<td>W = 0</td>
<td>W = 6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>p = .0022</td>
<td>p = .0180</td>
<td>p = .0001</td>
</tr>
<tr>
<td>INFERRED IDIOMATIC</td>
<td>PLAY TASK</td>
<td>9.3%</td>
<td>12%</td>
<td>10.3%</td>
</tr>
<tr>
<td></td>
<td>DEFINITION TASK</td>
<td>19.2%</td>
<td>15.7%</td>
<td>18%</td>
</tr>
<tr>
<td></td>
<td>DIFFERENCE</td>
<td>W = 20</td>
<td>W = 13</td>
<td>W = 61.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>p = .0413</td>
<td>n.s.</td>
<td>p = .0348</td>
</tr>
<tr>
<td>LITERAL</td>
<td>PLAY TASK</td>
<td>4.9%</td>
<td>4.6%</td>
<td>4.8%</td>
</tr>
<tr>
<td></td>
<td>DEFINITION TASK</td>
<td>3%</td>
<td>5.6%</td>
<td>3.9%</td>
</tr>
<tr>
<td></td>
<td>DIFFERENCE</td>
<td>W = 0</td>
<td>W = 9</td>
<td>W = 20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td>INFERRED LITERAL</td>
<td>PLAY TASK</td>
<td>2.5%</td>
<td>3.7%</td>
<td>2.9%</td>
</tr>
<tr>
<td></td>
<td>DEFINITION TASK</td>
<td>6.4%</td>
<td>14.8%</td>
<td>9.3%</td>
</tr>
<tr>
<td></td>
<td>DIFFERENCE</td>
<td>W = 6</td>
<td>W = 1.5</td>
<td>W = 12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>n.s.</td>
<td>n.s.</td>
<td>p = .0192</td>
</tr>
<tr>
<td>FUZZY</td>
<td>PLAY TASK</td>
<td>11.8%</td>
<td>24.1%</td>
<td>16%</td>
</tr>
<tr>
<td></td>
<td>DEFINITION TASK</td>
<td>4.9%</td>
<td>22.2%</td>
<td>10.9%</td>
</tr>
<tr>
<td></td>
<td>DIFFERENCE</td>
<td>W = 2.5</td>
<td>W = 10</td>
<td>W = 32.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>p = .0108</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
</tbody>
</table>
The results indicate that both the "SPD" and "ASP" subgroups scored significantly more "idiomatic" items on the play task than on the definition task. The "SPD" group also scored significantly more "inferred idiomatic" items on the play task than on the definition task.

The number of items scored "literal" and "inferred literal" did not differ significantly across tasks for either subgroup.

The "SPD" subgroup scored significantly more "fuzzy" items on the play task than on the definition task. This difference did not reach significance for the "ASP" subgroup.

The differences between the number of "SPD"/"ASP" items scored "appropriate" and "inappropriate" on each of the two tasks are presented in Table 53 below. The frequency data for these codes are included in Table 51 above.

<table>
<thead>
<tr>
<th>GROUP</th>
<th>PLAY TASK</th>
<th>DEFINITION TASK</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;SPD&quot;</td>
<td>W = 4.5, p = .0007</td>
<td>W = 33.5, p = .0418</td>
</tr>
<tr>
<td>&quot;ASP&quot;</td>
<td>W = 12, n.s.</td>
<td>W = 7, n.s.</td>
</tr>
</tbody>
</table>

The "SPD" subgroup was found to have scored significantly more "appropriate" items than "inappropriate" items on both tasks.

The "ASP" subgroup scored more "appropriate" than "inappropriate" items on the play task and more "inappropriate" than "appropriate" items on the definition task. However, these differences did not reach significance.
6. BETWEEN-IDIOMS ANALYSIS

The proportion of each group scoring "appropriate" for each individual idiom is presented in Tables 54, 55, 56 and 57 below.

**TABLE 54: PLAY TASK - LEVEL OF "APPROPRIACY" FOR INDIVIDUAL IDIOMS BY GROUP**

<table>
<thead>
<tr>
<th>% OF GROUP</th>
<th>MJ</th>
<th>MI</th>
<th>LD</th>
<th>SP</th>
</tr>
</thead>
<tbody>
<tr>
<td>≥90%</td>
<td>got on to</td>
<td>took off</td>
<td>dropped in</td>
<td>dropped in</td>
</tr>
<tr>
<td></td>
<td>dropped in</td>
<td>gave a hand</td>
<td>gave a hand</td>
<td></td>
</tr>
<tr>
<td></td>
<td>went round</td>
<td>went round</td>
<td>got on to</td>
<td></td>
</tr>
<tr>
<td></td>
<td>gave a hand</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥80%</td>
<td>took off</td>
<td>kept an eye on</td>
<td>under lock &amp; key</td>
<td>got on to</td>
</tr>
<tr>
<td></td>
<td>under lock &amp; key</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>moved up</td>
<td></td>
<td></td>
<td>went round</td>
</tr>
<tr>
<td></td>
<td>turned over</td>
<td></td>
<td></td>
<td>took off</td>
</tr>
<tr>
<td></td>
<td>kept an eye on</td>
<td></td>
<td></td>
<td>kept an eye on</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥70%</td>
<td>pick up</td>
<td>under lock &amp; key</td>
<td>took off</td>
<td>gave a hand</td>
</tr>
<tr>
<td></td>
<td>moved up</td>
<td></td>
<td>on the spot</td>
<td></td>
</tr>
<tr>
<td></td>
<td>turned over</td>
<td></td>
<td>kept an eye on</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>moved up</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>pick up</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥60%</td>
<td>ran over</td>
<td>ran over</td>
<td>turned over</td>
<td>on the spot</td>
</tr>
<tr>
<td></td>
<td></td>
<td>turned over</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ran over</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥50%</td>
<td>on the spot</td>
<td>on the spot</td>
<td>turned over</td>
<td>turned over</td>
</tr>
<tr>
<td></td>
<td>pick up</td>
<td></td>
<td>ran over</td>
<td>pick up</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ran over</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥40%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥30%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥20%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥10%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;10%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

192
TABLE 55: DEFINITION TASK – LEVEL OF "APPROPRIACY" FOR INDIVIDUAL IDIOMS BY GROUP

<table>
<thead>
<tr>
<th>% OF GROUP</th>
<th>MJ</th>
<th>MI</th>
<th>LD</th>
<th>SP</th>
</tr>
</thead>
<tbody>
<tr>
<td>290%</td>
<td>gave a hand</td>
<td>gave a hand</td>
<td>kept an eye on</td>
<td>under lock &amp; key</td>
</tr>
<tr>
<td></td>
<td>pick up</td>
<td>kept an eye on</td>
<td>took off</td>
<td>under lock &amp; key</td>
</tr>
<tr>
<td></td>
<td>kept an eye on</td>
<td>dropped in</td>
<td>dropped in</td>
<td>under lock &amp; key</td>
</tr>
<tr>
<td>280%</td>
<td>gave a hand</td>
<td>under lock &amp; key</td>
<td>under lock &amp; key</td>
<td></td>
</tr>
<tr>
<td>270%</td>
<td>pick up</td>
<td>kept an eye on</td>
<td>kept an eye on</td>
<td>dropped in</td>
</tr>
<tr>
<td>260%</td>
<td>ran over</td>
<td>ran over</td>
<td>ran over</td>
<td>took off</td>
</tr>
<tr>
<td></td>
<td>moved up</td>
<td>dropped in</td>
<td>gave a hand</td>
<td></td>
</tr>
<tr>
<td>250%</td>
<td>got on to</td>
<td>took off</td>
<td>ran over</td>
<td></td>
</tr>
<tr>
<td>240%</td>
<td>went round</td>
<td>moved up</td>
<td>moved up</td>
<td></td>
</tr>
<tr>
<td>230%</td>
<td>got on to</td>
<td>moved up</td>
<td>pick up</td>
<td>on the spot</td>
</tr>
<tr>
<td>220%</td>
<td>turned over</td>
<td>pick up</td>
<td>turned over</td>
<td>got on to</td>
</tr>
<tr>
<td>210%</td>
<td>went round</td>
<td>on the spot</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;10%</td>
<td>went round</td>
<td>went round</td>
<td>went round</td>
<td></td>
</tr>
</tbody>
</table>
While the ranking of idioms on the play task and definition task, respectively, was similar across groups, the absolute number of "appropriate" scores obtained on individual idioms differed across groups. The proportion of subjects within each group who scored "appropriate" for each idiom is presented in Tables 56 and 57 below.

<table>
<thead>
<tr>
<th>IDIOM</th>
<th>MJ</th>
<th>MI</th>
<th>LD</th>
<th>SP</th>
</tr>
</thead>
<tbody>
<tr>
<td>RAN OVER</td>
<td>53.3%</td>
<td>53.3%</td>
<td>40%</td>
<td>34.6%</td>
</tr>
<tr>
<td>TOOK OFF</td>
<td>86.7%</td>
<td>100%</td>
<td>66.7%</td>
<td>73.1%</td>
</tr>
<tr>
<td>GAVE A HAND</td>
<td>93.3%</td>
<td>100%</td>
<td>93.3%</td>
<td>69.2%</td>
</tr>
<tr>
<td>MOVED UP</td>
<td>80%</td>
<td>60%</td>
<td>60%</td>
<td>26.9%</td>
</tr>
<tr>
<td>GOT ON TO</td>
<td>100%</td>
<td>93.3%</td>
<td>93.3%</td>
<td>84.6%</td>
</tr>
<tr>
<td>ON THE SPOT</td>
<td>46.7%</td>
<td>46.7%</td>
<td>66.7%</td>
<td>46.2%</td>
</tr>
<tr>
<td>TURNED OVER</td>
<td>80%</td>
<td>53.3%</td>
<td>46.7%</td>
<td>38.5%</td>
</tr>
<tr>
<td>UNDER LOCK &amp; KEY</td>
<td>86.7%</td>
<td>66.7%</td>
<td>86.7%</td>
<td>73.1%</td>
</tr>
<tr>
<td>PICK UP</td>
<td>66.7%</td>
<td>46.7%</td>
<td>60%</td>
<td>38.5%</td>
</tr>
<tr>
<td>KEPT AN EYE ON</td>
<td>86.7%</td>
<td>93.3%</td>
<td>86.7%</td>
<td>73.1%</td>
</tr>
<tr>
<td>DROPPED IN</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>96.2%</td>
</tr>
<tr>
<td>WENT ROUND</td>
<td>100%</td>
<td>100%</td>
<td>93.3%</td>
<td>76.9%</td>
</tr>
</tbody>
</table>

194
TABLE 57: DEFINITION TASK – PROPORTION OF GROUPS SCORING "APPROPRIATE" ON EACH IDIOM

<table>
<thead>
<tr>
<th>IDIOM</th>
<th>MJ</th>
<th>MI</th>
<th>LD</th>
<th>SP</th>
</tr>
</thead>
<tbody>
<tr>
<td>RAN OVER</td>
<td>66.7%</td>
<td>60%</td>
<td>66.7%</td>
<td>57.7%</td>
</tr>
<tr>
<td>TOOK OFF</td>
<td>60%</td>
<td>93.3%</td>
<td>40%</td>
<td>65.4%</td>
</tr>
<tr>
<td>GAVE A HAND</td>
<td>100%</td>
<td>100%</td>
<td>80%</td>
<td>61.5%</td>
</tr>
<tr>
<td>MOVED UP</td>
<td>66.7%</td>
<td>40%</td>
<td>33.3%</td>
<td>42.3%</td>
</tr>
<tr>
<td>GOT ON TO</td>
<td>60%</td>
<td>40%</td>
<td>33.3%</td>
<td>23.1%</td>
</tr>
<tr>
<td>ON THE SPOT</td>
<td>53.3%</td>
<td>20%</td>
<td>13.3%</td>
<td>30.8%</td>
</tr>
<tr>
<td>TURNED OVER</td>
<td>66.7%</td>
<td>26.7%</td>
<td>40%</td>
<td>26.9%</td>
</tr>
<tr>
<td>UNDER LOCK &amp; KEY</td>
<td>93.3%</td>
<td>93.3%</td>
<td>80%</td>
<td>88.5%</td>
</tr>
<tr>
<td>PICK UP</td>
<td>100%</td>
<td>73.3%</td>
<td>26.7%</td>
<td>34.6%</td>
</tr>
<tr>
<td>KEPT AN EYE ON</td>
<td>100%</td>
<td>100%</td>
<td>73.3%</td>
<td>73.1%</td>
</tr>
<tr>
<td>DROPPED IN</td>
<td>100%</td>
<td>93.3%</td>
<td>66.7%</td>
<td>72%</td>
</tr>
<tr>
<td>WENT ROUND</td>
<td>40%</td>
<td>13.3%</td>
<td>6.7%</td>
<td>3.9%</td>
</tr>
</tbody>
</table>
By adding up the number of children in each group who scored "appropriate" for each individual idiom, it was possible to rank the idioms from those best understood to those worst understood by each group. In order to test the concordance of the four groups' idiom rankings, Kendall's Concordance Coefficient was calculated for each task. The results are presented in Table 58 below.

**TABLE 58: WITHIN-TASK CONCORDANCE OF GROUP PERFORMANCES ON IDIOMS RANKED IN TERMS OF "APPROPRIACY" LEVELS (Kendall's Concordance Coefficient)**

<table>
<thead>
<tr>
<th>Play Task</th>
<th>Definition Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W = 0.456$</td>
<td>$W = 0.591$</td>
</tr>
<tr>
<td>$p = .0009$</td>
<td>$p = .0001$</td>
</tr>
</tbody>
</table>

The significant concordance coefficient found for both tasks indicates that the relationship between the individual idioms on each individual task was very similar across all four groups. In other words, each idiom represented a very similar level of difficulty relative to the other idioms for each of the four groups.

In order to assess the concordance of idiom ranks across the two tasks, Spearman's Rank Correlation Coefficient was computed for each group across the two tasks. The results are presented in Table 59 below.

**TABLE 59: CORRELATION OF IDIOM RANKINGS ACROSS TASKS (BY GROUP) (Spearman's Rank Correlation)**

<table>
<thead>
<tr>
<th>MJ</th>
<th>MI</th>
<th>LD</th>
<th>SP</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho = 0.044$</td>
<td>$\rho = 0.396$</td>
<td>$\rho = 0.148$</td>
<td>$\rho = 0.12$</td>
</tr>
<tr>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

The results indicate that there is no correlation between each group's individual idiom rankings on the play task and on the definition task (cf within-task ranking above).
In the course of considering the results presented above, it became necessary to complete several additional analyses. The aims and results of these miscellaneous analyses are presented below.

(a) BETWEEN-GROUPS DIFFERENCES EXCLUDING DATA FOR "ON THE SPOT"

When the results relating to individual idioms were examined, it was evident that the idiom, "on the spot", had attracted considerably more "literal" and "inferred literal" scores than had any other idiom (see between-idioms discussion in Chapter 6). One possible explanation of this result was that the presence of a small dot on the play set might have misled some children into producing "literal" or "inferred literal" demonstrations even though they actually knew the target idiomatic meaning. In order to assess the impact of the idiom "on the spot" on the outcome of the study, between-groups differences were calculated for the "inappropriate", "literal" and "inferred literal" codes excluding the data for this idiom. The results are presented in Tables 60 and 61 below.
On this limited data set, the SP group scored significantly more "inappropriate" items than did the MJ, MI and LD groups, respectively. The SP group also scored significantly more "literal" items than did the MJ group. The difference in "literal" scores was not, however, significant between the SP group and the MI and LD groups, respectively. The SP group did not differ significantly from any of these groups on the code, "inferred literal".

The LD group did not differ significantly from the MJ or MI group on any of these codes.

The MJ and MI groups did not differ significantly on any of these codes.
On this limited data set, the SP group was found to have scored significantly more "inappropriate" and "inferred literal" items than did the MJ group. The number of "literal" items scored by the SP and MJ groups did not differ significantly. The SP group did not differ significantly from the MI or LD groups on any of these codes.

The LD group scored significantly more "inappropriate" items than did the MJ and MI groups, respectively. The number of LD "literal" and "inferred literal" items did not, however, differ significantly from those scored by the MJ or MI group.

The MI group scored significantly more "inappropriate" items than did the MJ group. The two groups did not differ significantly on "literal" or "inferred literal" scores, however.
As discussed in Chapter 6 below, there is a possibility that some children may have scored "appropriate" on the play task for certain idioms merely by demonstrating the words adjacent to the idiom. If operating, such a strategy would have yielded "appropriate" scores for the susceptible idioms on the play task, but would have given rise to "repeat", "narrative" or "repeat-narrative" scores on the definition task. It was thus desirable to examine the relationship between "appropriate" play task scores and "repeat", "narrative" or "repeat-narrative" scores on the definition task for the susceptible idioms. These results are presented in Table 62, below. (The codes, "repeat", "narrative" and repeat-narrative", are abbreviated on the table as "R/N/R-N").

<table>
<thead>
<tr>
<th>IDiom</th>
<th>Total Number of Items &quot;Appropriate&quot; on Play Task</th>
<th>Number of &quot;Appropriate&quot; Play Task Items Not Coded &quot;R/N/R-N&quot; on Definition Task</th>
<th>Number of &quot;Appropriate&quot; Play Task Items Coded &quot;R/N/R-N&quot; on Definition Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOOK OFF</td>
<td>18</td>
<td>17</td>
<td>1</td>
</tr>
<tr>
<td>ON THE SPOT</td>
<td>12</td>
<td>11</td>
<td>1</td>
</tr>
<tr>
<td>DROPPED IN</td>
<td>25</td>
<td>22</td>
<td>3</td>
</tr>
<tr>
<td>WENT ROUND</td>
<td>20</td>
<td>14</td>
<td>6</td>
</tr>
</tbody>
</table>

These results indicate that the majority of children who scored "appropriate" for these idioms on the play task did not score "repeat", "narrative" or "repeat-narrative" for the corresponding item on the definition task.
In order to explore the relationship between the Checklist language features (Section G, see Appendix C) and SP play task performance, a series of Point Biserial Correlations were calculated. Table 63 below presents the results for those features significantly correlated with at least one of the codes considered.

The features included in Table 63 are abbreviated as follows:

- G3 "abnormal prosody (eg. inappropriate stress, rhythm, intonation or speed of speech)"
- G6 "overly formal"
- G7 "dominates conversations"
- G8 "very passive in conversations"
- G11 "asks questions repetitively"
- G12 "does not ask for clarification when not understanding"
- G17 "pays attention to irrelevant, non-salient details"
- G19 "tends to end conversations/interactions very abruptly"
- G21 "has rigid concept boundaries"
- G26 "uses full sentence forms rather than ellipsis (eg. Adult: 'Have you brought your book?' Child: 'I have brought my book', rather than just 'Yes')"
- G28 "uses pronouns (eg. "them" and "it") without making it clear who's being referred to"
- G29 "has poor non-verbal communication (eg. doesn't use eyes, gesture or facial expression communicatively)"
- G33 "does not understand indirect speech acts (eg. Adult: 'Do you have the time?' Child: 'Yes' rather than telling time)"
<table>
<thead>
<tr>
<th>FEATURE</th>
<th>APPROPRIATE</th>
<th>INAPPROPRIATE</th>
<th>LITERAL</th>
<th>FUZZY</th>
<th>AMBIGUOUS</th>
</tr>
</thead>
<tbody>
<tr>
<td>G3</td>
<td>$r_{pb} = -0.41$</td>
<td>$r_{pb} = 0.289$</td>
<td>$r_{pb} = 0.372$</td>
<td>$r_{pb} = 0.237$</td>
<td>$r_{pb} = 0.282$</td>
</tr>
<tr>
<td></td>
<td>$p = .0187$</td>
<td>n.s.</td>
<td>$p = .0305$</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td>G6</td>
<td>$r_{pb} = -0.011$</td>
<td>$r_{pb} = -0.211$</td>
<td>$r_{pb} = -0.325$</td>
<td>$r_{pb} = -0.111$</td>
<td>$r_{pb} = 0.342$</td>
</tr>
<tr>
<td></td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
<td>p = .0437</td>
</tr>
<tr>
<td>G7</td>
<td>$r_{pb} = -0.113$</td>
<td>$r_{pb} = 0.231$</td>
<td>$r_{pb} = -0.241$</td>
<td>$r_{pb} = 0.376$</td>
<td>$r_{pb} = -0.154$</td>
</tr>
<tr>
<td></td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
<td>p = .0290</td>
<td>n.s.</td>
</tr>
<tr>
<td>G8</td>
<td>$r_{pb} = -0.358$</td>
<td>$r_{pb} = 0.193$</td>
<td>$r_{pb} = 0.401$</td>
<td>$r_{pb} = -0.037$</td>
<td>$r_{pb} = 0.336$</td>
</tr>
<tr>
<td></td>
<td>p = .0563</td>
<td>n.s.</td>
<td>$p = .0211$</td>
<td>n.s.</td>
<td>p = .0469</td>
</tr>
<tr>
<td>G11</td>
<td>$r_{pb} = -0.378$</td>
<td>$r_{pb} = 0.386$</td>
<td>$r_{pb} = -0.119$</td>
<td>$r_{pb} = 0.51$</td>
<td>$r_{pb} = 0.076$</td>
</tr>
<tr>
<td></td>
<td>p = .0284</td>
<td>p = .0256</td>
<td>n.s.</td>
<td>p = .0039</td>
<td>n.s.</td>
</tr>
<tr>
<td>G12</td>
<td>$r_{pb} = -0.297$</td>
<td>$r_{pb} = 0.457$</td>
<td>$r_{pb} = 0.355$</td>
<td>$r_{pb} = 0.379$</td>
<td>$r_{pb} = -0.175$</td>
</tr>
<tr>
<td></td>
<td>n.s.</td>
<td>$p = .0095$</td>
<td>$p = .0374$</td>
<td>p = .0280</td>
<td>n.s.</td>
</tr>
<tr>
<td>G17</td>
<td>$r_{pb} = -0.355$</td>
<td>$r_{pb} = 0.16$</td>
<td>$r_{pb} = 0.021$</td>
<td>$r_{pb} = 0.165$</td>
<td>$r_{pb} = 0.382$</td>
</tr>
<tr>
<td></td>
<td>p = .0375</td>
<td>n.s.</td>
<td>n.s.</td>
<td>p = .0272</td>
<td>n.s.</td>
</tr>
<tr>
<td>G19</td>
<td>$r_{pb} = -0.446$</td>
<td>$r_{pb} = 0.495$</td>
<td>$r_{pb} = 0.152$</td>
<td>$r_{pb} = 0.494$</td>
<td>$r_{pb} = 0.029$</td>
</tr>
<tr>
<td></td>
<td>p = .0112</td>
<td>p = .0051</td>
<td>n.s.</td>
<td>p = .0052</td>
<td>n.s.</td>
</tr>
<tr>
<td>G21</td>
<td>$r_{pb} = -0.553$</td>
<td>$r_{pb} = 0.536$</td>
<td>$r_{pb} = 0.262$</td>
<td>$r_{pb} = 0.494$</td>
<td>$r_{pb} = 0.155$</td>
</tr>
<tr>
<td></td>
<td>p = .0017</td>
<td>p = .0024</td>
<td>n.s.</td>
<td>p = .0052</td>
<td>n.s.</td>
</tr>
<tr>
<td>G26</td>
<td>$r_{pb} = -0.378$</td>
<td>$r_{pb} = 0.3$</td>
<td>$r_{pb} = -0.119$</td>
<td>$r_{pb} = 0.403$</td>
<td>$r_{pb} = 0.207$</td>
</tr>
<tr>
<td></td>
<td>p = .0284</td>
<td>n.s.</td>
<td>n.s.</td>
<td>p = .0207</td>
<td>n.s.</td>
</tr>
<tr>
<td>G28</td>
<td>$r_{pb} = -0.481$</td>
<td>$r_{pb} = 0.617$</td>
<td>$r_{pb} = 0.273$</td>
<td>$r_{pb} = 0.52$</td>
<td>$r_{pb} = -0.095$</td>
</tr>
<tr>
<td></td>
<td>p = .0064</td>
<td>p = .0004</td>
<td>n.s.</td>
<td>p = .0032</td>
<td>n.s.</td>
</tr>
<tr>
<td>G29</td>
<td>$r_{pb} = -0.375$</td>
<td>$r_{pb} = 0.382$</td>
<td>$r_{pb} = 0.501$</td>
<td>$r_{pb} = 0.146$</td>
<td>$r_{pb} = 0.076$</td>
</tr>
<tr>
<td></td>
<td>p = .0297</td>
<td>p = .0271</td>
<td>p = .0045</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td>G33</td>
<td>$r_{pb} = -0.478$</td>
<td>$r_{pb} = 0.47$</td>
<td>$r_{pb} = 0.233$</td>
<td>$r_{pb} = 0.476$</td>
<td>$r_{pb} = 0.124$</td>
</tr>
<tr>
<td></td>
<td>p = .0068</td>
<td>p = .0077</td>
<td>n.s.</td>
<td>p = .0070</td>
<td>n.s.</td>
</tr>
</tbody>
</table>
As evidenced by these results, significant negative correlations were found between SP "appropriate" play task scores and features G3, G8, G11, G17, G19, G21, G26, G28, G29 and G33 on the Checklist. Significant positive correlations emerged between SP "inappropriate" play task scores and Checklist features G11, G12, G19, G21, G28, G29 and G33. Further significant positive correlations were found between SP "literal" play task scores and Checklist features G3, G8, G12 and G29. SP "fuzzy" play task scores correlated significantly with Checklist features G7, G11, G12, G19, G21, G26, G28 and G33. SP "ambiguous" play task items correlated significantly with Checklist features G6, G8 and G17.

(d) RELATIONSHIP BETWEEN NUMBER OF SP TRIALS AND SP PLAY TASK PERFORMANCE

In order to investigate the relationship between play task performance and the number of trials completed by each SP child at the start of the play task (see Appendix F), Spearman Rank Correlations were calculated. The results are presented in Table 64 below.

<table>
<thead>
<tr>
<th>APPROPRIATE</th>
<th>INAPPROPRIATE</th>
</tr>
</thead>
<tbody>
<tr>
<td>rho = -0.403</td>
<td>rho = 0.587</td>
</tr>
<tr>
<td>p = 0.0207</td>
<td>p = 0.0008</td>
</tr>
</tbody>
</table>

The results indicate that there was a significant negative correlation between the number of SP trials and the number of SP "appropriate" scores on the play task. There was also a significant positive correlation between the number of SP trials and the number of SP "inappropriate" items.
CHAPTER 6: DISCUSSION

This chapter summarises the results of the study and considers them with reference to previous literature and to the hypotheses outlined in Chapters 3 and 4. It begins by examining the results with regard to the relationships between the four subject groups. It then seeks to account for the considerable level of competence found among SP subjects on the play task. Thereafter, possible play task effects on the SP group's performance are discussed. Sources of breakdown at each stage of idiom acquisition and processing are then explored, before closer attention is paid to within-group results as a means of identifying more specific patterns of underlying difficulty. Finally, between-idioms and between-tasks results are summarised and discussed.

1. IDIOM COMPREHENSION IN THE MAINSTREAM GROUPS

The mainstream children's results on the play task demonstrate a non-significant trend that is consistent with expectations. The older MJ children scored a greater number of "appropriate" items and, more specifically, "idiomatic" items than did the MI group (MJ: 81.7% "appropriate", 72.8% "idiomatic"; MI: 76.1% "appropriate", 63.9% "idiomatic"). The direction of this difference is that predicted by Hypothesis (i), but the discrepancy between the two groups fails to reach significance.

The developmental trend evidenced by these results is consistent with many previous studies of idiom comprehension (Ackerman, 1982; Brasseur & Jiminez, 1989; Brinton et al., 1985; Cacciari & Levorato, 1989; Douglas & Peel, 1979; Gibbs, 1987; Johnson, 1985; Lodge & Leach, 1975; Levorato & Cacciari, 1992; Nippold & Martin, 1989; Nippold & Rudzinski, 1993; Nippold & Taylor, 1995;

1 Hypothesis (i): "The MJ group will score more "appropriate" items and "idiomatic" items on both tasks than will the MI group."
Prinz, 1983; Strand & Fraser, 1979). This literature does, however, lead one to expect a more dramatic increase in idiom comprehension between the ages of 6 and 11 than was actually found on the play task.

One explanation for the smaller-than-expected performance discrepancy relates to the nature of the idioms investigated. The idioms included in this study are notable, firstly, for having been drawn from children's television and teacher language samples. In contrast, previous studies have all drawn their idioms from dictionaries or from other idiom comprehension studies. Their idiom selection has been based almost exclusively on linguistic patterns or on adults' judgements of idiom frequency. Given that adult and adolescent ratings of frequency have been found to differ significantly (Nippold & Rudzinski, 1993), it is possible that many of the "familiar" idioms used in previous child studies were not in fact all that familiar to the subjects. Differences in perceived frequency among children and adults might reflect adults' use of a limited set of idioms in interaction with children or rapid historical changes in idiom usage. The fact that the idioms in this study were drawn from current samples of classroom and children's television language allows us to be more confident that these idioms are in current usage and are used in contexts to which children of this age range have regular exposure.

The second important difference between these idioms and those used in earlier studies concerns the range of structures included. With the sole exception of Strand & Fraser (1979), no previous studies known to this author have examined comprehension of phrasal verb idioms such as "pick up" and "ran over". (As observed in Chapter 2, several studies have included phrasal verb idioms apparently inadvertently, but these have not been a focus of analysis.) Strand and Fraser (1979) found nothing inherent in the linguistic structure of their idioms to explain why one should be learned before another. It is possible, however, that had they examined children younger than five years of age, they
might have found evidence of phrasal verb idiom comprehension preceding that of other idiom structures. Indeed, samples of child language in the literature reveal little use of clause idioms (eg. "kick the bucket") and adjectival/adverbial phrase idioms (eg. "on the spot") among young children. Very early use of phrasal verb idioms can, however, be observed:

(a) "How does this turn on?" (examining a box with a buzzer)  
Tom (2;6) (Garvey, 1984, 177)

(b) Girl: (Your mommy's) coming  
Boy: No, she's not going down\textsuperscript{1} to pick me up. Joan is.  
Boy (3;0) (Garvey, 1991, 73)

(c) "I'll take care of the fire, okay?"  
Boy (3;6) (Garvey, 1984, 95)

(d) "Naughty boys scream, don't they, and wake their own babies up."
Donna (3;10) (Tizard & Hughes, 1984, 163)

(e) "How long before our friends come back?"  
Melissa (4;4) (Garvey, 1984, 38)

While further research is required to explore the order of idiom acquisition with regard to internal structure, it may be that the inclusion of phrasal verb idioms in this study partially accounts for the MI subjects performing as well as they did. (The relationship between idioms in this study is discussed below.)

It is further possible that the inclusion only of concrete idioms (demonstrable with props) served to reduce the difficulty of the play (and definition) task relative to those of previous studies which include abstract idioms. It may well be that concrete idioms are more easily understood by young children than are abstract idioms.

The non-significant discrepancy between the MI and MJ groups may also mask a difference in depth of comprehension. Following

\textsuperscript{1}This might be an immature version of "coming down".
Levorato (1993) and Wittgenstein (1958, 1970 cited in Johnson, 1985), we might expect that children (and adults) build up layers of semantic and pragmatic knowledge about an idiom as it is encountered in an increasing number of contexts. Thus, while the MI children were able to demonstrate knowledge of these idioms on the play task with a considerable degree of accuracy, their understanding might well have been shallower than that of the older MJ children. Some support for this theory derives from comparisons of MI and MJ definitions. Consider, for example, the following pair of definitions for the idiom "kept an eye on":

MI4 (7;0): "watched him all day"
MJ11 (11;6): "not watching him all the time but usually going back and checking on him every / ten twenty minutes or so"

The nature of the play task was clearly revealed to facilitate the demonstration of idiom comprehension. Whereas similar numbers of "appropriate" and "inappropriate" items were scored by the MI and MJ groups on the play task, on the definition task the MI group scored significantly fewer "appropriate" items than did the MJ group (p = 0.0253), and more than twice as many "inappropriate" items (p = 0.0055). Thus, as hypothesized\(^1\), MI idiom comprehension is indeed found to be significantly poorer than that of the older MJ children when measured on a more conventional methodology making expressive and metalinguistic demands. (Between-tasks differences are discussed fully below.)

While the depth of MI knowledge regarding these idioms may have been shallower than that of the MJ children, their mean of 9.13 (out of 12) idioms scored "appropriate" on the play task is consistent with the hypothesis that 6- and 7-year old children already have a considerable grasp of idiomaticity (Johnson, 1985; Levorato, 1993; Strand & Fraser, 1979). Such comments as that

\(^1\) Hypothesis (i): The MJ group will score more "appropriate" items and "idiomatic" items on both tasks than will the MI group.
Hypothesis (ii): The MJ group will score fewer "inappropriate" items on the definition task than will the MI group.
quoted below clearly indicate an awareness of both literality and idiomaticity. Consistent with Levorato's (1993) model of figurative language development, this child (aged 6;11) demonstrates a capacity to suspend literal interpretation in an idiomatic context.

A: Sally kept an eye on him all day. What do you think "kept an eye on" means?
MI15: it means look look after him
A: uhu
MI15: once my mum told me to look keep an eye on Vicky and I said "I can't my eye doesn't come out" (in a very animated voice)¹

Non-significant differences between the MI and MJ groups on "literal" and "inferred literal" scores suggest further that any remaining preference for literal interpretation (Strand & Fraser, 1979) among 6- and 7-year olds may not apply to very common idioms. (The MI group scored 3.9% of their play task items "literal" as opposed to 3.3% in the MJ group (n.s). Both mainstream groups scored "inferred literal" for 1.1% of their play task items.) Low scores on the codes "ambiguous" and "fuzzy" provide further support for the 6- and 7-year olds having adequate pragmatic skill to select the idiomatic meaning over the literal meaning in context. Were these skills lacking, four options would have been open to the MI children. Firstly, they may have opted to demonstrate both the literal and idiomatic meanings, thereby scoring "ambiguous-idiomatic/literal". Secondly, they could have chosen to act out the idiom in a manner that did not relate to either the literal or the idiomatic sense, yielding a "fuzzy" score. Thirdly, they might have chosen one of the two options at random. This would have given a child an equal chance of scoring "idiomatic" or "literal". Finally, the child might have fallen back on a literal strategy, the response to ambiguity previously reported for young children (Johnson, 1985). Since the MI-MJ differences on the codes "ambiguous-idiomatic/literal", "fuzzy", "literal" and "inferred literal"

¹It is unclear whether this anecdote is intended to comment on her own previous lack of idiomatic understanding, or whether it is a repeat of what was originally intended as a joke.

208
were all non-significant, we may conclude that the MI children had a considerable grasp of the semantic and pragmatic knowledge required to interpret these idioms in context.

While the MI group appeared, collectively, to have negotiated the transition to idiomatic comprehension, there was nevertheless some degree of within-group variance on the broad and narrow codes (see Tables 4, 6 and 8 in Chapter 5). The variance was not accounted for by age or syntactic comprehension (no correlations were found between age and TROG scores, respectively, and play task performance). It is likely that differential rates of exposure, learning and cognitive development all contributed to the within-group variation. Differences in early (and later) idiom acquisition may also indirectly reflect differences in socio-economic status. Given the established link between idiom comprehension and literacy in normally-developing children (Nippold, 1988 cited in Ezell & Goldstein, 1991), it would be interesting for future research to explore this question further.

2. IDIOM COMPREHENSION IN THE LD GROUP AS COMPARED WITH THE MAINSTREAM GROUPS

The results of the play task reveal that the LD group's idiom comprehension is less well-developed than that of the MJ group, but is not significantly poorer than that of the MI group.

The difference between the LD and MJ groups on the broad code, "appropriate", fails to attain significance, but only by a narrow margin (LD: 74.4%, MJ: 81.7%, p = 0.0606). Once the "inferred idiomatic" and "appropriate no action" data is excluded, there emerges a clearly significant difference between the two groups on the code "idiomatic" (LD: 61.1%, MJ: 72.8%, p = 0.0186).

The failure of the LD group to score as many "idiomatic" items
as did the MJ group is consistent with Hypothesis (iii). The expectation of poorer LD performance was based on the anticipated impact of general comprehension difficulties in many LD children, and on the lower average age of the LD than MJ children (MJ group ranged from 10;6-11;6 with a mean age of 10;9, LD group ranged from 8;6 - 11;2, mean age 9;11). Importantly, Hypothesis (iii) did not derive from the expectation that idiom comprehension would present more difficulty than literal comprehension for the LD group.

The results confirm that the LD group was not more literal than was the MJ group. On the contrary, the LD group scored fewer "literal" and "inferred literal" items on the play task than did the MJ group (although the absolute numbers were very small and the difference was not significant). Only 1.7% of LD items scored "literal" (mean 0.2, s.d. 0.41) and 0.6% scored "inferred literal" (mean 0.07, s.d. 0.26).

Overall, approximately three quarters (74.4%) of the LD items achieved an "appropriate" code (mean 8.93 out of a possible 12, s.d. 1.39). (This corresponds to 61.1% "idiomatic", 11.7% "inferred idiomatic" and 1.7% "appropriate no action".) These results clearly demonstrate that the LD subjects had achieved a considerable grasp of idiomaticity. Nevertheless, their play task performance still evidenced weakness in comparison with the normally-developing children. The LD children were an average of two years and ten months older than the MI children. The period that separates the eldest MI and youngest LD children (7;1-8;6) has been associated with a spurt in idiom comprehension among normally-developing children (Johnson, 1985; Strand & Fraser, 1979). Despite this age advantage, the LD group scored approximately the same number of "idiomatic", "inferred idiomatic" and "appropriate no action" items as did the MI group (LD: 61.1%, 11.7% and 1.7%, respectively; MI: 63.9%, 11.1% and 1.1%, respectively; all non-significant differences).

---

1 Hypothesis (iii): The MJ group will score more "appropriate" items and "idiomatic" items on both tasks than will the LD group.
While producing fewer "appropriate" actions, the LD group was not found to have scored significantly more "literal" or "inferred literal" scores than the MI (or MJ) group. To this extent, the results are consistent with those of Johnson (1985) who found that language-disordered children's idiom comprehension lagged behind that of normally-developing children by approximately two or three years, but was consonant with their overall receptive language ability. Strand (1982 cited in Johnson, 1985) and Vance and Wells (1994) similarly reported that idiom comprehension did not appear to present more difficulty than did any other comprehension for their groups of language-disordered children.

Interestingly, the weakness of the LD group relative to the MI group is further evidenced by a significantly higher level of "fuzzy" scores among the LD subjects (p = 0.0152). (At 13 out of 180 (7.2%, mean 0.87, s.d. 0.74), the number of LD "fuzzy" items is, however, still quite small.) As discussed above, "fuzzy" scores may reflect difficulty in choosing between a literal and idiomatic meaning. The non-significant difference on the code "ambiguous-idiomatic/literal" between the LD group and the MI and MJ groups, respectively, militates against this explanation, however. Alternatively, a "fuzzy" score might also occur when a child recognises the contextual anomaly created by a literal interpretation but does not know any other meaning for the given expression. In terms of this interpretation, the difficulty would lie with a limited idiom repertoire rather than with difficulty in using context to inform the selection of idiomatic over literal interpretation. (As discussed below, more general difficulty in using context may also give rise to a limited idiom vocabulary.)

While "fuzzy" items might reflect areas of semantic and/or pragmatic weakness, the significantly higher level of LD than MI "fuzzy" items might also relate to difficulty in other skills necessary for this task. General syntactic and lexical comprehension difficulties, a lapse in concentration, short-term memory problems, failure to discern the words on the tape, being
unable to think of an appropriate portrayal of the idiom, or feeling unable to ask for clarification could all lead a child to produce a "fuzzy" demonstration of an idiom. Certainly, in at least two of the 13 LD cases of "fuzzy", the problem clearly does not stem from a difficulty with idiom comprehension per se. The following "fuzzy" actions were both followed a few minutes later by clearly "idiomatic" definitions.

Example 1 (LD7, 10;5):

"The policeman put him in a cell and kept him under lock and key"

# lowering the bike rider into the cell with one hand and holding the policeman in the other
* pushes the cell door closed with the hand that holds the policeman, stands the policeman in the police station, touches the lock and key without lifting them, fetches Sally from the road and stands her outside the police station, HA

Example 2 (LD5, 8;7):

"Sally kept an eye on him all day"

# Rover lies in the basket in the house; car parked inside the house; Sally lies in the garden; removing Mum from the car
* removes Mum from the car, reverses the car out of the house and parks it as at the start, stands Mum against the side wall of the house, HA

In Example 1, it appears that LD7 registered the need to do something with the lock and key. Given that he later demonstrated his familiarity with the idiom, it is possible that he simply could not hold both clauses of this long sentence in memory at once. It is further possible that this child failed

---

1 The symbol # denotes the position of the props at the moment when the idiom is heard. The symbol * denotes the child's action from the time the idiom is heard. "HA" ("hands away") denotes the child moving his or her hands away from the play set.

2 There is no evidence of other LD children having had difficulty processing this sentence. It is possible, however, that others may have compensated for any difficulty by making use of their world knowledge about locks and keys.
to discern the words of the idiom during the play task but succeeded in doing so when he heard it again during the definition task.

In the case of LD5 (Example 2), at least two possible explanations may be furnished for the child's "fuzzy" action despite subsequently demonstrating knowledge of the idiom. With the car occupying the entire floor space in the house, the child was unable to enact the sentence. It may be that, in the course of moving the car, the child became distracted or forgot the sentence and felt reluctant to request its repetition. Alternatively, this child may not have been able to think of a way of demonstrating "kept an eye on" and may therefore have chosen to manipulate other props in order to satisfy the researcher.

As illustrated by these examples, some LD difficulty may have arisen from task factors unrelated to comprehension. In assessing the likely impact of this finding on the validity of the study, it is important to note that each of the 15 LD children achieved an "appropriate" score for at least seven of the twelve idioms. No individual LD child scored "fuzzy" for more than two idioms. Where "fuzzy" was coded on the play task (13 out of 180 items), the corresponding item on the definition task almost always scored an "inappropriate" code (11 out of 13).

In other words, for most of the LD "fuzzy" items, there is no evidence of underlying task-related difficulties⁴. The fact that the LD "fuzzy" items were spread across eight different idioms on the play task further supports the view that there was nothing inherent in the linguistic contextualisation of the idioms that posed a problem for this group. Accurate demonstration of all the literal language in the story indicates that the children did understand what was required of them in the task.

Given that a large proportion of definition task "inappropriate" scores correspond to "appropriate" play task scores, no firm conclusions can be drawn in this regard.

213
While there is evidence for only very limited interference from task demands on the play task, the results clearly indicate that the LD children were greatly disadvantaged by the requirements of the definition task. In contrast to the play task, the results of the definition task revealed a significantly lower level of LD than MI "appropriate" items \( (p = 0.0253) \) and a significantly higher level of "inappropriate" items in the LD group as compared with the MI group \( (p = 0.0468) \). (These discrepancies confirm Hypotheses (v) and (vi).) The discrepancy between the LD and MJ groups was, not surprisingly, even greater \( ("appropriate" \ p = 0.0002, \ "inappropriate" \ p = 0.0002) \). (These results confirm Hypotheses (iii) and (iv).)

Importantly, despite the poor performance of the LD group on the definition task (discussed in relation to between-tasks differences, below), this group was still not found to score significantly more "literal" or "inferred literal" definition items than did either the MI or MJ groups. This confirms the play task conclusion that the LD children, collectively, were not any more literal in their interpretations than were those in the mainstream groups. Although the precise degree of receptive language difficulties in these children was not assessed in this study, its findings are consistent with previous research which has found no outstanding difficulty with regard to idioms in language-disordered children (Johnson, 1985; Strand, 1982 cited in Johnson, 1985; Vance & Wells, 1994).

\[\text{Hypothesis (v): "The MI group will score more "appropriate" items on the definition task than will the LD group."} \]

\[\text{Hypothesis (vi): "The MI group will score fewer "inappropriate" items on the definition task than will the LD group."} \]

\[\text{Hypothesis (iii): "The MJ group will score more "appropriate" items and more "idiomatic" items on both tasks than will the LD group."} \]

\[\text{Hypothesis (iv): "The MJ group will score fewer "inappropriate" items on the definition task than will the LD group."} \]
3. IDIOM COMPREHENSION IN THE SP GROUP AS COMPARED WITH THE MAINSTREAM GROUPS

As hypothesized\(^1\), the SP group displayed significantly poorer performance on the play task than did either of the mainstream groups. Whereas the MJ and MI groups scored "appropriate" for 81.7% and 76.1% of their items, respectively, only 60.9% of SP items earned an "appropriate" code (SP-MJ \( p = 0.0005 \); SP-MI \( p = 0.0070 \)). The weakness in SP idiom comprehension was further reflected in a relatively high level of "inappropriate" items. In contrast to 8.3% of the MJ items and 9.4% of the MI items, as many as 25% of the SP group's items scored "inappropriate" (SP-MJ \( p = 0.0006 \); SP-MI \( p = 0.0012 \)).

This pattern of results is consistent with a large body of literature which characterises children with semantic-pragmatic difficulties as having problems in comprehending non-literal language (Bishop & Adams, 1989; Bishop & Rosenbloom, 1987; Culloden et al., 1986; Hyde-Wright & Cray, 1991; Jones, Smedley & Jennings, 1986; McTear, 1991; McTear & Conti-Ramsden, 1992; Rapin & Allen, 1987; Shields, 1991; Smedley, 1989). Interestingly, however, these results do not accord with those reported by the only previous systematic study of idiom comprehension in children with semantic-pragmatic difficulties (Vance and Wells, 1994).

As outlined in Chapter 2, Vance and Wells used a multiple choice task to compare the comprehension of a group of SLI children - which included a subgroup of SPLD (semantic-pragmatic language disorder) children - with that of normally-developing children matched on receptive language. No significant difference in the two groups' comprehension of a mixed set of idioms and metaphors (ten items in total) emerged. Furthermore, the two subgroups of SLI (SPLD and non-SPLD) were not found to differ significantly

\(^1\) Hypothesis (vii): The SP group (and subgroups) will score fewer "appropriate" items and "idiomatic" items on both tasks than will the MI or MJ group.

Hypothesis (viii): The SP group (and subgroups) will score more "inappropriate" items and "literal" items on both tasks than will the MI or MJ group.
from each other. In the absence of reported data comparing each subgroup with the normally-developing controls, it is possible only to extrapolate that the SPLD group did not differ significantly from the control group either.

It has been argued above that there are a number of features of Vance and Wells' methodology that might have led to the idiom ability of SP children being overestimated. Firstly, the subjects in Vance and Wells' study had all been pretested on a semantic anomaly task. One SLI subject who failed the semantic anomaly task was excluded. (It is not clear from the report whether this subject was classified SPLD or non-SPLD.) Since the ability to recognise semantic anomaly is essential for understanding idioms in context, exclusion on these grounds would reduce the likelihood of finding idiom comprehension impairment in the group. Having included only children capable of detecting semantic anomaly, Vance and Wells then reduced the pragmatic demands of the task by priming their subjects to select the idiomatic meaning. Together, these features may well have inflated the performance of the SPLD children on the idiom comprehension task. It is further notable that Vance and Wells' study used a small sample of SP children. Although 17 SLI subjects completed the task, only seven SPLD children and seven non-SPLD children were included in the subgroup comparison.

In their report, Vance and Wells conclude that

performance on a structured task of non-literal comprehension would not seem to be a good indicator of the difficulties that so-called SPLD children appear to have with understanding non-literal language in conversation (pp. 39-40).

While the play task data of the present study were also gathered on a structured task, the context in which they were gathered and in which the idioms were presented approximated natural conversation to a greater extent than did that of Vance and Wells. The subjects in the Vance and Wells study were alerted to the presence of idioms and instructed to focus on the
idiomatic meaning. On the play task, in contrast, the presence of idioms in the story was not pointed out, nor were the children aware of what it was that the researcher was examining. Indeed, the SP children did not appear to realise that the play task was a test at all. The nature of the contextualising narrative further increased the similarity between this task and normal conversation. Whereas Vance and Wells presented each idiom in a separate one- or two-sentence context, the idioms in this study were all embedded into a single narrative.

Despite the discrepancy in the between-groups results presented by Vance and Wells (1994) and those presented here, some degree of consistency is evident in relation to the analysis of errors. The three multiple choice options presented to Vance and Wells' subjects took the form of an idiomatic, a literal and a distractor picture, respectively. The authors report that the SLI group did not display a greater tendency to literal interpretations than did the control group. They note further that the subgroup of SPLD children did not show a greater tendency to make literal errors rather than distractor errors.

Similarly to Vance and Wells, the present study found that the SP group did not score significantly more "literal" or "inferred literal" items than did the mainstream groups. A "literal" action was recorded for 4.8% of SP items, 3.9% of MI items, and 3.3% of MJ items (differences not significant). "Inferred literal" was coded for 2.9% of SP items, 1.1% of MI items, and 1.1% of MJ items (differences not significant). Thus, contrary to Hypothesis (viii), the SP group did not emerge as being more literal than the mainstream groups.

Importantly, however, the SP group was found to have scored

---

1 Their distractor choice represented neither the literal nor the idiomatic meaning of the expression. For example, for the idiom "raining cats and dogs", the distractor picture showed cats and dogs on the ground with no rain.

2 Hypothesis (viii): "The SP group (and subgroups) will score more "inappropriate" items and "literal" items on both tasks than will the MI or MJ group."
significantly more "fuzzy" items than were recorded for either of the mainstream groups. The MJ and MI groups scored "fuzzy" for 3.3% and 2.2% of their items, respectively. The SP group scored "fuzzy" for 16% of their total number of items. The differences between the SP and MJ groups (p = 0.0002) and between the SP and MI groups (p = 0.0000) are both highly significant.

As discussed above with respect to the LD group, "fuzzy" codes may be argued to reflect immaturities in the semantic and pragmatic knowledge necessary for interpreting idioms in context. While this interpretation was based on examining the options open to a child in idiomatic transition, it also finds support in previous studies of idiom comprehension. It was noted in Chapter 2 that the tendency to select or offer a meaning which is neither the target idiomatic sense nor the literal sense has been found to be typical of children aged between 5 and 7 (Cacciari & Levorato, 1989, 1992). It has been suggested that this response to idioms is characteristic of children in transition between complete literality and a full grasp of idiomaticity. While arguably not directly comparable to these "not idiomatic but also not literal" codes, the "fuzzy" code does also encompass actions which are neither idiomatic nor literal demonstrations of the target idiom.

Interestingly, Vance and Wells' study also included a "not idiomatic nor literal" option (their "distractor" picture). As noted above, the authors found that the SPLD children did not make more literal errors than distractor errors. Unfortunately, the authors do not report on whether the number of distractor errors was greater than the number of literal ones - as was the case in the present study. It is further regrettable that Vance and Wells do not present between-groups differences on distractor choices.

While comparative literature pertaining to "fuzzy"-type responses in children with semantic-pragmatic difficulties is lacking, between-tasks comparisons provide some important clues as to what
factors underlie "fuzzy" codes in the SP group. Examination of the corresponding play and definition task items reveals that 30 of the 50 (60%) SP items scored "fuzzy" on the play task also scored an "inappropriate" code on the definition task. This pattern of results is consistent with an hypothesis that the children lacked the semantic and/or the pragmatic knowledge required to interpret these idioms in context\(^1\). In at least some cases of "fuzzy" play task items, however, it could also be argued that skills unrelated to idiom comprehension per se were implicated.

When the "fuzzy" items are closely examined, it is possible to identify at least three types of actions which attracted this code:-

(a) Actions which represented part of the sentence but failed to demonstrate the idiomatic, literal or any other figurative sense of the expression;

(b) Actions which represented a part of the ensuing narrative but failed to demonstrate the idiomatic, literal or any other figurative sense of the expression; and

(c) Actions which were unrelated to any sense of the idiomatic expression, to the rest of the sentence or to the ensuing narrative.

Of the 50 "fuzzy" items scored by the SP group on the play task, ten were of type (a), eight were of type (b) and 17 were of type (c). As evidenced in the discussion of the following examples, each of these "fuzzy" types is suggestive of a different set of interpretations. It is, however, sometimes difficult to draw firm conclusions from these actions even at an individual level.

\(^1\) Given that a large proportion of definition task "inappropriate" scores correspond to "appropriate" play task scores, no firm conclusions can be drawn in this regard.
Example of "fuzzy" action type (a) (SP23, 6;5):

In response to the sentence "Sally gave her mum a hand to carry Rover to the white car", SP23 lifts Rover into the car without involving either Sally or her mum.

One interpretation of this action is that the child failed to process the first half of the sentence or was unable to retain it in memory. The evidence does, however, militate against a memory explanation. The same child was found capable of remembering other long and more complex sentences and, furthermore, displayed no evidence of memory failure in relation to any of the literal sentences. (It might, of course, be argued that less familiar language is more difficult to retain in memory, but this in itself suggests that the idiomatic language was not familiar.)

While a general memory impairment is unlikely to account for this "fuzzy" action, this action may plausibly be explained in terms of a momentary lapse in concentration. It should be noted, however, that the researcher did repeat sentences when it appeared that a child had been momentarily distracted (see Appendix J).

Another possible interpretation of this "fuzzy" action relates to immaturity in play skills. Studies of symbolic play development indicate that the ability to employ a doll as the agent of action on another doll or toy represents a higher level of play development than does the child him/herself playing agent of action on a recipient doll or toy (Bretherton, 1984 cited in Rescorla & Goossens, 1992; Garvey, 1991; Lowe, 1975; Watson & Fischer, 1977 cited in Rescorla & Goossens, 1992). Prior to the development of doll-as-agent play, a child might plausibly have difficulty in demonstrating the idiom "gave a hand". In this case, however, the child engaged in doll-as-agent demonstrations in other instances. (The use of a play task to

---

1In normally-developing children, doll-as-agent play usually emerges between the ages of 2;6 and 3;0.
assess children with semantic-pragmatic difficulties is discussed more fully below.)

Even if acting with props did not in itself present problems to this child, the dramatic context could have influenced his play task performance. As noted above, the procedure was presented to the children as the making of a video rather than as a test. To the author, the idiom represented the most important part of any target sentence. For the child, unaware that this was a test of idiom comprehension, the idiom may not have appeared to be the most salient bit of a given sentence. Thus, on hearing the sentence, "Sally gave her mum a hand to carry Rover to the white car", the child may have considered it important only that he get the dog into the car. How this was achieved may not have seemed significant.

In view of the plausibility of this salience argument, it would be unwise to assume pathology or immaturity in a child like SP23, solely on the grounds that he took the dog to the car himself without involving either Sally or her mum. It is possible, however, that this child did have a tendency to focus on parts of a sentence rather than on the whole.

While any of these factors may plausibly have contributed to the "fuzzy" demonstration of this idiom, it remains possible that the child simply did not know the idiomatic meaning despite recognising that a literal demonstration would not make sense. Enacting another part of the sentence may reflect an attempt to hide comprehension failure, desire to please an adult stranger and not disrupt the video, or an attempt to guess the meaning of an unknown idiom.

In fact, in this particular case, lack of semantic knowledge can be ruled out. SP23 went on to produce an "idiomatic" definition

---

1 Attention to parts rather than to the whole might well underlie type (a) actions which do not represent consecutive words in the sentence. For example, in response to the same sentence, "Sally gave her mum a hand to carry Rover to the white car", several SP children depicted Mum carrying Rover to the car (without involving Sally at all).
of this expression in the second task. While clearly having the semantic knowledge to interpret this idiom, the child may nevertheless have faltered in the process of retrieval. This processing may have been facilitated on the definition task by the repetition of the idiom and the author's directing of the child's attention to it.

Consideration of this type (a) item highlights the fact that a failure to demonstrate knowledge does not necessarily imply a lack of knowledge. In actions of type (b), however, idiom comprehension difficulty does appear to be the more likely cause of the "fuzzy" action.

Example of "fuzzy" action type (b) (SP12, 8;0):

When SP12 hears the sentence, "Sally turned the bike rider over to the policeman", the bike rider is seated on the bike, the policeman stands nearby, and Sally sits in the car some distance away. The child then moves the policeman till he is standing adjacent to the bike. He then immediately moves the bike rider on the bike, with the policeman alongside it, towards the police station.

Clearly, the child has omitted the idiom sentence entirely and instead demonstrates the ensuing sentence, "The policeman took the bike rider to the police station". This had been heard once before (in the first reading of the story), but was not played during the action until after this demonstration.

In such cases as these, it is difficult to account for the "fuzzy" action in terms of memory failure, loss of concentration or salience factors. Actions such as these are, however, consistent with an idiom comprehension problem. Indeed, this theory is supported by the finding that all but one of the "fuzzy" actions of this type were followed by a "literal", "inferred literal", "narrative" or "repeat-narrative" definition. The remaining case was coded "ambiguous-idiomatic/OFS" on the definition task.
Type (c) "fuzzy" actions, like type (b), were almost always followed by an "inappropriate" definition. This lends support to the theory that these usually reflected a difficulty in interpreting the given idiom. The demonstration of an entirely unrelated action might, however, reflect a difficulty of impulse control or a lapse in concentration. Children who produced actions such as the one below often appeared confused and quite at a loss as to what they might do to represent the sentence. Actions of this type were sometimes interspersed with comments from the child which revealed an effort to make sense of the idiom or to reconstruct the sentence entirely.

Example of "fuzzy" action type (c) (SP22, 9;0):

When the sentence, "Sally turned the bike rider over to the policeman", is played, SP22 is holding Mum and the policeman in the air. The bike rider is on the bike in the garden. On hearing the sentence, the child stretches out the policeman's arms, brings the bike to the police station, lifts it and finally parks it between the police station and the vet's office. Still holding Mum and the policeman, she knocks them together and then holds them face to face in the air.

For the most part, the interpretation options outlined with regard to "fuzzy" action types (a), (b) and (c) are within-child problems (either of idiom comprehension or of other skills required for the play task). In relation to two further types of "fuzzy" action, it is also necessary to consider the inherent difficulty in coding play data.

Two further "fuzzy" action types can be identified:–

d) Actions for which some relationship to the idiomatic or literal sense could be argued but which might plausibly reflect a different idiom or merely a response to other words in the sentence;

e) Actions which took place in the context of the props being so positioned that further action was unnecessary. Had the unrelated action not occurred, these items would have scored "appropriate no action".
Represented by twelve items, actions of type (d) could be interpreted in terms of within-child idiom comprehension difficulty. They might also, however, reflect a child's intended demonstration of the idiomatic or literal meaning being obscured by an immature or idiosyncratic style of play, or by the child's haste. This is illustrated by the following action depicting the sentence "Sally gave her mum a hand to carry Rover to the white car".

Example of "fuzzy" action type (d) (SP12, 8;0):

SP12 lifts Rover between Mum's arms, motions with Mum and Rover in mid-air towards Sally who stands approximately 20cm away, and then takes Mum and Rover to the car.

The dilemma in coding this action centres on whether or not the child's motioning towards Sally was intended to represent Sally's role in carrying Rover to the car. It is only with the benefit of corroborating evidence from the definition task - the corresponding item scored "inferred idiomatic" - that it is possible to rule out an idiom comprehension problem in this case\(^1\). In other cases, however, the definition task data does not provide such clear guidance.

The difficulty in interpreting play data is highlighted further in relation to type (e) "fuzzy" actions. It was notable that this action type was associated solely with the idiom, "kept an eye on". When the sentence, "Sally kept an eye on him all day" was played, several SP children already had the props appropriately placed for an idiomatic demonstration i.e. Sally was standing close to and facing the basket in which Rover lay. Had these children not acted at all, these items would have been coded "appropriate no action". Instead, perhaps feeling obliged to do something, or being caught up in the creative task, the children performed an action unrelated to the idiom, such as taking Sally's mum into the garden or adjusting the position of

---

\(^1\)Each task was coded independently without reference to the data for the other task.
the car. The fact that this action type corresponded to a single idiom, coupled with the finding that all three instances of it were followed by "appropriate" definitions indicates that these "fuzzy" items were probably false positives).

In all, 14 SP items coded "fuzzy" on the play task were associated with "appropriate" definitions. Thus, in at least 28% of cases where SP children scored "fuzzy" on the play task, they did in fact have knowledge of the given idiom. Their failure to demonstrate this knowledge may reflect a difficulty in processing (some) idioms in context, a difficulty in another task skill such as memory, attention or the ability to discern the words on tape, or a combination of the two. The 14 items were spread across 13 different children and both subgroups. They encompassed all five "fuzzy" action types, but were associated commonly with type (e) and rarely with types (b) and (c). This is consistent with type (e) actions reflecting coding difficulties and action types (b) and (c) being more likely to reflect actual difficulty in interpreting idioms in context.

Clearly, the conclusions that can be safely drawn with respect to the "fuzzy" results are limited. At least 14 of the 50 items are probably false positives. It may be that some of the remaining "fuzzy" items also reflect difficulty in skills other than idiom comprehension per se. Importantly, however, such skills as short-term memory, attention and focusing on the whole rather than on parts of a sentence are all essential to idiom comprehension in context, even though they may not be necessary for idiom comprehension in isolation.

Given that at least some of the "fuzzy" SP items appear to be

---

1 In the context of this report, a "false positive" is an item which was not scored "appropriate" despite actually being understood by the child.

2 Three out of three type (e) "fuzzy" actions were associated with "appropriate" definitions. Three out of 24 type (b) or (c) actions corresponded to "appropriate" definitions.

3 It remains possible that idiom processing difficulty hindered the children's first attempts to retrieve the idiomatic meaning (on the play task), but that the repetition of the idiom and of the processing may then have yielded an "appropriate" definition.
artifacts of the play task itself, we may conclude that the SP group's idiom knowledge (at least at a semantic level) is in fact marginally better than it might appear from the broad code play task results. Nevertheless, it should be borne in mind that the SP group is, on average, 30 months older than the MI group. As such, any inferiority (or even equality) in the SP group's play task performance relative to that of the MI group indicates some impairment in idiom comprehension relative to the normally-developing children. (Even when all "fuzzy" items are removed from the "inappropriate" scores, the SP group still scores proportionately more "inappropriate" items than does the MI group.)

While the results of the play task indicate the presence of some idiom comprehension difficulty in the SP group as a whole, the comparison of SP "appropriate" and "inappropriate" results is testimony to their considerable competence in relation to these twelve idioms. In line with the characterisation of SP children in terms of idiom comprehension impairment, it was predicted that the SP group would score more "inappropriate" than "appropriate" items on the play task (Hypothesis (xv)). The analysis did indeed reveal a highly significant difference (p = 0.0001). Crucially, however, this difference was in the opposite direction to that predicted. The SP group achieved an "appropriate" score for 60.9% of its play task items (mean = 7.31, s.d. = 2.33). Only 25% (mean = 3, s.d. = 2.02) of their play task items were "inappropriate". (The remaining 14.1% were "ambiguous". No significant between-groups differences were found on this code.)

Overall, then, the SP group emerged on the play task as significantly less able than the MJ and MI groups to demonstrate comprehension of these twelve idioms. They were not, however, significantly more literal in their demonstrations. These results are consistent with an immature idiom vocabulary and difficulty in retrieving and selecting idiomatic meanings in

1When all "fuzzy" scores are removed, "inappropriate" still accounts for 9% of the SP group's items and 7.2% of the MI items (difference not significant).
context. Despite its relative weakness, the SP group scored significantly more items "appropriate" than "inappropriate".

4. IDIOM COMPREHENSION IN THE SP GROUP AS COMPARED WITH THE LD GROUP

As discussed in Chapter 2, the relationship between semantic-pragmatic difficulties and language disorders is a subject of continuing controversy. Resolution of the debate inevitably demands that two questions be answered: (a) can children with semantic-pragmatic difficulties be distinguished from those with autism and Asperger syndrome? and (b) are children with semantic-pragmatic difficulties different in some measurable respect from those with language disorders? Thus far, only one study has addressed the latter question from the perspective of idiom comprehension (Vance and Wells, 1994). Having tested both groups on a pictorial multiple choice task, Vance and Wells (1994) reported that no significant differences between the two groups were found with regard to idiom comprehension. The results did, however, suggest that a subgroup of children in both groups were characterised by idiom comprehension that was impaired relative to their general receptive language ability.

The results of the two tasks employed in the present study present two very different pictures of the SP-LD relationship. The findings of the definition task portray two groups not distinguished by their respective scores on any broad or narrow code. (The SP group did in fact score marginally more "appropriate" items and marginally fewer "inappropriate" ones on the definition task than did the LD group"). The differences are very small and not significant, however.) While the differences between the SP and LD groups were not significant, both groups scored significantly fewer "idiomatic" items and significantly

^1"Appropriate": SP 48.2% vs LD 46.7%, n.s. "Inappropriate": SP 42.4% vs LD 45%, n.s.
more "fuzzy" items than did either mainstream group. Additionally, the SP and LD groups both scored significantly more "narrative" items than did the MI group.

The results of the definition task are consistent with the findings of Vance and Wells (1994). Critically, however, they differ strikingly from those yielded by the arguably more valid play task.

As hypothesized⁴, the SP group emerged on the play task as significantly less "appropriate" (p = 0.0091) and significantly more "inappropriate" (p = 0.0015) than the LD group. At the narrow code level, the two groups differed significantly on "idiomatic" (p = 0.0337) and "fuzzy" (p = 0.0124) codes. Interestingly, the differences on the codes "literal" (p = 0.0546) and "inferred literal" (p = 0.0739), while not attaining significance, did come closer to doing so than did any other pairwise difference on these codes. This suggests that exploration of these differences in future research may well prove fruitful.

The results of the play task thus indicate that the LD group's ability to demonstrate comprehension of these idioms is poorer than that of the normally-developing children (as discussed above), but to a significantly lesser degree than is the case for the SP group. Indeed, the evidence from both Mann-Whitney Tests and Kth Neighbour discriminant analysis supports a distinction between SP and LD groups on the basis of idiom comprehension.

The dramatic difference in SP-LD relationship depicted by the two tasks in this study highlights the extent to which differences between these groups may be masked by an inappropriate task. It is noteworthy that the greater similarity between SP and LD groups on the definition task was due to a decline in LD results rather than an improvement in SP results. Indeed, both groups

⁴Hypothesis (ix): The SP group (and subgroups) will score fewer "appropriate" items and "idiomatic" items on the play task than will the LD group.

228
performed less well on the definition task than on the play task. The decline was, however, greater for the LD group than for the SP group (see between-tasks discussion, below).

The between-tasks discrepancy in relation to the SP and LD groups supports the view, expressed in Chapter 2, that formal tests of idiom comprehension which employ definition tasks (Wiig & Secord, 1989; Thorum, 1986) put both SP and LD children at a disadvantage. As such, the validity of these tests' findings with regard to idiom comprehension in SP and LD children is seriously threatened.

It is notable, however, that the pictorial multiple choice methodology employed by Vance and Wells (1994) did not involve verbal expression on the subjects' part. How do we explain the discrepancy between the results reported by Vance and Wells and those of the play task? In order to account for this difference, it is helpful to examine the narrow codes on which the SP and LD groups differed significantly on the play task. In addition to a significant difference in "idiomatic" scores, the two groups also differed significantly in their respective numbers of "fuzzy" items. Whereas the LD group scored "fuzzy" for 7.2% of their play task items, the SP group scored a total of 16% "fuzzy" items (p = 0.0124).

As discussed above, "fuzzy" items may reflect difficulty in retrieving idioms in context (even if the child has adequate semantic knowledge of the idiom) and/or a deficit in a related skill such as memory or attention. While it is possible that the SP group was more impaired than the LD group in task-related skills other than idiom comprehension, there is no evidence to substantiate such a theory. It is more likely that the SP-LD discrepancy in "fuzzy" codes reflects a discrepancy in pragmatic skills.

It was noted above that Vance and Wells' methodology reduced the pragmatic demands of idiom comprehension in natural conversation.
by alerting the children to the presence of idioms and priming them to select the idiomatic meanings. In doing so, the authors largely eliminated that source of SP-LD difference in idiom comprehension which probably underpins the distinction between them in the present study.

In summary, the findings from the present study contradict those reported by Vance and Wells (1994). With respect to these SP and LD children, at least, it is possible to differentiate the two groups on the basis of their play demonstrations of twelve common idioms.

5. ACCOUNTING FOR THE SP GROUP'S CONSIDERABLE COMPETENCE ON THE PLAY TASK

While consistent with the literature which portrays SP children as having relative difficulty with idioms, these results present a considerably more optimistic picture of SP idiom ability than does most of the commentary in this regard. Indeed, the SP group scored more than twice as many "appropriate" items (60.9%) as "inappropriate" ones (25%, p = 0.0001). They were found, furthermore, not to be significantly more literal than any of the other three groups. In order to understand the better-than-expected performance of the SP group in this study, it is informative to compare the purpose of this study with that of the previous literature. As noted in Chapter 2, the identification of idiom comprehension as an area of difficulty for children with semantic-pragmatic difficulties derives either from case studies or from overviews of this population. The aim of these studies was predominantly to characterise a clinical subgroup. As such, their emphasis has been on the disabilities associated with semantic-pragmatic difficulties. The concern of the present study was not with what the SP group could not do, but rather with what they were capable of doing in facilitative circumstances.
Bearing in mind the different purpose of this study and that of the previous literature, the following discussion explores firstly, the aspects of previous commentary that might have led to an underestimation of idiom competence in children with semantic-pragmatic difficulties. Secondly, it examines the characteristics of the play task which may have facilitated SP idiom comprehension and enabled this group to demonstrate their ability. Finally, it considers the issue of whether or not the play task could have overestimated idiom comprehension in the SP group.

(a) HOW MIGHT THE UNDERESTIMATION OF IDIOMATIC COMPETENCE IN CHILDREN WITH SEMANTIC-PRAGMATIC DIFFICULTIES HAVE ARISEN IN PREVIOUS COMMENTARY?

The predominant emphasis on disability in the literature leads to the underestimation of idiom ability in children with semantic-pragmatic difficulties in both a direct and indirect manner. The direct effect is evident in the literature's attention to comprehension failures despite evidence of idiom comprehension success. The indirect effect manifests in what is arguably an overly-rapid acceptance of an idiom-disability view. Such unquestioning acceptance has perhaps deterred authors and clinicians from closer appraisal of idiom comprehension in children with semantic-pragmatic difficulties, and of their own definitions of idiomaticity.

As emphasised in Chapter 2, clause idioms such as "kick the bucket" and "spill the beans" represent the most commonly investigated idioms but are, in fact, just one of several types of idioms. Teacher and television language samples gathered for this study indicate that clause idioms are actually far less common in at least some contexts of interaction with children than are other idiom types such as phrasal verb idioms¹. Yet,

¹Teachers were found to use few clause idioms but produced a large number of phrasal verb idioms.
discussions with language unit teachers and SLTs during the study revealed that many did not recognise phrasal verbs, such as "line up" or "come on", as being idiomatic. It is thus likely that many clinicians fail to reflect on the children's comprehension of this important group of idioms when passing judgement on their idiom competence. The unintentional inclusion of idiomatic phrasal verbs in previous studies, as noted in the literature review, indicates that even researchers in the field may not always recognise the idiomatic status of these expressions.

Failure to recognise the idiomaticity of phrasal verb idioms may account too for the conspicuous absence of comment on the considerable use of idioms by children with semantic-pragmatic difficulties. Interestingly, the clinical examples included in the literature contain many illustrations of appropriate and spontaneous use of idioms by children considered to have semantic-pragmatic difficulties. As exemplified below, such usage is even evident in young children:

(i) "And he blows up balloon and just blow up another one" (John, age 4;6, Smith & Leinonen, 1992, 270)

(ii) "Do it all over again" (Peter, age 6;5, Smith & Leinonen, 1992, 285)

(iii) "I DID though when I was run over by a car" (age 9, Bishop & Adams, 1989, 250)

(iv) "He took the ball off Ian and he told him off." (Simon, age 10/11, Smith & Leinonen, 263)

(v) "Mummy doesn't let me watch them at seven-fifteen, because she watches 'Coronation Street' at half seven ... she watches 'Coronation Street' all the time." (McTear, 1985, 135).

In addition to the overlooking of idiom competence in children with semantic-pragmatic difficulties, the portrayal of this population as predominantly literal may also stem from the conflation of miscomprehension with literal comprehension. It was notable that the SLTs' Checklist judgements about individual subjects' literality (Checklist feature G32, see Appendix C) did
not correlate with actual literality on the play task. It is possible that this reflects inherent problems in the Checklist itself, but it may also reflect a tendency to assume that "getting it wrong" implies "getting it literal". The importance of this distinction is illustrated with reference to an example presented by Hyde-Wright and Cray (1991). During a practical session, the therapist said to the group, "Good, we are all ready to begin. Off you go." In response, a child with semantic-pragmatic difficulties in the group "left his seat and made to leave the room" (p. 86). The authors interpret this child's response as a manifestation of literal idiom comprehension. Certainly the child's reaction can be thus interpreted. However, it is also possible that this child has confused the target meaning of the idiom with its alternative idiomatic meaning: leave (a place) (Cowie and Mackin, 1993, Oxford Dictionary of Current Idiomatic English Vol I). If the latter is true, then this child's problem might stem from an inability to use context to work out the appropriate meaning or from a restricted idiom vocabulary, rather than from fundamental literality.

As illustrated by this discussion, the discrepancy between the "predominantly literal" and "considerable competence" portrayals of idiom comprehension in children with semantic-pragmatic difficulties probably stems in part from differences in focus and definition. It is further possible that the discrepancy reflects the consideration of different samples of this population. It was noted at the outset of this report that the term "semantic-pragmatic difficulties" probably encompasses a heterogeneous group of children with a considerable range of symptoms and underlying problems. Certainly this view was supported in this study. As discussed below (see within-group discussion), the 26 SP subjects displayed substantial variation in their ability to demonstrate accurate comprehension of these twelve idioms. It may be that the children described in the case studies and group overview literature are in fact more impaired in this regard than were many of the SP children who participated in this study.
Alternatively, it could be that the real-life contexts in which judgements of idiom difficulty have been made provide less support for a child with semantic-pragmatic difficulties than did the play task context. Jordan (1990) observes that whereas an artificial test situation might prove more difficult than real life for normally-developing children, a child with autism may fare better in a structured test situation. Vance and Wells (1994) conclude that a structured test of idiom comprehension might mask the problems that children with semantic-pragmatic difficulties have in this regard in normal conversation. The play task has been argued above to approximate real life to a greater extent than do previous methodologies. The discrepancy between judgements of poor comprehension reported in the literature and demonstrations of considerable idiom competence on the play task may nevertheless reflect differences in the data-gathering contexts.

Finally, it needs to be borne in mind that idiom comprehension norms are lacking. Indeed, the uniqueness of every idiom may make it impossible to develop generalizable norms for idiom comprehension. In the absence of such norms, it may be difficult to assess whether a young child with idiom comprehension difficulties is delayed or really age-appropriate.

(b) HOW MIGHT THE PLAY TASK HAVE FACILITATED IDIOM COMPREHENSION AND THE ABILITY TO DEMONSTRATE IT IN THE SP GROUP?

Between-tasks comparisons in the literature highlight the enormous impact that different types of tasks may have on the outcome of idiom comprehension assessment. The play task was devised in response to the inability of conventional approaches to tap the ceiling of a child's ability. The aim was to select a set of idioms frequently used with young children and to present them in a manner which facilitated idiom comprehension without artificially reducing its pragmatic demands.
As noted in Chapter 4, the nature of the play task constrained the types of idioms that could be included. Since it was necessary that the items be demonstrable with concrete props, idioms such as "see red" and "be on top of the world", whose idiomatic sense relates to an abstract state of mind, were all excluded. It may be that the less abstract nature of the idioms included made these items easier for the SP children to understand.

The concrete visual support afforded by the play set and props is likely to have further facilitated comprehension on this task. Previous studies have indicated that the visual skills of children with semantic-pragmatic difficulties tend to surpass their auditory skills (Jacklin, 1993; McTear, 1985). It was evident that most of the SP children scanned the play set appropriately during the first reading of the story. This indicates that the children recognised the connection between the words and their concrete referents. It is probable that the concrete play set provided an anchor for the transitory auditory input and thereby facilitated attention and memory on the play task.

Attention deficits and poor listening skills are likely to have been further offset by the repetition built into the methodology of this study. At the outset, the entire story was heard. It was then heard again, sentence by sentence, as the child acted it out. When the video was replayed, the story was heard for a third time and each sentence containing an idiom was once again repeated during the definition question.

In addition to reinforcing the words of the idioms, the repetition and concrete visual support may also have strengthened the overall story context for the SP children. It is difficult to assess the impact of the context on the SP group given the typical impairment in this regard reported in children with

---

¹Twelve of the 26 SP children scanned the set for at least three quarters of the story. A further 10 children scanned it intermittently (see Appendix F).
semantic-pragmatic difficulties (see sources of idiom difficulty discussion, below). Nevertheless, whatever ability these children had to recognise contextual anomaly and to integrate world knowledge and contextual information in interpreting idioms would probably have been maximized by the rich narrative.

The richness of the story context is likely to have been increased by the embedding of all twelve idioms into a single narrative. As noted in Chapter 2, the effect of context length on idiom comprehension in children has not, to this author's knowledge, been investigated. Schweigert & Moates (1988) found, however, that adult idiom comprehension was facilitated by the presentation of idioms in paragraph contexts rather than in single-sentence contexts. A study of metaphor comprehension in adults (Ortony et al., 1978) further demonstrated that contexts 33 to 60 words in length facilitated comprehension relative to contexts comprising only three to eleven words.

In addition to the importance of context length, the extent to which a contextualising narrative biases the idiomatic or literal sense of an expression has been shown to have a significant effect on idiom comprehension in children (Ackerman, 1982; Cacciari & Levorato, 1989; Levorato & Cacciari, 1992). Whilst the idiomatic interpretation of all of the idioms in this study made greater sense than did their respective literal counterparts, the extent to which this was true varied slightly. It was notable that the idiom which attracted the highest number of "literal" and "inferred literal" actions ("on the spot") was also the one whose context was least obviously idiomatically-biasing (see between-idioms discussion below).

Overall, it can be seen that there are a number of characteristics of the play task which might have facilitated idiom comprehension in the SP subjects. As important, however, is the extent to which the play task enabled the SP children to demonstrate their understanding. By making no verbal expression or metalinguistic demands on the subjects, this task increased
the likelihood of the children's considerable idiom comprehension being recognised.

(c) COULD THE PLAY TASK HAVE OVERESTIMATED THE SP GROUP'S IDIOM COMPREHENSION?

While a range of plausible explanations for the better-than-expected SP results have been proposed, it remains necessary to consider the possibility that the SP group's idiom comprehension was not really as proficient as the results would indicate. The critical question is whether the play task produced a large number of false negatives.

A search for possible sources of false negative results yields three possibilities. The first concerns the code "inferred idiomatic". While the criteria for "idiomatic" actions were relatively explicit, the coding guidelines for "inferred idiomatic" actions were necessarily more open-ended. This raises the possibility that the researcher might have made greater allowances for SP children when scoring the code "inferred idiomatic". The results do, however, militate against this proposal in two ways. Firstly, it is notable that the number of "inferred idiomatic" items scored by the SP group on the play task was not significantly different from that scored by any of the other three groups. Secondly, not one item scored "inferred idiomatic" by the author was coded differently by the second coder. As noted in Chapter 4, the independent, second-coder was blind to the classification of subjects and was unfamiliar with the nature of semantic-pragmatic difficulties. There is thus no reason to suspect that the "inferred idiomatic" results of the SP group are inflated.

A second criticism that might be levelled at the play task is that the demonstration of some idioms required the children to

---

1 In the context of this report, a "false negative" is an item scored "appropriate" despite not actually being understood by the child.
use props in a predictable way. In other words, the child might plausibly score "idiomatic" merely by performing the obvious action with the props mentioned in the sentence. This possibility is clearly illustrated by the idiom "under lock and key". Closing the lock around the open edge of the only door on the set may have seemed the obvious thing to do with a lock and key to a child who did not know the meaning of the idiom.

It is certainly conceivable that some children did score "appropriate" in this way. Indeed, "under lock and key" was associated with a high number of "appropriate" scores (19 out of 26 SP children). Critically, however, 17 of these 19 children also provided an "appropriate" definition of this idiom, thereby proving that their play task demonstrations were grounded in true understanding.

Another way of exploring this possible source of false negatives is by examining the idioms on which the SP group scored the greatest number of "appropriate" codes. Between-idioms analysis (discussed fully below) reveals that six idioms (apart from "under lock and key") were scored "appropriate" by at least half of the SP group: "dropped in", "got on to", "went round", "took off", "kept an eye on" and "gave a hand". Of these, it is only "got on to" that may be argued to correspond to the obvious or commonsense action in the circumstances. Interestingly, this idiom was demonstrated considerably better on the play task than it was defined by the SP group. Notably, however, this was true of all four groups, suggesting that this idiom may simply have been difficult to define (see further discussion below). That the other idioms on which the SP group performed well are not clearly susceptible to this strategy undermines the explanatory force of this obvious-action argument.

The obvious-action argument is further questionable in the light of SP children having performed a number of actions during the play task which involved the use of props in an unusual manner. For example, one SP child substituted Sally for Rover early on
in the story, allowing her to be placed in the dog's basket, be carried sideways and examined by a vet (who had been explicitly introduced as "a doctor who helps animals when they are ill or injured"). Other SP children drove the car into the house and vet's office or parked the motorbike upside down. Indeed, given the propensity of children with semantic-pragmatic difficulties to problems with inferencing (see below), it is unlikely that many SP children were able to be guided by probability in their play task demonstrations.

A more concerning potential source of false negatives is the use of a circumventing strategy. For certain idioms, it was possible for a child to perform the target action merely by demonstrating the words adjacent to the idiom:

(i) The bike rider took off without stopping.
(ii) Suddenly the bike rider stopped on the spot.
(iii) The policeman dropped in to visit Sally and Rover.
(iv) Rover went round everyone and licked them all.

For each of these sentences, a child could score "idiomatic" by accurately demonstrating the underlined words even if s/he ignored the idiom (emboldened).

The likelihood of widespread use by the SP subjects of this "ignore-the-idiom" strategy is immediately undermined by the demands it places on the child. In order to use this strategy efficiently, the child needs to realise that s/he does not understand the idiom string, to recognise the incongruence produced by a literal interpretation and, probably, to be aware of the unity of the expression. As discussed below, each of these demands may be problematic for a child with semantic-pragmatic difficulties.

Nevertheless, some small measure of support for this theory derives from a comparison of the idioms susceptible to this strategy with those below which are not susceptible to such an
approach:-

(i) The bike rider didn't see Rover and ran over him.
(ii) (The bike rider wanted his motorbike at the police station.) The policeman went to pick it up.
(iii) (Sally put Rover into his basket.) Sally kept an eye on him all day.

It is notable that three of the four susceptible idioms were associated with high levels of "appropriate" scores in the SP group ("took off": 73.1%; "dropped in": 96.2%; "went round": 76.9%). The fourth ("on the spot") was not well understood (46.2%). Of those not susceptible, one attracted a high level of "appropriate" scores ("kept an eye on": 73.1%) while the other two were poorly demonstrated ("ran over": 34.6%; "pick up": 38.5%).

While these results appear consistent with the operation of an ignore-the-idiom strategy, it is important to bear in mind that the idioms "ran over" and "pick up" were both associated with a high level of "ambiguous" scores on the play task. The definition task results confirm that the vast majority of SP children who scored "ambiguous" on the play task for "ran over", did in fact understand this idiom (see between-idioms discussion below). The pattern of results for the susceptible as opposed to the non-susceptible idioms is thus less clear than it might initially appear.

An alternative means of monitoring the effect of an "ignore-the-idiom" strategy is to examine the relationship of codes across the two tasks for the relevant idioms. While this strategy would allow a child to score "appropriate" on the play task, it would not enable the child to score an "appropriate" code on the definition task. Instead, the child's repetition or paraphrase of other words in the sentence would yield a code of "repeat", "narrative" or repeat-narrative". Thus, for idioms susceptible to this strategy, we would expect there to be a consistent
relationship between the items coded "appropriate" on the play task and those coded "repeat", "narrative" or "repeat-narrative" on the definition task.

When the play task "appropriate" items are compared with their definition task counterparts, it is evident that no such relationship exists. Twenty play task demonstrations of the idiom "went round" scored "appropriate". Of these, only six correspond to "repeat", "narrative" or "repeat-narrative" definitions. For "on the spot", only one of the eleven "appropriate" play task items corresponded to a "repeat", "narrative" or "repeat-narrative" definition. In the case of "took off", only one out of 18 play task "appropriate" items corresponded to a "repeat", "narrative" or "repeat-narrative" definition. This between-tasks pattern of results is in fact contrary to that which we would expect to find if the SP play task competence was the result of the "ignore-the-idiom" strategy.

Further between-tasks comparisons provide additional evidence that is inconsistent with a false-negative account of the SP play task results. The idiom "dropped in" differed from the eleven other idioms in that for this one idiom a repeat of a word in the sentence ("visit") was considered to be an "idiomatic" definition and was not coded "repeat" (see Appendix I). Thus, for this item alone, it was possible for the subjects to get away with using the "ignore-the-idiom" strategy on both the play task and the definition task. If the strategy was operating in the case of this susceptible idiom, one would expect that the level of "appropriate" scores should not fall on the definition task. On the contrary, "dropped in" was associated with a drop from 96.2% "appropriate" on the play task to 72% "appropriate" on the definition task. For this idiom, at least, the evidence militates against the SP play task competence being the result of an "ignore-the-idiom" strategy. Instead, it is probable that superior performance on the play task relative to that on the definition task reflects task differences.
It is impossible to prove beyond doubt that an "ignore-the-idiom" strategy was not used by any children for any of the idioms. Similarly, some degree of interference from an "obvious-action" strategy cannot be categorically ruled out. The evidence does not, however, support the operation of either strategy. If used at all, they do not appear to have had a substantial effect on the outcome of the SP group's results. We may therefore conclude that the better-than-expected performance of the SP group on the play task does indeed reflect true ability.

6. ACCOUNTING FOR THE SP GROUP'S POOR PLAY TASK PERFORMANCE RELATIVE TO THE MAINSTREAM AND LD GROUPS: CONSIDERATION OF PLAY TASK DEMANDS

The finding that the SP group performed significantly less well on the play task than did each of the other three groups is, as discussed above, consistent with a large body of literature which characterises this population as having difficulty in comprehending idioms. While it is probable that the SP group's relatively poor performance on the play task stems largely from a real problem with idioms, it is nevertheless important to consider those aspects of the play task which may have negatively influenced SP performance on this task.

(a) MOTIVATION

Like most idiom assessment procedures, the play task demanded a considerable level of co-operation from all of the subjects. As noted in Chapter 4, two SP subjects (both subsequently found to have been diagnosed with classic autism) failed to complete the task and were eliminated from the study. The remaining 26 SP children all completed the play task as well as the subsequent definition task. With only one exception, there was no need for the researcher to urge any child to continue. SP20, alone,
required occasional persuading to make a video rather than watching a particular cartoon with which he had an obsession. The extent to which the task was inherently rewarding to the children was evidenced in the reluctance among some to leave the session!

Attracting the children's attention and sustaining their motivation was largely accomplished by focusing on the making of a video. Not only were many children excited by the prospect of starring in a video, but their belief that the session was organised for this purpose detracted from the sense that this was a test situation. In fact, there was no indication during the play task that any child realised that it was a test. This was important in that the children were consequently unlikely to perceive this task as one on which it was possible to fail.

(b) THE SOCIAL CONTEXT

While every effort was made to make the children feel socially comfortable during the procedure, being alone in the presence of two adult strangers would certainly intimidate many children. In this respect, an "(absence of) normal fear of strangers" (Bishop and Rosenbloom, 1987) typical of some children with semantic-pragmatic difficulties would have operated in the researcher's favour. If the social context influenced performance at all, it is likely that the mainstream and LD children would have been affected to a greater extent than would the SP group. Indeed, many of the "SPD" children engaged in inappropriately intimate interactions with the researcher and camera-person.

While the presence of strangers was less likely to distress the SP children, pressure to communicate and/or make eye contact with the researcher could have proved distressing to some of the SP subjects. It is one of the strengths of the play task that there is virtually no need for the child to engage in social
interaction. With the sole exception of a "yes" or "no" response to the question, "Do you understand what I want you to do?", the children were not required to communicate with the researcher at all. The only other interactive skill demanded by the play task was turn-taking. It was necessary that the children be able to listen, then act, then listen again, and so on until the story was complete. For the most part, the SP children who participated in this study displayed no difficulty in this regard.

(c) GRASPING THE TASK REQUIREMENTS

As outlined in Chapter 4, the play task began with the playing of the entire story. This was intended to give the child an overview of the story and a sense of its length. It was notable that twelve of the 26 children scanned the play set visually for at least three quarters of the story as it was played for the first time. A further ten children were noted to have scanned the play set intermittently, but for at least a quarter of the story (see Appendix F). The fact that these children appeared to be focusing on the specific props mentioned in the story at any given time indicates that they did draw the connection between the story and the associated props.

Following the first playing of the story, the children were instructed to "make the people do whatever the story says". This was followed by a demonstration enacted by the researcher and a minimum of two trials. Where the two trial items were not accurately enacted, additional trials were completed. The SP subjects required between two and five trials (mean 2.96, s.d. = 0.96; see Appendix F).

Following the trials, the children were reminded to "make the people do whatever the story says" and were told that they might use any of the things on the table. The researcher then asked the children whether they understood what she wanted them to do.
Many of them paraphrased the task instruction at that point.

Despite the initial trials, several children failed to act at all when the first sentence was played. (This applied to several MI and LD children too.) In these instances, the researcher prompted the child by saying, "You do it." Apart from this initial uncertainty, the children all appeared to have grasped the requirements of the task. This initial hesitation might have been avoided by playing the trial items on tape - as the story would be played - rather than saying them to the children.

Given that the SP children did not otherwise display any difficulty with the task itself, it is interesting that a significant positive correlation was found between the number of trials completed for the SP children and their "inappropriate" scores on the play task (rho = 0.587, p = 0.0008). In interpreting this finding, it is important to bear in mind that correlation does not necessarily imply causation. It may be that general comprehension difficulties in some children undermined both their initial grasp of the task and their idiom comprehension difficulty.

While the SP children included in the analysis all demonstrated their understanding of what was required of them, it is noteworthy that the two classically autistic children eliminated from the study did not appear to have grasped the requirements of the task at all. This suggests that the play task is probably not suitable for use with more severely affected children with autism.

(d) PLAYING SYMBOLICALLY

The use of a play-based methodology to assess children with semantic-pragmatic difficulties was, in itself, experimental. Vance and Wells (1994) assert that
as limitations of play are characteristic of children with semantic-pragmatic disorder (Bishop and Rosenbloom, 1987), any procedure involving acting-out with toys might prejudice their performance (p. 31).

In the light of this assertion, it is illuminating to observe that the SP children in this study displayed no impairment in acting out any of the literal language contained in the story. Indeed, as evidenced by the results, they attained a high level of "appropriate" acting-out even with respect to the idiomatic language.

This finding appears less surprising when considered in the light of in-depth research on symbolic play in autistic children. Following Ungerer and Sigman (1981 cited in Baron-Cohen, 1987), Baron-Cohen (1987) draws a distinction between "pretend play" and "functional play". "Pretend play" includes using an object as though it were a different object, attributing properties to objects which do not have those properties, and referring to absent objects as though they were present. "Functional play", in contrast, refers to the use of objects - including miniature representations of objects - in a manner appropriate to their conventional use. Employing this distinction, Baron-Cohen (1987) found that autistic children produced significantly less pretend play than did normal and Down's syndrome children, but that their functional play did not differ significantly.

To date, no comparative investigation of functional and pretend play in children considered to have semantic-pragmatic difficulties has, to this author's knowledge, been conducted. It is interesting to note, however, that the subject of Jones et al.'s (1986) case study was reported to score above 36 months on the Symbolic Play Test (Lowe & Costello, 1976) when at the chronological age of 3;7. As Lewis & Boucher (1988) point out, the Symbolic Play Test is in fact a test of functional play. This, in conjunction with the finding that the SP children coped well with the play task, suggests that the conclusions drawn with respect to autistic children may well apply to those with
The demonstrated ability of the SP group to cope with the play demands of this task is also consistent with the findings of Lewis and Boucher (1988). Confirming the findings of Gould (1986 cited in Lewis & Boucher, 1988), Lewis and Boucher (1988, 1990) reported that the elicited and instructed play of autistic children were relatively unimpaired. In spontaneous conditions, autistic children demonstrated significantly impaired functional and "symbolic" play (the latter equates to Baron-Cohen's "pretend play") relative to normal controls. However, when play was elicited (e.g. "show me what you can do with these toys"), the autistic children were found to be unimpaired in either functional or symbolic play. Symbolic play was similarly unimpaired relative to normal controls in the instructed play condition (e.g. "show me how you put petrol in the car")

Interestingly, the play task in the present study represents a form of instructed play. The finding that the SP children diagnosed "ASP" coped with its demands supports the conclusions of Lewis and Boucher (1988) and Gould (1986 cited in Lewis & Boucher, 1988). That the "SPD" subgroup also demonstrated an ability to play in this way suggests that these findings regarding instructed play might be equally true of children considered to have "semantic-pragmatic disorder/difficulties".

While the SP children in this study all coped with the play demands of this task, their style of play was, on occasion, noticeably more meticulous than that of the other children. This was evident for example, in the response of some SP children to discovering that Sally's mum did not easily fit behind the wheel of the car. Most of the mainstream and LD children were content to have the doll stand in the driver's seat while the car was "driven". A number of SP children, in contrast, delayed further action and persisted in manipulating Mum for up to several

---

1 Instructed functional play could not be assessed as so much of it had occurred in the preceding elicited play condition.
minutes until such time as she was appropriately seated.1

Fortunately, such delays almost always occurred in relation to literal sentences in the story. When they did occur in relation to idioms, there was no evidence of their having affected the outcome of the action. For example, a child who took a long time to close the lock when demonstrating the idiom "under lock and key", would still have scored "idiomatic" once the action was completed.

While very precise in some respects, the play of some SP children displayed a degree of immaturity also noted in the MI and LD groups. In demonstrating the sentence "He moved them (the bandages) up", for example, some children used the vet to move the bandages whereas others performed the action themselves. As noted above, play involving the child as the agent of action on a doll precedes the use of a doll as the agent of action on another doll in normally-developing children (Bretherton, 1984 cited in Rescorla & Goossens, 1992; Garvey, 1991; Lowe, 1975; Watson & Fischer, 1977 cited in Rescorla & Goossens, 1992). While doll-as-agent play emerges in normally-developing children between 2;6 and 3;0 (Lowe, 1975), some evidence of self-as-agent play was still observed in the MI, LD and SP groups and even, to a more limited extent, among the MJ children. Given that this feature is age-related, children were not scored down for demonstrating idioms by means of self-as-agent play provided that the action could still be identified as clearly "idiomatic".

(e) FINE MOTOR SKILLS

In addition to making allowances for immaturities in play style, the coding guidelines sought, where possible, to make allowances for immaturities or impairments in fine motor co-ordination. Demonstration of the sentence, "He moved them (the bandages) up",

---

1This meticulous attention to detail might itself be argued to reflect literalness of a sort.
for example, required the child to push the bandages across/along the surface of the table. Given that the table was only 10.5cm long, considerable fine motor skill was necessary if the child was to accomplish this without allowing any of the bandages to fall off the table. For this reason, a code of "idiomatic" was scored even if one or two of the bandages dropped off the table as the child pushed them across, provided that s/he refrained from lifting them off the floor (which might have represented a literal demonstration of "moved up").

Despite such allowances, fine motor impairment could be argued to place a child at a disadvantage on the play task. Studies of children diagnosed with "SPD" are unclear on the matter of fine motor co-ordination. Bishop and Rosenbloom (1987) refer to impairment in gross motor skills but do not mention fine motor skills. McTear (1985) describes a single child's difficulty in manipulating a knife, fork and hammer in addition to impairment in a range of gross motor activities. Among the SP children in this study, no substantial fine motor difficulties were apparent. On occasions, some children did knock buildings with sufficient force to loosen them and caused other minor damage to props. These incidents caused delay but did not affect idiom demonstrations. The researcher ensured that the children were focused on the task before continuing. While fine motor problems did not appear to cause appreciable difficulty in these children, their association with Asperger syndrome (Green, 1990; Manjiviona & Prior, 1995; Szatmari, 1991) suggests that the potential influence of this factor needs to be borne in mind when this task is used.

(f) EFFECT OF PRESENTING THE STORY ON TAPE

Although the recorded reading was clear and all of the subjects were reported to have normal hearing, it occasionally appeared that children were unable to discern the words of a sentence on tape, even following a repeat of the sentence. This appears to
be part of the problem for SP19 (10;2) when he responds to the sentence, "Rover went round everyone and licked them all":

SP19 grasps the policeman and says, "What shall I do?" The author replies, "Do whatever the story says." The child holds the policeman leaning into the basket and says, "Then what does he do?" The author replies, "Let's listen again" and replays the sentence. The child stands the policeman and lifts Rover into the garden. He says, "went round everyone and looked for cars". He drives the car to the edge of the garden, saying "If that's what it is", then drives the car away. The child holds Rover standing in the road and says "can't understand what licked the mall" is. I'm sure it's something like watched out for cars." The child then puts Rover in the basket and says "there you go".

In addition to isolated problems of clarity, the use of a tape-recording might further affect idiom comprehension by reducing the nonverbal clues that indicate idiomatic usage. Ezell and Goldstein (1991) suggest that the recognition of idiom form may be partially accomplished through monitoring of a speaker's eye-blinking. It might, however, be argued that even if this proposal is confirmed in normally-developing children, the extent to which children with semantic-pragmatic difficulties are likely to be helped by monitoring eye-blinking is minimal.

(g) ATTENTION

Like any other assessment procedure, the play task is inevitably vulnerable to the effects of poor attention skills. This is an important factor since the literature suggests that poor attention and distractibility are common problems in children considered to have semantic-pragmatic difficulties (Bishop & Rosenbloom, 1987; Culloden et al., 1986; McConachie & Jones, 1991) or Asperger syndrome (Szatmari, 1991). Certainly some of

---

1 This relates to the previous action for the sentence, "The policeman patted Rover."

2 If, as is suggested by SP19's remark, he perceived the clause as "licked the mall", his inability to get from this to "licked them all" is indicative of greater inferencing difficulties.
the SP children who participated in this study appeared to have difficulty in this regard. As far as possible, the researcher strove to monitor children's attention during the task and replayed sentences when it was obvious that a child had been distracted. In some instances, children requested a repeat of a sentence themselves (see Appendix J). It is likely that this compensated for some instances of concentration lapse, but the possibility that attention problems did interfere with some subjects' performance cannot be ruled out.

Aside from general attention difficulties, lapses in concentration might also have stemmed from fatigue. The likelihood of interference from fatigue effects is partially reduced by the fact that the children undertook the play task before the definition task. Nevertheless, the length of time required by most of the SP children on the play task exceeded that required by the mainstream groups\(^1\). It could thus be argued that the SP group was more likely to suffer fatigue. The results confirm, however, that there was no deterioration in SP performance as the play task progressed. On the contrary, the three idioms which occur last in the story ("kept an eye on", "dropped in" and "went round") are all among the "best six" for the SP group, scored "appropriate" by between 73% and 96% of the SP children. There is thus no support for the SP group's play task results being negatively influenced by fatigue.

(h) SHORT-TERM MEMORY IMPAIRMENTS

Inevitably, a child who could not hold an entire sentence in memory even for a few seconds would have been disadvantaged by the play task procedure. Short-term memory plays at least two roles in successful performance on the play task. Assuming that the child is capable of interpreting a given idiom accurately, it is necessary that s/he hold the idiom sentence in memory long

\(^1\) The mainstream children usually completed the play task in approximately ten minutes. The SP and LD children usually took approximately 15 minutes.
enough to portray this interpretation. At an even more fundamental level, short-term memory is essential to the use of the narrative context to guide a child's interpretation of the idiom (see comprehension of connected discourse discussion, below).

The inconclusive literature on short-term memory in Asperger syndrome, high-functioning autism and semantic-pragmatic difficulties (see McTear, 1985; Lincoln, Dickstein, Courchesne, Elmasian & Tallal, 1992; Ozonoff et al., 1991) suggests that memory ability and disability in this population is probably variable. Among the SP children who participated in the present study, there was little indication of short-term memory impairment. Indeed, literal stretches of narrative were almost always accurately demonstrated. Furthermore, idioms embedded in long sentences were not associated with fewer "appropriate" scores than were those presented in short sentences. As discussed above, the concrete visual support afforded by the play set and the built-in repetition may have compensated for minor degrees of memory impairment.

(i) COMPREHENSION OF CONNECTED DISCOURSE

While the comprehension of connected discourse will be argued below to be an essential aspect of all idiom comprehension in context, its importance on the play task may be greater than in assessments which use pictorial multiple choice methodologies. In order to demonstrate the idiomatic actions accurately, it was necessary that the participants in this study not only grasped within-sentence syntax, but were also able to make connections across the narrative. It was important that children be able to follow who the protagonists were, to make sense of referents and to maintain a sense of theme and topic.

1 In cognisance of reference difficulty among children with semantic-pragmatic difficulties, the story aimed to minimise the use of pronouns, especially where their referents might be ambiguous.
The literature on processing of connected discourse in children with semantic-pragmatic difficulties and Asperger syndrome indicates that this aspect of the play task might well have placed SP children at a disadvantage (McTear, 1989; Smith & Leinonen, 1992; Szatmari, 1991). Interestingly, most of the SP children consistently used the appropriate dolls and props and defined many of the idioms in terms of story events. This suggests that these particular SP children followed the story with a considerable degree of accuracy. In some instances, this was confirmed by children retelling the story to the researcher or to a teacher after the session.

Nevertheless, it is difficult to assess the extent to which the SP group did suffer relative to other children as a result of difficulties with connected discourse. It is also very difficult to tease out the role of such comprehension in the play task as opposed to its involvement in all idiom comprehension in context.

7. SOURCES OF IDIOM COMPREHENSION IMPAIRMENT IN CHILDREN WITH SEMANTIC-PRAGMATIC DIFFICULTIES

As outlined in Chapter 2, accurate idiom comprehension is dependent, firstly, on the child's adequate acquisition of the cognitive framework for idiomaticity, secondly, on the child's adequate acquisition of individual idiomatic expressions, and thirdly, on the child's ability to retrieve and select the idiomatic meaning in context. In order to address the question, "Where and why might these SP children have difficulty in comprehending idioms?", it is helpful to begin by exploring the potential problems that could arise at each of these three stages.

---

1 The two autistic children eliminated from the study were considered by their teacher to lack the necessary skills for following a story of this complexity.
Idiomatic competence involves not only the ability to acquire and interpret individual idioms, but also a fundamental grasp of the concept of idiomaticity. If a child is to have a chance of interpreting idioms accurately in context, it is necessary that s/he possess the knowledge that the same phonetic pattern may correspond to more than one meaning (semantic duality) and that surface and intended meanings may not always be one and the same.

The attainment of semantic duality by at least some of the SP children was forcefully illustrated by such definitions as follow:-

Example 1 (SP5, 9;5):

A: The bike rider took off without stopping. What do you think "took off" means?
SP5: it means um well in if it was an aeroplane then it would have left the runaway (sic) if it was a rocket it would have gone up into space if it was a motorbike then you would just rode on and not even bothered to stop

Example 2 (SP19, 10;2):

A: The bike rider took off without stopping. What do you think "took off" means?
SP19: oh I don't know what you do with a bike but an aeroplane goes (over) the air that's took off

In addition to a grasp of semantic duality, the child must know that the surface meaning of an utterance may not correspond to that which the speaker or writer intended to convey. Such understanding is itself dependent on a grasp of the "co-operative principle" (Grice, 1975, 1978 cited in Levinson, 1983) and of adequate theory of mind to recognise that others may have intentions.

While theory of mind may be impaired to varying degrees in
children with semantic-pragmatic difficulties (see discussion below), the fact that each of the 26 children in the SP group scored "appropriate" for at least four idioms suggests that the cognitive framework for idiomaticity was in place. This is further supported by the observation that each child within the SP group used at least one (and, in some cases, many) idioms spontaneously and appropriately during the session

(b) ACQUISITION OF INDIVIDUAL IDIOMATIC EXPRESSIONS

The relative competence of the SP group on the play task indicates that these children have some ability to acquire individual idiomatic expressions. Nevertheless, their relative comprehension difficulty may have stemmed from their being familiar with fewer idioms than were the other children. A limited idiom repertoire might itself relate to inadequate exposure to idioms and/or to a failure to learn from such exposure.

(i) Lack of exposure to idioms

Based on the acquisition-via-exposure hypothesis (Strand & Fraser, 1979), the theory that the SP group was less familiar with these idioms as a result of lesser exposure to them is consistent with recommendations that teachers avoid idioms and other forms of non-literal language in interaction with language-disordered children (Blue, 1981; Gross, 1994). However, given the extent to which idioms are used in the English language, it is debatable whether such advice can be heeded, let alone whether this is a desirable goal. As illustrated by the following comment reported in Bruck and Ruckenstein (1981), it cannot be assumed that idiom usage by teachers is reduced in response to

---

1 This usage included clause idioms (e.g. "stopped dead", "didn't move a muscle", "be as good as gold"), adverbial phrase idioms (e.g. "straight away", "in the first place", "on her own") and, predominantly, phrasal verb idioms (e.g. "look after", "tell off", "popped in", "blow up").

---

255
such recommendations:

One therapist reported that despite continuous reminders to teachers to use short sentences and simple vocabulary when communicating with LD (language disordered) children, the teachers followed these instructions only for brief periods of time and then reverted to normal patterns (p. 218).

Interestingly, most of the language unit teachers tape-recorded for this study did claim to use idioms very rarely in the classroom. The recordings indicated, however, that the number and types of idioms used by language unit and mainstream teachers were very similar. Discussions with teachers revealed that many had not included phrasal verb idioms in these estimations. Indeed, most of their idiom output was of the phrasal verb type.

Two points emerge from this finding which militate against an exposure-based theory of the SP group's idiom difficulty. Firstly, even in cases where language unit teachers were attempting to avoid idioms, their pupils were nevertheless exposed to many idiomatic expressions. Secondly, despite the language unit children being exposed to many more phrasal verb idioms than clause idioms in these classrooms, the children from these units did not display greater difficulty in relation to those infrequently used (clause idioms) than to those very frequently used (phrasal verbs idioms).

The validity of the exposure-based view is further undermined by the fact that, in many instances, the SP and LD children were pupils in the very same classes. Since these children were in contact with the same teachers and SLTs, it is unlikely that their exposure to idioms would have differed greatly, at least in the school environment.

A further reason for rejecting an exposure-based account of the SP group's difficulties is that several of the idioms included were drawn from children's television. There is no reason to suppose that SP children watch less television than any of the other groups of children investigated.
It is difficult to assess the extent of exposure to idioms by SP children outside of the educational context. To this author's knowledge, no research has examined parental use of idioms in interaction with SP children. Studies of parent interaction with language-impaired and autistic children indicate that parents tend to adjust their language to the needs of their children (Conti-Ramsden, 1991; Conti-Ramsden & Friel-Patti, 1983, 1984; Horsborough, Cross & Ball, 1985 cited in Conti-Ramsden, 1991; Wolchik & Harris, 1982). Interestingly, while not discussing the use of idioms, Conti-Ramsden (1991) observes that mothers used fewer double function terms in interaction with their language-impaired children than in interaction with normal siblings of the same language stage. It may well be that parents try to avoid idioms when talking to their children with semantic-pragmatic difficulties, if it appears that they do not understand them. However, given that it would be very difficult to avoid all idioms and that phrasal verb idioms may not be recognised by many adults as being idiomatic, it is likely that these children are exposed to a great number of idioms in their parents' discourse.

Many of the SP and LD children in this study were in fact boarders at their schools. A week-long observation by the author in the residence attached to a language unit confirmed that both the LD and SP children were being exposed to many idioms in interaction with residential carers. It appeared, however, that the SP children's interactions with other children were very limited. It is plausible that their exposure to idioms in interaction with peers is reduced relative to other children. Nevertheless, in the light of the undermining arguments presented and the fact that the idioms included in the story are very common in the English language, a purely exposure-based argument does not seem to offer a satisfactory account of the discrepancies in idiom comprehension found in this study.
(ii) Failure to learn from idiom exposure

While exposure levels alone may fail to account adequately for the children's difficulties in acquiring idioms, there is reason to suspect that children with semantic-pragmatic difficulties may benefit less than others from their idiom exposure.

In order to acquire an unfamiliar idiom, the child must, at the outset, recognise that the literal meaning of the expression does not make sense in the context. As outlined in Chapter 2, the "global elaboration hypothesis" (Levorato & Cacciari, 1992) proposes that the recognition of contextual anomaly alerts the child to search for a meaning that does make sense. This process is part of an overall drive to search for coherence and relevance in all texts and discourse (Sperber & Wilson, 1986).

Once alerted to the contextual anomaly created by a literal interpretation of an expression, the normally-developing child can use exposure as an opportunity for asking the meaning of an unfamiliar idiom or for inferring its meaning from the context. Repeated exposure to the given idiom allows the child to check the appropriacy of an inferred or learned meaning and to seek further help, if necessary.

If, as has been suggested, children with semantic-pragmatic difficulties lack the drive, the confidence or the ability to seek "central coherence" (Frith, 1989; Smith & Leinonen, 1992), then it is likely that the detection of a contextual anomaly and the ensuing search for a relevant interpretation will be obstructed. Whereas the search for relevance is continually reinforced in the idiom-normal child, its absence or weakness in a child with semantic-pragmatic difficulties is inevitably self-perpetuating. Indeed, one might argue that the SP child's expectation of relevance diminishes as s/he increasingly experiences discourse and texts that do not make sense. In their discussion of hyperlexia in children with high-level semantic difficulties, Spence, Fleetwood, Geliot, Wrench, Earles & Searby
(1989) observe that such children are not only reluctant to admit their poor understanding but may also fail to recognise the importance of understanding what they say or hear.

Indeed, the responses of the SP children in this study suggest that some may lack insight into their own comprehension difficulties. Following Dollaghan (1987), it may be argued that these children displayed difficulty not only with respect to comprehension, but also with regard to comprehension monitoring. In the following examples, SP21 and SP19 (both diagnosed "ASP") comment favourably on their own performance immediately after providing an incorrect definition.

Example 1 (SP21, 8;7):

A: The policeman went to pick it up. What do you think "pick up" means?
SP21: um/ pu- pu- put (it) up is that right?
A: say that again
SP21: put up
A: uhu
SP21: getting good at this

Example 2 (SP19, 10;2):

A: The policeman went to pick it up. What do you think "pick up" means?
SP19: easy everything's easy these questions it means pick up the motorbike and carry it there
A: uhu
SP19: yeah/ no hard questions

It is notable that such comments were observed only in the SP group. Clearly it is possible that children in the other groups similarly lacked insight into their comprehension failures but recognised the social inappropriacy of self-praise in the presence of a stranger. It is also possible that the SP children quoted in the above examples were aware of their errors and that their comments were an attempt to divert attention from their lack of understanding. Such a strategy does, however, involve theory of mind skills which cannot be assumed to be present in
children with Asperger syndrome (see discussion below). It is also possible that these remarks were not intended to comment on the preceding definition. They may have been general observations on the task as a whole. This interpretation is supported by the fact that some children, such as SP11 (9;6) quoted below, repeatedly objected to being questioned even though they were performing very well on the definition task.

A: He put him in a cell and kept him under lock and key. What do you think "under lock and key" means?
SP11: under lock and key? lock him in with a lock and with a lock and a key
A: uhu
SP11: so he can't get out
A: uhu
SP11: I know a lot of those tho- these things besides I've always known all these things (silence) besides I don't really need to be questioned about all this.

As important as insight into comprehension failure is the nature of the child's response to that failure. As illustrated by the following examples, some SP children did acknowledge their comprehension failure. Such admissions were, however, strikingly rare among the SP subjects.

Example 1 (SP8, 9;1):

A: The policeman dropped in to visit Sally and Rover. What do you think "dropped in" means?
SP8: like you drop someone in
A: uhu so what does "drop in" mean?
SP8: don't know (yawns)

Example 2 (SP19, 10;2):

A: The bike rider took off without stopping. What do you think "took off" means?
SP19: oh I don't know what you do with a bike but an aeroplane goes (over) the air that's took off
A: uhu
SP19: yes but what did it took off without stopping

As these examples demonstrate further, the SP children generally
made a concerted effort to respond to the author's questions even when aware of not knowing the target meaning. Indeed, their response pattern on the definition task bore much resemblance to the "strategies for maintaining a semblance of interactional competence" outlined by McTear (1985, 141):—

(1) if the demands of informativeness and accuracy cannot be easily fulfilled, then give any answer which is potentially appropriate;

(2) if no such answer can be given then fulfil the requirements of interaction by using a non-elliptical response and/or a stereotype acknowledgement;

(3) if more is required, then repeat your own preceding utterance;

(4) if all else fails, then retreat by saying "don't know" or by changing the topic.

That the SP children displayed an awareness of the pragmatic requirement to respond to a question was a positive feature of their definitions. Doing so even when they did not know the answer is also not unusual. Indeed, studies with normally-developing children (and adults) have repeatedly shown that they will consistently proffer an answer to nonsense questions or to those for which they have inadequate information1 (Finn, cited in Pratt, 1990; Hughes & Grieve, 1980; Piaget, 1928 cited in Pratt, 1990; Pratt, 1990). What is important, however, is that the SP children very rarely actively alerted the researcher to their lack of understanding. In contrast, the MI and LD children were far more likely to admit to not knowing an idiom2 even if they did then go on to provide some (inadequate) definition3.

In addition to admitting to not understanding, the MI and LD children were also occasionally observed to ask what a given

---

1These include questions that are bizarre (eg. "Is a nose quicker than a knee?"), contain nonsense words (eg. "Are there more Wugs or Glugs?"), or provide inadequate information (eg. "One day there were two flies crawling up a wall. Which fly got to the top first?")

2The MJ children usually indicated their apparent lack of comprehension by remaining silent. When prompted, these children were frequently noted to whisper, mumble or utter an "inappropriate" definition at speed.

3The number of comprehension failure acknowledgements exceeds the number of "don't know" items since some admissions of comprehension failure were followed by a definition which fulfilled the criteria for another code.
idiom meant. Bishop et al. (1994) point out that requests for help are an uncommon feature of children's conversations. The fact that some of these children did ask for help from a stranger suggests that they would be even more likely to do so in interaction with familiar adults and peers. In their everyday lives, this type of response to comprehension failure would likely yield an explanation of an unfamiliar idiom. Among the SP children, several requests for clarification were noted (see Appendix J), but on no occasion did a SP child explicitly ask what an idiom meant. Whether resulting from a desire to avoid detection, linguistic or social difficulties or failure to recognise the problem (see McTear & Conti-Ramsden, 1992), not seeking help will inevitably reduce the number of explanations received.

In such circumstances, the child becomes increasingly reliant on an adult's spontaneous teaching of idioms and/or on their own ability to infer the meaning of idioms. The enormous demands of idiom acquisition by inference are likely, however, to prove insurmountable to many children with semantic-pragmatic difficulties. Even the prerequisites for idiom inferencing are likely to present difficulty: the child must notice the contextual anomaly, have sufficient motivation and confidence to seek a relevant meaning and recognise that the individual words of the expression function as a unit.

How we come to recognise idioms as lexical items is still a matter of speculation at this point. Van Lancker and Canter (1981) observed that when speakers tried consciously to distinguish idiomatic and literal uses of the same phrases, they uttered the idioms at greater speed, with shorter inter-word pauses, and with fewer or weaker pitch accents than when the same words were intended literally. Adult listeners were able to discriminate idiomatic and literal meanings even when they heard just one of the pair in isolation. It may be that the phrasing of idioms helps children to identify them as units of meaning. In the absence of child studies in this regard, it is unclear at
what age children may be helped by prosodic clues. Difficulty in interpreting prosody observed in many SP children (Culloden et al., 1986) would, however, put this group at a considerable disadvantage relative to normally-developing children.

It is further plausible that normally-developing children might recognise that the words of a given idiom "go together" because they encounter the same combination of words (in some cases in an inflexible syntactic form) used repeatedly in similar contexts. Inadequate attention to context and/or auditory memory problems would hinder the observation of this pattern by some children with semantic-pragmatic difficulties. Grasping the unitary nature of idioms would be made even more difficult for a child who tended to focus on the part rather than the whole (McConachie & Jones, 1991; Rapin & Allen, 1987).

Given this theoretical support for children with semantic-pragmatic difficulties failing to recognise idioms as single lexical items, it is interesting to consider definitions of the type quoted below:-

Example 1 (SP24, 8;5):

A: Sally turned the bike rider over to the policeman. What do you think "turned over" means?
SP24: when you going round that what's turning means

Example 2 (SP21, 8;7):

A: He put him in a cell and kept him under lock and key. What do you think "under lock and key" means?
SP21 a key's something what you put on a lock means um you can't go get in

The evidence here is admittedly weak, but the underlined sections could be argued to reflect deconstruction of idioms by these SP children. It may prove fruitful for future studies to consider whether the breaking up of idioms into their constituent parts is typical of children with semantic-pragmatic difficulties.
Even if the child is able to grasp the unity of the idiomatic expression, the process of inferring its meaning still requires attention to context, integration of world knowledge and an ability to abstract from the specific to the general. For example, in order to infer the meaning of the idiom, "skating on thin ice", a child needs, firstly, to have and apply at least three pieces of world knowledge: thin ice is prone to cracking, falling through cracked ice into cold water threatens one's wellbeing, and skating on thin ice increases the likelihood of its cracking. While bearing all of these facts in mind, the child needs, secondly, to identify the theme of vulnerability within the context before, thirdly, integrating these two sources of information to arrive at an abstraction from the specific literal situation to the general idiomatic meaning.

In normally-developing children, transparent idioms such as "skating on thin ice" have been found to be better understood than are metaphorically opaque idioms (Gibbs, 1987; Nippold & Rudzinski, 1993; Nippold & Taylor, 1995). Among children with semantic-pragmatic difficulties, whose inferencing skills are characteristically weak (see discussion below), there is no reason to expect an acquisition advantage for transparent idioms over opaque ones. The composition of the idiom set in this study does not allow for meaningful comparison of transparent and opaque idiom comprehension. By focusing on this difference, future research might, however, yield information pertinent to the maximising of idiom acquisition in children with semantic-pragmatic difficulties.

Given the demands of inference and the reduced likelihood of their asking for help, it is probable that children with semantic-pragmatic difficulties are reliant to a very great extent on others' observation of their difficulty and on spontaneous teaching of idioms by adults and peers. Even explicit teaching may not necessarily guarantee idiom comprehension in children with semantic-pragmatic difficulties, however. In addition to possible problems in generalizing from
one situation to another (Culloden et al., 1986; Hyde-Wright & Cray, 1991; McTear, 1989), the child may need to overcome the additional hurdle of rigid concept boundaries (Hyde-Wright & Cray, 1991; McConachie & Jones, 1991; McTear, 1991; Smedley, 1989). McTear (1991, 25), for example, describes a child who says "the clock is by the wall", possibly because he believes that the preposition "on" should be used only to refer to locations on the horizontal plane. In learning idioms such as "get on to", "get on with", "run on" and "come on", this child may need to confront not only the non-literalness of these phrases, but also the rigid concept boundary which s/he may perceive in relation to the preposition "on".

The usefulness of an explicit definition to a child with semantic-pragmatic difficulties will inevitably be further influenced by the nature of the definition. As illustrated by many of the paraphrases provided in this study, it is often difficult to define an idiom without using another. Often, the fact that we have used an idiom - particularly if it is a phrasal verb idiom - may escape our notice. It may be that many idioms which we think we have taught to children with semantic-pragmatic difficulties are still as unclear to them as they ever were.

(iii) Conclusions

As evidenced by this discussion, the skills required for learning new idioms are numerous. The acquisition of semantic knowledge inevitably requires an array of pragmatic skills which may be impaired to varying degrees in different children with semantic-pragmatic difficulties. The results of this study indicate that SP children were not incapable of acquiring idioms, but might have been doing so at a slower rate and through fewer channels of learning than other children. The twelve idioms included in this study are all very commonly used in English. Frequent exposure to these may have facilitated the SP children's inferencing of their meanings and/or increased the likelihood of
their comprehension failure being noticed by parents, teachers and clinicians. Future research might fruitfully compare comprehension of these common idioms with those to which children are likely to have been less frequently exposed.

(c) RETRIEVAL AND SELECTION OF IDIOMATIC MEANINGS IN CONTEXT

Assuming the semantic knowledge of an idiom has been acquired, the interpretation of idioms in context demands that the appropriate meaning be retrieved from the mental lexicon and be selected over any other literal or idiomatic candidates.

As discussed in Chapter 2, idiom processing theory has derived almost entirely from adult studies. It is therefore only possible to speculate on the process of idiom interpretation in children and on the manner in which semantic-pragmatic difficulties might impact on this process.

(i) Retrieval of the idiomatic meaning from the mental lexicon

The mechanism by which listeners retrieve an idiomatic meaning already encoded in their mental lexicons remains an issue of considerable debate. At this stage, it is unclear whether idiomatic meanings are retrieved before, after or simultaneously with literal meanings, or if each of these is associated with different types of idioms. The differences inherent in these theories do, however, have important implications for idiom comprehension in children with semantic-pragmatic difficulties. Indeed, further research in this field might succeed in shedding some light on the processing debate.

The "literal-first" hypothesis (Bobrow & Bell, 1973) argues that the idiomatic meaning is only retrieved once a literal interpretation has been found to produce a contextual anomaly. Clearly, if the child does not realise that the literal meaning
fails to make sense in the context - as might occur in a child who pays inadequate attention to context or is unable to use it effectively - the idiomatic meaning may never even be retrieved. A rather different scenario is suggested by the "idiom-first" hypothesis (Gibbs, 1980, 1985, 1986) and by the "simultaneous-processing" view (Swinney & Cutler, 1979). In terms of these theories, the idiomatic meaning (or meanings) should at least be retrieved from the mental lexicon, provided that the child recognises the idiomatic string as a unit of meaning. As discussed above, the recognition of the unity of idiomatic expressions may be undermined by difficulties in interpreting prosodic and contextual information and a tendency to focus on parts rather than on the whole.

Thus, it can be seen that the process of idiom retrieval may never even be triggered in some children with semantic-pragmatic difficulties. If the process is triggered, the efficiency with which it proceeds may depend on such features as the relative flexibility of the idiom (Gibbs & Gonzales, 1985), the relative transparency of the item (Mueller & Gibbs, 1987), its degree of compositionality (Mueller & Gibbs, 1987), the frequency with which the idiom occurs in the language (Flores d'Arcais, 1993), the likelihood of the expression being used idiomatically rather than literally (Cronk & Schweigert, 1992; Popiel & McRae, 1988; Tompkins et al., 1992), and the point at which the expression can be identified as an idiom (Cacciari & Tabossi, 1988).

The mechanism by which these factors influence processing is still unclear. It is thus difficult to assess their impact on idiom processing in children with semantic-pragmatic difficulties. It is evident, however, that at least some of the features which facilitate rapid idiom processing involve the use of context. As discussed above, the apparently facilitative effect of metaphorical transparency on normally-developing children may not apply to children who have difficulty in integrating world knowledge and contextual information. Thus, while transparent idioms may be processed more quickly than
opaque ones by normally-developing children\(^1\), transparent idiom processing by children with semantic-pragmatic difficulties may not be expedited in this manner.

Slower idiom processing in children with semantic-pragmatic difficulties may be explained further by a theory proposed by Cacciari and Tabossi (1988) (reviewed in Chapter 2). The authors hypothesize that the speed at which an idiomatic meaning is accessed is determined by the speed at which the configuration is identified. If the point of idiom identification is reached early in the expression, then any remaining words may not be processed literally. As observed by Flores d'Arcais (1993), the speed at which the "point of idiom uniqueness" is attained is partially determined by the context. Thus, in the context of talk about secrets, the meaning of "spill the beans" becomes available earlier (probably immediately after the word "spill") than it would without such a context. Failure to make predictions from the context on the basis of world knowledge and previous linguistic experience may therefore delay access to the idiomatic meaning and the speed at which this meaning can be selected. If this is true, then the encountering of idioms in discourse is likely to make it difficult for listeners with semantic-pragmatic difficulties to keep up with speakers.

(ii) Selection of the idiomatic meaning in context

The process of interpreting idioms in context requires not only that the child be able to retrieve semantic knowledge from memory but also that s/he be able to select from all possible meanings associated with the given string. The alternative candidates may include a literal interpretation and/or other figurative senses.

As discussed in Chapter 2, the presentation of idioms within an

\(^1\) At present, this processing speed advantage has only been recorded in adults, but child studies suggest that transparent idioms are better understood than opaque ones (Gibbs, 1987; Nippold & Rudzinski, 1993; Nippold & Taylor, 1995).
idiomatically-biasing context has repeatedly been found to facilitate normally-developing children's selection of idiomatic meaning over literal meaning (Ackerman, 1982; Levorato & Cacciari, 1992). In order to assess the extent to which children with semantic-pragmatic difficulties might similarly be guided by contextual information, it is helpful to consider the manner in which a normally-developing child might have used the narrative context in this study. By way of example, let us consider the interpretation of "took off" within the following context:

The bike rider didn't see Rover and ran over him. The bike rider took off without stopping.

Provided that the child has understood the idiom "ran over", the first sentence has the potential to increase considerably the child's ability to predict the ensuing action. To benefit fully from this contextual information, it is necessary that the child be able to theorise about the bike rider's state of mind after the accident and integrate relevant world knowledge. Firstly, the child needs to realise that the running over of Rover was accidental. Integration of the appropriate world knowledge and experience would immediately enable the child to predict two possible outcomes: the bike rider might not be severely punished for his action because it was accidental, or the bike rider would be punished because he should have taken action to prevent an accident (eg. driving slowly or looking carefully). On the basis of this grasp of the situation, the child might theorise as to the bike rider's feelings and beliefs after running over Rover. It is likely that fear of repercussions would be one state of mind that the child attributes to the bike rider. If the child understands that actions are filtered through beliefs and feelings (Wellman, 1993), then this empathy, possibly coupled with the child's knowledge of hit-and-run accidents, sets up a limited range of possible next actions. The bike rider might stop because he believes he will not be punished, because it is the moral thing to do, or because he believes the accident was witnessed. Alternatively, his fear of the repercussions might
lead him to attempt escape. By the time that the second sentence is played, the child already has a limited set of possible next-actions predicted. S/he is thus primed to interpret the idiom, "took off", in its idiomatic sense.

In contrast, for a child who has impaired theory of mind, lacks the necessary world knowledge, or has difficulty in integrating world knowledge into current context, the possibilities for the ensuing action are wide open. Deprived of this rich background information, such a child may encounter the second sentence almost as though it had occurred in isolation.

The theory of mind skills required for the full potential of the context to be realised are likely to have been present in the MJ and MI children. By the age of four, most normally-developing children are able to demonstrate a grasp of the causes of emotion (Harris, Johnson, Hutton, Andrews & Cooke, 1989 cited in Baron-Cohen, 1993) and to understand that action is filtered through belief (Wellman, 1993). The use of world knowledge to guide interpretation is also evident in normally-developing children from as early as two or three years of age (Nelson & Gruendel, 1979; Olson & Nickerson, 1979; Pace & Feagans, 1984 all cited in McTear & Conti-Ramsden, 1992; Strohner & Nelson, 1974 cited in Tager-Flusberg, 1981).

The prediction of emotion from events by children with semantic-pragmatic difficulties is not well discussed in the literature. Informal school-based observation by this author of children diagnosed with "semantic-pragmatic disorder/difficulties" revealed substantial difficulty among this group in predicting emotion from events. The inability to "put themselves in the place of others" in role play is described by Hyde-Wright and

---

1 For a more in-depth discussion of the role of world knowledge in the comprehension process, see for example Campbell & Bowe (1977), McTear (1989), McTear and Conti-Ramsden (1992), Milosky (1990) and Paul (1990).

2 Emotion recognition (from facial expression and tone of voice, for example) is, however, commonly cited as presenting difficulty to children considered to have semantic-pragmatic difficulties (Bishop & Rosenbloom, 1987; Culloden et al., 1986; Hobson, 1985a,b cited in Baron-Cohen, 1993; Smith & Leinonen, 1992).
Cray (1991) as typical of all the children with semantic-pragmatic difficulties they have encountered in their language unit. Among children with Asperger syndrome, the application of theory of mind skills has been found to be generally lacking (Happe, 1994a,b). Baron-Cohen (1991 cited in Baron-Cohen, 1993) did, however, find that autistic children were unimpaired in their ability to judge emotion caused by a situation.

As noted above, inferencing requires not only that emotion be judged, but also that consequent action be predicted on the basis of emotion. This requires a grasp of cause and effect (which may be impaired in children with semantic-pragmatic difficulties) as well as the ability to integrate world knowledge with the current context and to use inferencing skills to produce a theory (or set of theories) about the ensuing action. The literature indicates that high-functioning autistic children are often insensitive to varying probability levels (Tager-Flusberg, 1981). Children with semantic-pragmatic difficulties have also been widely described as being substantially impaired in their ability to draw inferences on the basis of world knowledge and contextual information (Bishop & Rosenbloom, 1987; Conti-Ramsden & Gunn, 1986; Culloden et al., 1986; Hyde-Wright & Cray, 1991; McTear, 1989; McTear & Conti-Ramsden, 1992). As with high-functioning autistic children, it is still unclear whether or not children considered to have semantic-pragmatic difficulties can be distinguished from SLI children on the basis of inferencing.

---

1 Eisenmajer & Prior (1991 cited in Prior, 1992) suggest that not all children with Asperger syndrome are unable to pass theory of mind tests. Others propose that theory of mind develops eventually in individuals with Asperger syndrome (Bowler, 1992; Ozonoff et al., 1991 both cited in Happe, 1995). It is acknowledged, however, that even if children with Asperger syndrome are able to pass theory of mind tests, they may nevertheless fail to apply these skills in real life. Interestingly, Happe's (1994b) study suggests that autistic individuals who are able to pass a battery of false belief and deception tests may be less impaired in understanding idioms in context than other forms of non-literal language. Happe's inclusion of subjects ranging in age from 8;9 to 45;1 (mean 20;6) and the fact that only two idioms were investigated prevents the drawing of conclusions with respect to SP children. The study is reviewed in Chapter 2.

2 Smedley (1989) suggests that difficulty in expressing and understanding causal relationships may stem from a linguistic problem with respect to conjunctions like "because", rather than from a nonverbal cognitive impairment.

3 See Paul (1990) with regard to the debate on the impairment of world knowledge usage in high-functioning autistic children relative to that in SLI children.
skills (see Bishop & Adams, 1992)\(^1\).

At this point, it is certain that at least some children with semantic-pragmatic difficulties do have substantial difficulty in drawing inferences on the basis of contextual information and world knowledge. It is likely that the use of context to trigger the retrieval of idiomatic meanings and their selection was undermined by these impairments in at least some of the SP subjects. It should be borne in mind, however, that the number of literal items scored by this group was not significantly greater than that scored by the mainstream groups. The SP group also scored a greater-than-chance number of "appropriate" scores. These findings suggest that when faced with the choice of an idiomatic and literal meaning, the SP children were able to use context to make the appropriate selection at least some of the time. Further evidence of some ability in this group to think about the mental state of others and to integrate contextual information and world knowledge derives from the following definitions:

Example 1 (SP14, 8;9):

A: The bike rider took off without stopping. What do you think took off means?

SP14: It means trying to get away so they don't see who done it.

Example 2 (SP5, 9;5):

A: Sally kept an eye on him (Rover) all day. What do you think kept an eye on means?

SP5: It it means that she looked after him and never let him out of her/ sight

A: uhu

SP5: because she didn't want um all that to happen again

\(^1\)As is often the case with semantic-pragmatic difficulties, the investigation of inferencing difficulties in SP children is inevitably somewhat circular since, by definition, a child with inferencing difficulties has a pragmatic difficulty.
As emphasized by Leinonen (1991), the fact that a child has knowledge does not ensure that the knowledge will be mobilized. It may be that some SP children had the potential to use context to guide accurate idiom interpretation but failed to pay attention to the context, focusing instead on single sentences of the narrative. Indeed, it is possible that the pausing of the story after each sentence during the play task may have encouraged such behaviour.

While there are clearly a number of mechanisms by which idiom processing can be undermined by a failure to "see the whole picture", it is interesting to consider the possibility of an impact in the opposite direction. In discussing the role of lexical phrases in communication, Nattinger and DeCarrico (1992, 32) make the point that

> it is our ability to use lexical phrases...that helps us to speak with fluency. This prefabricated speech has both the advantage of more efficient retrieval and of permitting speakers (and hearers) to direct their attention to the larger structure of the discourse, rather than keeping it focused narrowly on individual words as they are produced.

While referring specifically to lexical phrases rather than idioms - in terms of their definition, the two overlap - Nattinger and DeCarrico's point raises the possibility that the search for coherence by children with semantic-pragmatic difficulties may be hindered by their inefficient processing of idioms. Whereas the use of idioms may afford normally-developing children more processing time in which to concentrate on the bigger picture, children with semantic-pragmatic difficulties may require considerably longer processing time just to interpret the idiom itself.

(iii) Conclusions

The pragmatic demands of idiom retrieval and selection clearly have the potential to obstruct accurate idiom processing. The
results of the play task confirm that these 26 SP children were capable of retrieving and selecting idiomatic meanings in context. Their variable results suggest, however, that these aspects of idiom comprehension are not of an all-or-nothing nature. It may be that their development in these SP children has been partial but still remains immature. Alternatively, following Leinonen (1991), we may find that many of these children had the requisite "knowing that" elements of idiom comprehension, but fell down in not "knowing how" to mobilise this knowledge. They may have known, for example, that groups of words can function together to produce a different meaning, that speakers do not always intend exactly what they say and that "pick up" can mean "fetch". They may have had greater difficulty in knowing how to recognise a unit of meaning and how to work out when "pick up" means "fetch" rather than "lift", however. What is clear is that many different sites and sources of breakdown can be identified. It may therefore be too simplistic to address the question "Where does idiom breakdown occur?" to the SP group as a whole.

8. WHERE AND WHY DID THE SP CHILDREN BREAK DOWN IN COMPREHENDING IDIOMS? EVIDENCE FROM THE WITHIN-SP GROUP RESULTS

As discussed above, discriminant analysis and Mann-Whitney Tests reveal significant differences between the SP group as a whole and each of the other three groups. Nevertheless, there is evidence of considerable variance within the SP group data. Indeed the standard deviations (on individual codes) for the SP group are almost invariably substantially larger than are those for the MJ, MI and LD groups, respectively.

Greater variance among SP subjects than among mainstream ones is not a surprising result. Whereas the ages of children in the MJ and MI groups vary within only twelve months, the SP group spans
the entire range 6;6 to 11;3. Significant correlations between SP age and play task performance confirm that the variance in this group stems partially from age differences ("appropriate" rho = 0.402, p = 0.0210; "inappropriate" rho = -0.483, p = 0.0063). In the mainstream groups, where age differences are only slight, no significant correlations between age and performance (on either task) were found. The greater variance within the SP group than within either mainstream group can therefore be attributed in part to the wider age range spanned by the SP group.

A positive correlation between age and idiom comprehension is consistent with the developmental literature reviewed in Chapter 2. Cognitive and linguistic development, along with more opportunity for exposure to idioms and to explicit teaching by parents, teachers and clinicians, might each account in part for this improvement. What is interesting, however, is that the mainstream children at each end of the age range spanned by the SP group (MI and MJ groups) did not differ significantly in their comprehension of these idioms. In fact, all but two of the 6- and 7-year old children scored "appropriate" for two thirds or more of these idioms. In contrast, it was not until the age of 10 that most SP children scored "appropriate" for at least two thirds of these idioms.

While the SP age-performance correlations are significant, their magnitude indicates that age differences alone cannot account for all of the variance within the SP group. Indeed, some 6- and 7-year old SP children did very well on the play task while some of the eldest SP children performed very poorly.

As outlined in Chapter 2, the heterogeneity of the SP group related not only to age, but also to diagnosis. The two subgroups comprised children diagnosed as having "semantic-pragmatic disorder/difficulties" ("SPD") and children diagnosed

---

1 Significant age-performance correlations in the SP group were also found on the definition task: "appropriate" rho = 0.562, p = 0.0014; "inappropriate" rho = -0.584, p = 0.0009.
as having Asperger syndrome or high-functioning autism ("ASP")\(^1\), respectively. Interestingly, a discriminant analysis of the play task data revealed that 14 out of 17 (82.4%) " SPD" children and eight out of nine (88.9%) " ASP" children could be assigned to the correct subgroup on the basis of their play task results. Statistical examination of the results confirm that the " SPD" subgroup scored significantly more play task items "appropriate" (65.7%) than "inappropriate" (20.1%; \( p = 0.0007 \)). The difference in the " ASP" subgroup between "appropriate" (51.9%) and "inappropriate" (34.3%) was considerably smaller (n.s.). A series of Mann-Whitney Tests revealed further that the " SPD" and " ASP" subgroups differed significantly from each other on the codes "idiomatic" and "fuzzy". The " SPD" subgroup scored "idiomatic" for 55.9% of their items (mean 6.71, s.d. = 2.23). The " ASP" subgroup scored only 39.8% of their items "idiomatic" (mean 4.78, s.d. = 1.86; \( p = 0.0294 \)). The code "fuzzy" accounted for 11.8% of the " SPD" items (mean 1.42, s.d. = 0.87), but for as many as 24.1% of the " ASP" items (mean 2.89, s.d. = 2.26; \( p = 0.0363 \)). The two subgroups did not differ significantly on any other codes, including "literal" and "inferred literal".

When the two subgroups were compared individually with the MJ, MI and LD groups, it emerged that both scored fewer "appropriate" and more "inappropriate" items than did any of these three groups. However, the magnitude of the difference is greater for the " ASP" subgroup than it is for the " SPD" subgroup. Indeed, the number of "appropriate" items scored by the " SPD" subgroup (65.7%) is not significantly less than that scored by the " LD" subgroup (74.4%). The "appropriate" difference is, however, significant between the LD group and the " ASP" subgroup, who scored only 51.9% "appropriate" (\( p = 0.0020 \)). With regard to "inappropriate" results, the " SPD" subgroup scored significantly more than the LD group (" SPD" 20.1%, LD 10%, \( p = 0.0112 \)). This discrepancy was nevertheless smaller than was that between the

\(^1\)As noted in Chapter 1, the grouping together of children with the diagnoses "Asperger syndrome" and "high-functioning autism" reflects the fact that many children in this subgroup were described to the author as having "Asperger syndrome/autism" or "Asperger syndrome/high-functioning autism".
"ASP" subgroup and the LD group ("ASP" 34.3%, LD 10%, p = 0.0015).

Relative to the MI group, a similar pattern of results emerges. Both the "SPD" and "ASP" subgroups scored significantly fewer "appropriate" items and significantly more "inappropriate" items than did the MI group. Once again, the discrepancy is, however, greater between the MI and "ASP" subgroup than between the MI and "SPD" subgroup.

In the light of Vance and Wells' (1994) non-significant difference between SPLD children and non-SPLD children, it is interesting that the proportion of "appropriate" items scored by the "SPD" subgroup here should have been not significantly less than that scored by the LD group. (The mean ages of the two groups were very similar - "SPD" 9;10 and LD 9;11 - although the range was bigger for the "SPD" subgroup.) Notably, however, Vance and Wells report a higher rate of idiomatic scores among SPLD than non-SPLD children.

That there should be a significant difference between the "SPD" and "ASP" subgroups is perhaps less surprising than is the finding that the "ASP" subjects did worse than the "SPD" subjects. In discussing the diagnostic distinction, Bishop (1989) characterises Asperger syndrome in terms of lesser impairment in "meaningful verbal communication" (i.e. "those aspects of language concerned with meaning and use" p. 275) in combination with more abnormal "interests and social relationships". "Semantic-pragmatic disorder" is described in terms of more normal "interests and social relationships" alongside greater difficulties in "meaningful verbal communication". Adams (1991) has suggested that the language problems characteristic of autistic children stem from the use

---

1 Vance and Wells' SPLD ("semantic-pragmatic language disorder") subgroup equates to the "SPD" subgroup in the present study. The non-SPLD subgroup is the equivalent of the present LD group.

2 "SPD" subgroup ranged from 7;2 to 11;3. LD group ranged from 8;6 to 11;2.
of obsessive and stereotyped language and that these children rarely display semantic errors. In this sense, their language profile differs from children diagnosed with "semantic-pragmatic disorder" who are more likely to display semantic errors.

While the comparative literature does not focus specifically on comprehension of idioms, it might be argued that it leads one to expect poorer idiom comprehension in the "SPD" than "ASP" subgroup. How then might we explain the finding that the "ASP" performance was poorer than that of the "SPD" subgroup?

One possible explanation is that the two groups did not in fact differ substantially and that the findings are merely an artifact of the groups' composition. It is notable, for example, that the groups are relatively small (nine "ASP" subjects and 17 "SPD" subjects) and may not be representative of the wider population of children thus diagnosed. It is conceivable, although unlikely, that this study stumbled on particularly impaired "ASP" children and unusually idiomatic "SPD" children.

It is further evident that the "SPD" subjects were, on average, older than the "ASP" children ("SPD" age range = 7;2-11;3, mean 9;10; "ASP" age range 6;5-10;6, mean 8;8). It might thus be argued that the difference between the "ASP" and "SPD" subgroups stemmed from differences in age. It is notable, however, that two of three eldest "ASP" children (SP3, SP19 and SP25, all aged over 10 years) failed to reach the mean number of "appropriate" items scored by the "SPD" subgroup. Indeed, no relationship was found between age and performance within the "ASP" subgroup.

An alternative explanation of these results is that the two subgroups differed not in terms of their comprehension of idioms, but rather in the skills necessary for demonstrating their understanding on the play task. As discussed above, successful demonstration of idiom comprehension on this task required sustained motivation and attention, intact short-term memory, the ability to engage in directed functional play, adequate fine
motor skills, adequate comprehension of connected discourse, a grasp of the nature of the task, and management of the social demands of the situation.

Observation of the children during the task and examination of their data provide no support for a difference between the two subgroups with respect to motivation, fine motor skills or the ability to act out literal stretches of narrative accurately. The "ASP" children did not appear to have poorer turn-taking skills than did the "SPD" children and generally coped well with the need to listen and act in turn. At a social level, the "ASP" children were noted to engage in less verbal interaction and eye contact with the researcher and camera-person than did many of the "SPD" subjects. As discussed above, however, such interaction was not required for the task and would have been unlikely to affect its outcome. Of course, social difficulties had the potential to inhibit clarification requests. It was found, however, that the "ASP" children requested clarification more often than did the "SPD" subjects (see Appendix J).

As discussed above, sentences on the tape were repeated when it was apparent to the author that the child had been distracted or when a child requested clarification. Three repeats were recorded for the "SPD" subgroup, all of which followed clarification requests. The "ASP" subgroup required 17 repeats (nine of which were considered necessary by the author and eight of which followed clarification requests). The higher rate of clarification repeats among "ASP" subjects suggests that these children may have had more difficulty discerning the words on tape, remembering whole sentences or processing the connected discourse. The higher rate of repeats initiated by the author would be consistent with a theory that these children had more difficulty sustaining attention.

Interestingly, most of the "SPD", "ASP" and LD children were noted on the Checklist to display "poor turn-taking in conversation (eg. doesn't take his/her turn, does not allow others to have a turn)" (Feature G9).

Eight of the nine repeats of this type were required by just two children: SP22 and SP23.

279
"ASP" results, these hypotheses are not, however, well-supported by the results. Five of the "ASP" repeat items ultimately scored "idiomatic". Of the remaining twelve, eight scored an "inappropriate" narrow code on both tasks (suggesting that these were not false positives) and only four were "inappropriate" on the play task and "appropriate" on the definition task (probably false positives).

On the surface, the significantly greater number of "fuzzy" items scored by the "ASP" than "SPD" subgroup suggests that the former may in fact have had more difficulty in demonstrating comprehension. Division of the "ASP" fuzzy items into the five subtypes outlined above reveals, however, that the majority were of types (a) and (c)\(^1\). It was argued above that these (along with type (b)) are more likely than some others to reflect true idiom difficulty than interference from task factors. When the "SPD" subgroup's "fuzzy" items are similarly broken down into action types, it is evident that only 50% are of type (a), (b) or (c), as opposed to 76.9% of the "ASP" "fuzzy" items. This strengthens the hypothesis that the "ASP" subgroup had more true idiomatic difficulty than did the "SPD" subgroup.

In the light of these findings, the difference between the "ASP" and "SPD" subgroups does indeed appear to stem predominantly (if not entirely) from a true difference in idiomatic comprehension. Two possible explanations for this result may be proposed. The first relates to differential exposure to idiom therapy. Most of the "SPD" children were pupils in language units. As such, many had the benefit of daily sessions with a SLT and contact with teachers who were aware of a tendency to literal interpretation among these pupils. In contrast, most of the "ASP" subjects had less SLT input and were being educated in mainstream schools, often by teachers not specialised in teaching children with learning difficulties. It may have been that the therapeutic intervention received by "ASP" children focused to

\(^1\)Nine out of 26 (34.6\%) were of type (a) (relating to other words in the sentence) and eleven (42.3\%) were of type (c) (unrelated to the surrounding narrative).
a lesser extent on non-literal language than did that of the "SPD" children¹.

A second account of this result derives from the recognition that idiom comprehension is at least as much a pragmatic task as it is semantic. It may therefore be the case that greater impairment on Bishop's (1989) dimension of "interests and social relationships" is of more consequence to idiom comprehension than are abnormalities in "meaningful verbal communication". Bishop does not, herself, explain these terms in detail, but descriptions of "abnormality" with respect to "interests" in Asperger syndrome typically include highly restricted interests and inflexibility of thought (DSM-IV, APA, 1994; ICD-10, WHO, 1992; Gillberg & Gillberg, 1989; Wing, 1981). Since flexibility of thought is essential for idiom comprehension, it is likely that idiom comprehension will be limited in children constrained by rigidity of thought. Interestingly, analysis of the Checklist reveals that the "ASP" children were noted to have rigid concept boundaries (feature G21) substantially more often than were "SPD" subjects ("ASP" 44.4% vs "SPD" 23.5%). It was found, furthermore, that every child who had rigid concept boundaries (in either subgroup) performed poorly on the play task and that no child who scored eight or more (two thirds) idioms "appropriate" had rigid concept boundaries. A positive correlation between rigid concept boundaries and "inappropriate" scores was found to be significant (r_{pb} = 0.536, p = 0.0024). More specifically, significance was also reached for the positive correlation between rigid concept boundaries and "fuzzy" scores (r_{pb} = 0.494, p = 0.0052).

The association between rigid concept boundaries and idiom comprehension difficulty is not surprising. As discussed above, idiomatic interpretation rests on a foundation of semantic duality. A child who attaches a single, rigid meaning to lexical

¹The mainstream "ASP" children might of course have been exposed to more idioms from their peers than would children in language units. The extent to which they may have benefitted from slightly increased exposure is unclear, however.
items is likely to have great difficulty in acquiring and applying understanding of idiomaticity. Rigid concept boundaries might also inhibit the acquisition of individual idiomatic items if these contain words which are associated with a very rigid meaning in the child's mind.

In addition to rigid concept boundaries, the "ASP" and "SPD" Checklists revealed a considerable difference between the two subgroups with regard to prosody (feature G3). As many as 55.6% of "ASP" subjects were noted to have abnormal prosody as compared with only 17.6% of the "SPD" children. A significant positive correlation was found between abnormal prosody and "literal" performance within the SP group ($r_{pb} = 0.372, p = 0.0305$).

While the Checklist feature focused on the child's own prosody rather than on interpretation of prosody, it is likely that the two are correlated. Prosody, as discussed above, probably plays an important role in the recognition of idioms as units of meaning. As such, its impairment - or an impairment in the memory (or other) skills required to use previous prosodic experience - may impact on the learning of idioms and on the interpretation of acquired idioms in context.

A third feature which differentiated the "ASP" and "SPD" subjects was "poor nonverbal communication" (feature G29). In the "ASP" subgroup, 77.8% of subjects scored positive for this feature. This compared with 41.2% of the "SPD" subgroup. Like abnormal prosody, poor nonverbal communication was found to correlate significantly with "literal" scores ($r_{pb} = 0.501, p = 0.0045$). This correlation is interesting given that SP children's learning of individual idioms may be very reliant on adults' detecting of their comprehension failure. Such detection in relation to idioms may be very difficult if a child has poor use of facial expression, unless the idiom is uttered as part of an instruction. Particularly in combination with a reluctance to or difficulty in asking for help, this feature is likely to contribute to poor learning from idiom exposure.
A further feature which partially discriminated the "ASP" and "SPD" subgroups on the Checklist relates to the children's use of unestablished referents (feature G28). This feature was scored positive for 55.6% of the "ASP" subjects as compared with 35.3% of the "SPD" subjects. Of the eleven SP subjects reported to use unestablished referents, eight performed poorly on the play task (scoring "appropriate" for no more than six idioms). The correlation between the use of unestablished referents and "inappropriate" scores ($r_{pb} = 0.617$, $p = 0.0004$) as well as between unestablished referents and "fuzzy" scores ($r_{pb} = 0.52$, $p = 0.0032$) were both significant.

Given that the use of unestablished referents might relate to a theory of mind deficit or to a coherence impairment, its correlation with idiom comprehension is not surprising. As discussed above, theory of mind skills are necessary if a knowledge of idiomaticity is to be applied in context. Without an awareness of central coherence, the recognition of semantic anomaly and the search for a meaningful interpretation of the idiom are inevitably undermined. The selection of idiomatic over literal meaning or over alternative idiomatic senses of an expression is also facilitated by the application of theory of mind and the ability to make use of cohesive devices within discourse.

It has, thus far, been argued that many of the pragmatic features apparently impaired in a large number of "ASP" children play a central role in idiom comprehension in context. Their importance in idiom comprehension is underlined by the finding that most of these features were intact in the LD subjects who performed significantly better on the play task than did the SP group.

Among the other Checklist features on which the SP and LD groups differed substantially was feature G17: paying of attention to

---

1In the LD group, abnormal prosody, rigid concept boundaries and poor non-verbal communication were noted for 13%, 7% and 7% of subjects, respectively. This compared with 31%, 31% and 54%, respectively, of the SP subjects who scored positive for these features.

283
irrelevant, non-salient details. A weak but significant negative correlation was found between this feature and SP "appropriate" items \( (r_{pb} = -0.355, p = 0.0375) \). Interestingly, this feature was also associated with SP "ambiguous-idiomatic/literal" items \( (r_{pb} = 0.382, p = 0.0272) \). This suggests that some SP children may have had difficulty using the overall context to guide their selection between idiomatic and literal interpretations.

While more "ASP" than "SPD" subjects were reported to have abnormality in features important to idiom comprehension, many of these features were more often reported for "SPD" than for LD subjects. This is consistent with the finding that the "ASP" subgroup performed worst on the play task, followed by the "SPD" subgroup, and then by the LD group.

Interestingly, some Checklist features which were correlated with "literal" and "fuzzy" scores occurred with similar frequency in the SP and LD groups. The use of unestablished referents (feature G28) was recorded for 42% of SP subjects and 40% of LD subjects. Failure to request clarification (feature G12) was noted for 50% of SP subjects and 53% of LD subjects\(^1\). As discussed above, use of unestablished referents may reflect impaired theory of mind or coherence, both of which are central to the use of context to guide idiom comprehension. Failure to request clarification may limit a child's learning of new idioms. Both of these features may partially explain the weakness of the LD group's performance relative to the mainstream children.

While variability was evident within the LD group with respect to the use of unestablished referents and clarification requests, there was virtually no variation on the other features correlated with SP play task performance: abnormal prosody, rigid concept boundaries and poor non-verbal communication (almost all LD

\[ ^1 \text{This feature correlated significantly with SP "inappropriate" items (} r_{pb} = 0.457, p = 0.0095\), "literal" items (} r_{pb} = 0.355, p = 0.0374\) and "fuzzy" items (} r_{pb} = 0.379, p = 0.0280\). It did not discriminate "SPD" (47.1%) and "ASP" (55.6%) subgroups.\]
subjects scored negative for these features\(^1\)). The lack of variance on these skills may well explain why the LD group variance on the play task was very similar to that of the mainstream groups, despite its being a heterogeneous group (with respect to age and symptomatology). It would seem that most of the dimensions on which the LD subjects did vary do not play an important role in idiom comprehension.

In contrast, many of the features on which the SP group varied\(^2\) are fundamental to idiom comprehension or relate to it via a common underlying mechanism. Their greater representation among "ASP" than "SPD" subjects probably explains why the "ASP" subgroup performed significantly worse on the play task than did the "SPD" subgroup. While abnormality in a range of "critical" skills does appear to correlate with poor performance and to be more common among "ASP" children, the picture is not entirely clearcut. Both subgroups included children who were noted as having problems in these "critical" skills and who performed poorly, as well as children who were not recorded to have difficulty in these skill areas and did relatively well on the play task\(^3\). Indeed, the standard deviations for "appropriate" and "inappropriate" scores in both subgroups were considerably larger than those in the mainstream and LD groups.

It thus appears that play task performance within the SP group cannot be accounted for entirely by either diagnosis or age, individually. It may be that the "critical" skills are in fact influenced by age, possibly tending to reduce in severity over time, as has been described for some other language features associated with semantic-pragmatic difficulties and the autistic continuum (Bishop, 1989; Frith, 1991; Sahlén & Nettelbladt, 1993;)

\(^1\) Of the 15 LD subjects, 13 scored negative for abnormal prosody and 14 scored negative for rigid concept boundaries and poor non-verbal communication, respectively.

\(^2\) Many of the other features on which the SP group varied considerably did not appear to affect idiom comprehension significantly (see Appendix C and Table 63).

\(^3\) Abnormality in "critical skills" did not match perfectly with poor idiom performance on the play task although it did predict it to a significant degree.
The overall picture that emerges with regard to idiom comprehension in the SP group is one of considerable variability. Not only do these children vary in their overall play task scores, but the evidence suggests that they vary too with respect to the sites and sources of their idiom difficulty. The following consideration of two individual cases illustrates the importance of drawing on all sources of information in order to establish an idiom ability profile for individual children considered to have semantic-pragmatic difficulties.

In the case of SP3 (aged 10;1, diagnosed with Asperger syndrome/high-functioning autism), it is clear that the child has some grasp of idiomaticity. On the play task, this child scored "idiomatic" for five idioms ("took off", "got on to", "kept an eye on", "dropped in" and "went round"). A sixth idiom ("ran over") was scored "ambiguous-idiomatic/literal" but was subsequently defined accurately (see between-idioms discussion with regard to the high levels of ambiguity associated with this idiom). Of the remaining six idioms, three scored "literal" or "inferred literal" and three scored "fuzzy". One "inferred literal" and one "fuzzy" item subsequently scored "inferred idiomatic" on the definition task. It appeared as though the child may have had difficulty processing these on the play task.

The picture that presents on the idiom assessment is of a child who has the knowledge that words may have more than one meaning and that what is said may not always be the same as what the speaker intends to convey. Nevertheless, this knowledge is not always applied in practice. His literal interpretations may have arisen from a failure to notice the incongruity of the literal meaning in context, to retrieve the idiomatic meaning and/or to select the appropriate meaning for the particular context. Alternatively, this child may simply not have known an idiomatic meaning for these idioms and may thus have opted to accept their literal meanings. In the case of the "fuzzy" items, it is
perhaps more likely that the child did realise that the literal meaning could not be what was meant, but did not know any other meaning for the expressions.

Examination of the child's Checklist profile supports both the theory of a limited idiom repertoire and of difficulty in mobilising knowledge of idiomaticity in context. General passivity in conversation along with a tendency to stick to obsessive topics in conversation probably reduce this child's exposure to new idioms in the social context. If his abnormal prosody extends to difficulty in interpreting prosody, this may reduce his awareness of having encountered a unit of language when he is exposed to idioms. His tendency to attend to irrelevant details and probable difficulties in using some cohesive devices (unestablished referents) would greatly reduce this child's potential for inferring meaning from context. A noted tendency not to ask for clarification when failing to understand, particularly in combination with poor non-verbal communication, would further reduce the likelihood of others detecting his comprehension failure. It is possible that this child does not, himself, recognise his comprehension failures. Even when explicitly taught idioms, this child's apparently rigid concept boundaries might hinder his grasp of some idioms.

General inflexibility of thought (manifested in restricted and obsessive topics of interest, repetitive, uncreative play and rigid concept boundaries) might also interfere with the application of acquired idiomatic knowledge in context. The use of context to guide comprehension would also be hindered by attention to irrelevant details rather than to topic, difficulty with overall coherence (unestablished referents) and impaired ability to consider situations from another's point of view (lacks awareness of what his conversational partner can be expected to know and uses unestablished referents).

Thus, there is reason to suspect that the idiom comprehension difficulty displayed by this child reflects breakdown at several
points in the process of idiom acquisition and interpretation.

While only a month older than SP3 and scoring almost the same number of "idiomatic" items, SP19\(^1\) appears to have more limited difficulty with idiom comprehension. His six "idiomatic" scores indicate a grasp of the fundamentals of idiomaticity. A further "fuzzy" item was subsequently found to be well understood on the definition task. A second "fuzzy" action entirely unrelated to the narrative may reflect lack of familiarity with the idiom in the context of recognition that the literal meaning does not make sense. Of interest in this child’s performance is a high number of "ambiguous-idiomatic/literal" actions (four out of twelve as compared with a mean of 1.69 in the SP group).

Of the four "ambiguous" items, three involved a demonstration of both the idiomatic and the literal meaning (as opposed to one action which could have been interpreted as an idiomatic or a literal demonstration). This pattern suggests that SP19 has adequate semantic knowledge to interpret most of these idioms, but may lack the necessary pragmatic skills to choose between the idiomatic and literal meanings in context. Support for this interpretation derives from the Checklist profile. That he "gives more information than is necessary" and "seems to lack awareness of what (his) conversational partner can or can't be expected to know" indicates that this child may have a general theory of mind deficit. His utilisation of the context might therefore have been limited. It is further interesting that this child "pays attention to irrelevant, non-salient details", since a failure to attend to the whole narrative would reduce his chance of recognising contextual anomaly and of using the context to guide his selection of the idiomatic meaning. The acquisition of world knowledge by this child might also be restricted by his "(sticking) obsessively to one/few conversational topics".

A comparison of SP3 and SP19 illustrates a probable difference

\(^1\)SP19 was diagnosed as having "autism / query Pathological Demand Avoidance Syndrome".
in site of breakdown and source of difficulty with regard to idiom comprehension. If the proposed explanations of these children's difficulties were supported by further knowledge of the children, the emphasis of therapeutic intervention might differ despite the superficial similarity in their "idiomatic" scores. In both cases, the mobilisation of learned idiom knowledge in context would need to be addressed. In the case of SP3, however, a programme of explicit idiom teaching may also be desirable.

Critically, the results of this study indicate the need for an individual approach to characterising idiom comprehension ability and disability in children with semantic-pragmatic difficulties. In the 26 SP children included in this study, there was no evidence of complete or even predominant literality. Some children, such as SP3, demonstrated probable difficulty in learning idioms and in interpreting them in context. Others, like SP19, appeared to have more limited difficulty in using context to guide the selection of idiomatic meaning over literal meaning. In a small number of cases, there was no evidence of any idiom comprehension difficulty. SP16 and SP17 each scored ten "idiomatic" items, one "inferred idiomatic" and one "ambiguous-idiomatic/literal". SP25 scored "idiomatic" for nine idioms and "inferred idiomatic" for the remaining three. All of these children were aged 10;6 or older and were noted as having abnormalities in few of the "critical" skills.

In considering the performance of such children as SP16, SP17 and SP25, we need to bear in mind that these results cannot necessarily be generalized. The children's apparent ability with respect to twelve common, concrete idioms in this structured, visually-supportive task may mask greater difficulty with less commonly-used and/or more abstract idioms or even with the same idioms in a less facilitative natural context. It may also be that the SP children had a shallower understanding of these

\[1\] In both cases, the ambiguity arose from an action which could have been interpreted as either "idiomatic" or "literal" rather than from the child depicting both meanings.
idioms than did children in the other three groups. It is further possible that these SP children had minor idiom comprehension difficulties in their younger years, but that these have been overcome as their underlying problems reduced in severity naturally and/or as a result of therapeutic intervention.

While certainly we need to be cautious in generalizing from these results, the success of at least some SP children on the play task indicates that the overarching characterisation of children with semantic-pragmatic difficulties in terms of idiom comprehension difficulty may be inaccurate. It is almost certainly the case that children with semantic-pragmatic difficulties are heterogeneous with regard to idiom comprehension. It is probable that this heterogeneity reflects diverse underlying problems. As stressed by such authors as Fey and Leonard (1983), McTear and Conti-Ramsden (1992) and Prutting and Kirchner (1983, 1987), it is quite plausible that subgroups of pragmatic impairments will emerge as our insight into this group of children (and adults) increases. The results of this study support the hypothesis that the population of individuals considered to have semantic-pragmatic difficulties is a heterogeneous grouping whose symptoms - even if superficially similar - may be underpinned by different problems. The relationship between Asperger syndrome and "semantic-pragmatic disorder / difficulties" that emerges here is largely consistent with that hypothesized by Leinonen (1991): -

The common thread between (autism, Asperger syndrome and "semantic-pragmatic disorder") may be a problem of acquisition or functioning of pragmatic knowledge; the differences may lie in which aspects of this knowledge have failed to be acquired or is (sic) not functioning adequately (p.111).

---

1 As discussed above, we might expect that children build up layers of semantic and pragmatic knowledge about an idiom over time.
9. BETWEEN-IDIOMS DIFFERENCES

As noted in Chapter 2, considerable degrees of between-idioms difference are frequently reported in studies of idiom comprehension. The results of this study are in keeping with this pattern. On the play task, the proportion of subjects scoring "appropriate" for individual idioms ranged from 46.7% to 100% in the MJ and MI groups, from 40% to 100% in the LD group and from 26.9% to 96.2% in the SP group. Notably, at least 80% of the MJ subjects scored "appropriate" for all but three idioms ("pick up", "ran over" and "on the spot"). In the other three groups, a greater degree of between-idioms difference was evident.

While almost all the idioms were associated with smaller proportions of "appropriate" scores in the SP group than in the other three groups, there was only one idiom ("moved up") on which the SP group scored particularly few "appropriate" scores relative to the other three groups on the play task. In the case of "moved up", only 26.9% of SP subjects scored "appropriate". This compared with 60% of the LD and MI groups, respectively, and 80% of the MJ group.

In most cases, SP children who did not score an "appropriate" narrow code for this idiom scored "literal", "inferred literal" or "fuzzy". It is possible that the use of a pronoun in this sentence ("The vet moved them up so that there was enough space for Rover") created difficulty for some of the SP children. It may also be that the occurrence of the idiom early on in a long sentence caused some children to forget the idiom and to respond instead to the words "enough space for Rover". This is supported by the finding that many SP children demonstrated this item by pushing all the bandages off the table and lifting Rover onto the table. This action might also, however, be interpreted as a demonstration of a different expression ("moved off") or as

---

1 Wider ranges were recorded on the definition task. The MJ group ranged from 40% to 100%, the MI group ranged from 13.3% to 100%, the LD group ranged from 6.7% to 80% and the SP group ranged from 3.9% to 88.5%.
relating to the literal meaning of "moved up" (since Rover was lifted from the floor onto the table). Since it was not clear that the children in these instances were intending to demonstrate the idiomatic meaning of "moved up", these actions were coded "fuzzy". As noted above, "fuzzy" actions of type (d) cannot necessarily be interpreted as reflective of true idiom comprehension difficulty. As such, it would be unwise to draw any conclusions from the between-groups discrepancy found for this idiom.

While the proportion of children within each group scoring "appropriate" for individual idioms differed, it is interesting that the rank positions of each idiom relative to the others (in terms of "appropriate" scores) correlated significantly across all four groups (Kendall's $W = 0.456$, $p = 0.0009$). In other words, the difficulty of any given idiom relative to any other was very similar for the MJ, MI, LD and SP groups, respectively.

This result suggests that some idioms were inherently more difficult than were others for children of this age range to comprehend and/or to demonstrate. On the play task, five idioms ranked consistently in the top six for each of the four groups. These were "took off", "gave a hand", "got on to", "dropped in", and "went round". The six idioms ranked lowest for each group consistently included "ran over", "moved up", "on the spot", "turned over", and "pick up". "Under lock and key" ranked in the top six for all but the MI group, for whom it ranked seventh. "Kept an eye on" ranked at between sixth and eighth place (out of twelve) for all the groups.

While between-idioms variability is not in itself remarkable, the factors underlying this consistent pattern may hold important information with regard to early idiom acquisition and processing. It was noted in Chapters 2 and 4 that the idioms...
included in this study differed on at least two dimensions. Firstly, seven of the idioms had clearly literal counterparts whereas the counterparts of the other five idioms lay on the border between literal and idiomatic. Secondly, the idioms in this study encompassed several syntactic structures: phrasal verb idioms (e.g. "ran over"), clause idioms (e.g. "gave a hand") and adverbial phrase idioms (e.g. "under lock and key").

Viewed superficially, the rank positions of the twelve idioms appear to bear no relation to the status of these idioms on either of these two dimensions. Closer examination reveals, however, that the syntactic behaviour of the phrasal verb idioms ranked consistently in the top six differs from that of those in the bottom six. Firstly, the phrasal verb idioms in the bottom six ("ran over", "moved up", "turned over", "pick up") are all transitive verbs. Secondly, they all allow their object to be inserted between the verb and particle (e.g. "moved them up", "turned the bike rider over", "pick it up"). In contrast, the phrasal verb idioms best understood were either intransitive verbs ("took off", "dropped in"), or were transitive but did not permit object insertion ("got on to", "went round"). It is thus apparent that the phrasal verb idioms better understood were all more frozen syntactically than were those that were less well understood. This result is consistent with that of Gibbs (1987) who found that frozen idioms presented in context were better understood by 5- and 6-year old children than were more flexible idioms. Interestingly, the frozenness factor appeared to make no significant difference to the idiom comprehension of Gibbs' 8- and 9-year old children. In this study, the correspondence between relative frozenness and idiom comprehension appears to hold right across the age range 6;6 to 11;6.

In addition to their relative syntactic flexibility, idioms also vary with respect to lexical flexibility i.e. the extent to which

---

1 "ran over", "gave a hand", "moved up", "on the spot", "under lock and key", "kept an eye on", "dropped in"

2 "took off", "got on to", "turned over", "pick up" and "went round"
they will allow lexical substitution (Gibbs et al., 1989a). Several idioms included in this study allow a small degree of lexical substitution: "give/lend a hand", "move/shift up", "drop/pop/call in". There is, however, no clear correspondence between lexical flexibility and relative comprehension in these data.

As outlined in Chapter 2, idioms can be further differentiated in terms of semantic decomposition (Gibbs & Gonzales, 1985; Gibbs & Nayak, 1989; Gibbs et al., 1989a, 1989b; Nunberg, 1978 cited in Gibbs & Nayak, 1989) and metaphorical transparency / opacity. As these concepts relate primarily to clause idioms, of which only two are represented in this set, the results of this study can shed little light on the impact of these two dimensions.

With the exception of syntactic behaviour, the inherent properties of these idioms are limited in their ability to explain the pattern of between-idioms differences. It may be that all or some of the dimensions discussed above did contribute to the outcome of these results, but that their effect is interactive rather than individual. It is further possible that these and other factors work indirectly to affect comprehension. For example, syntactically frozen idioms may be better understood as a result of their being more familiar. This, in turn, may stem from their being heard in only one, or very few, syntactic forms (Gibbs & Gonzales, 1985).

Alternatively, the inferior performance associated with syntactically flexible idioms might not have stemmed from their relative familiarity but rather from their use in this context in a disrupted form. Three of the four phrasal verb idioms which allowed object insertion were in fact presented in this form in the story: "moved them up", "turned the bike rider over", "pick it up". It is possible that the splitting of the idiom in this manner made it more difficult to recognise the unity of these expressions. Interestingly, two of these three idioms ("moved up" and "turned over") were scored "appropriate" by considerably
more MJ subjects than by MI, LD or SP subjects. "Moved up" was scored "appropriate" by 80% of the MJ group as opposed to 60% of the MI and LD groups, respectively, and only 26.9% of the SP group. Similarly, "turned over" was scored "appropriate" by 80% of the MJ group, 53.3% of the MI group, 46.7% of the LD group and only 38.5% of the SP group. Coupled with the fact that all three of these idioms ranked in the bottom four for the SP group, this pattern of results suggests that SP children may have particular difficulty in recognising the unity of expressions when they are presented in disrupted form. MI and LD children may have had less difficulty than the SP children, but nevertheless scored less well for these idioms than did the older mainstream children.

In addition to exploring between-idioms differences in terms of relative levels of "appropriate" scores, it is further informative to consider the relationship of idioms with respect to the other broad codes. It is evident, for example, that three idioms were associated with a high number of "ambiguous-idiomatic/literal" scores: 40 for "ran over", 27 for "pick up" and 18 for "turned over" (each out of a possible 71). In the cases of "pick up" and "turned over", a range of codes other than "ambiguous" were also scored. "Ran over", however, was scored "idiomatic" or "inferred idiomatic" by every child who did not score "ambiguous-idiomatic/literal". Examination of the data reveals that many of the children who scored "ambiguous-idiomatic/literal" for "ran over" did in fact understand the idiom. Each group scored more "appropriate" items for this idiom on the definition task than on the play task. The high level of ambiguity associated with this idiom stemmed largely from the difficulty of coding a demonstration of "ran over". Many children depicted this idiom by knocking Rover to lying position with the front wheel of the bike and then driving the bike onto or across Rover's body. Some other children hit Rover with the front wheel and then moved the bike in a soaring motion over the dog. Certainly these demonstrations are consistent with an idiomatic interpretation. However, the action of moving the bike
rider over Rover's body might also be argued to relate to the literal interpretation. (This possibility was supported by the fact that some children later demonstrated the bike rider running without removing him from the bike.) The result was a large number of "ambiguous-idiomatic/literal" scores for this idiom, many of which were false positives.

Similar coding difficulties were noted to a lesser degree with respect to "turned over". In some instances, children made Sally turn the bike rider around to face the policeman when demonstrating the sentence, "Sally turned the bike rider over to the policeman." Unlike the cases in which the bike rider was rotated entirely unnecessarily (coded "inferred literal"), it was impossible to be certain in these instances whether the child was demonstrating the idiomatic or "inferred literal" sense of "turned over".

As noted above, difficulties associated with coding play data occasionally accounted for false positive "fuzzy" scores too. In four out of the five instances in which a "fuzzy" code was assigned to "kept an eye on", it was subsequently found that the children did know the idiom. These were the only examples of "fuzzy" action type (e) in which the props were already positioned to reflect the idiomatic meaning. The children scored "fuzzy" because they acted out some other action with other props. This may have stemmed from a feeling of obligation to respond to the sentence in some way.

When the idiom differences are considered from the perspective of relative literality, one in particular stands out from the eleven others. Not only did "on the spot" attract more "literal" codes among the SP group than did any other idiom, but it also

---

1Since this coding difficulty affected all four groups to a very similar degree, this factor is most unlikely to have affected the between-group results.

2Although all three of these cases involved SP children, the fact that only three out of 852 SP items were affected indicates that this factor could not have produced the significant discrepancy between the SP group and the other three groups.

296
attracted a high number of "literal" and "inferred literal" codes within each of the mainstream groups. Indeed this idiom accounts for six out of a total of nine MI items coded "literal" or "inferred literal" on the play task, and seven out of the eight MJ items coded "literal" or "inferred literal". Overall, 23 of a total of 71 subjects scored "literal" or "inferred literal" for this idiom.

In exploring this result, it is important to note that all but five of these 23 "literal" and "inferred literal" "on the spot" items also scored an "inappropriate" narrow code on the definition task (i.e. "literal", "inferred literal", "fuzzy", "narrative" or "repeat-narrative"). The consistently "inappropriate" results associated with this expression suggest that this idiom was either inherently more difficult than the others for children of this age group to acquire or was more difficult than other idioms to retrieve in this context.

In terms of its inherent properties, "on the spot" is notable for being one of two idioms which function as adverbs. It is difficult to see how this factor might have given rise to a high level of literality, however. The other adverbial phrase idiom, "under lock and key", did not attract many "literal" or "inferred literal" codes.

Aside from its syntactic function, "on the spot" stands out from the other idioms by virtue of having many layers of closely related meaning. In addition to its target idiomatic sense (immediately, without delay) and its literal sense (on the physical dot), the expression can be interpreted in three further senses. The first interpretation is "there at the time and place an event happened" as in "ITN had a man on the spot interviewing ...paratroopers as they dropped from the sky" (ODCIE Vol. 2, Cowie et al., 1985, 435). The second is that of performing some activity in a single place, as in "running on the spot". Thirdly, the figurative sense of "spot" as "a particular place (or) definite locality" (The Concise Oxford Dictionary, Sykes,
1982, 1027) could give rise to a child placing the bike rider on the accident spot (as some children did). As discussed in Chapter 2, it may be that the process of selecting a meaning in context is more complicated for expressions with multiple meanings. Indeed the results support an explanation based on confusion between figurative senses. In addition to the high level of literality, this idiom also attracted more "other figurative sense" codes than all of other idioms, together\(^1\). The many layers of meaning, all of which relate to "a place", might have made this a difficult idiom even for 10 and 11-year old mainstream children to learn and retrieve in context.

In addition to the close relationship between the different meanings of "on the spot", the frequency with which these meanings occur in everyday discourse is, impressionistically, fairly equal. In this sense, "on the spot" differs from such idioms as "gave a hand" and "under lock and key", which are far more likely to be heard in their idiomatic senses than in their literal senses. On the basis of their results with 5- to 12-year old children, Brinton et al., 1985 hypothesize that uncertain subjects may be swayed towards selecting the idiomatic sense of expressions with a higher likelihood of idiomatic than literal occurrence. A similar bias in adults is reported by Van Lancker and Canter (1981). Since the target sense of "on the spot" is probably not heard more frequently than any of the other meanings of the expression, this priming effect would not be likely to occur for this idiom.

As discussed above, the context within which an idiom occurs also has a strong potential for setting up an idiomatic interpretation. Like that of all the other idioms, the narrative context for "on the spot" supported its target idiomatic interpretation\(^2\). Interestingly, however, the contextual bias

\(^1\) "On the spot" attracted the only "other figurative sense" score on the play task and nine out of the 13 "other figurative sense" codes scored on the definition task. These were spread across all four groups.

\(^2\) The story, heard in its entirety just before the play began, reads, "The bike rider started running. Sally chased him. Suddenly the bike rider stopped on the spot. There was a policeman in the road."
for this idiom was arguably weaker than that for many of the other idioms. It may be that the pragmatic skills necessary in order to select one of several closely related figurative meanings in a less strongly biasing context exceed those possessed by many 10- and 11-year olds. The MJ subjects might have had the semantic knowledge required for interpreting "on the spot", but inadequately developed pragmatic skills to select the correct meaning in this context. This theory is consistent with the findings of Strand and Fraser (1979) and Levorato (1993). Strand and Fraser note that even 11-year old normally-developing children still switch between literal and idiomatic strategies when interpreting different idioms. On the basis of a study including fourth grade (9- to 10-year old) normally-developing children, Levorato concludes that the fact that they could still be fooled by (a) literal context into choosing a not entirely appropriate literal response implies that they had not completely acquired and systematically applied the figurative strategy, especially in the case of unfamiliar idioms.... We assume that when figurative competence is fully developed, the listeners choose the figurative interpretation whenever possible, basing the choice on their belief that whoever produced the message would not have intended to create any ambiguity and would not have meant a figurative expression to be interpreted literally (1993, 115).

While the particular pragmatic demands of this idiom are able to account for the high level of literality associated with it, it is notable that two of the seven MJ children who scored "literal" on the play task subsequently produced an "appropriate" definition. Given that the idiom was heard twice more in context before the definition, it may be that these children succeeded in working out the target meaning by the time they came to define it. The discrepancy might, however, also be argued to relate to the nature of the play set. Unlike any of the other idioms, the literal demonstration of "on the spot" involved the use of a prop not otherwise used during the play task: a little green dot

---

1 Levorato's "literal" context was one in which a literal interpretation was allowed but a figurative interpretation was more appropriate. It therefore resembles the context of "on the spot".
placed beside the road, close to the vet's office. It is arguable that some of the children may have seen the green dot and wondered why it had been placed there. It is possible that some of these children believed that the tester intended them to make use of the green dot and were therefore misled into incorporating it into their demonstration of "on the spot".

Clearly, it is impossible to know whether and to what extent any children attempted to outwit the researcher or to conform with what they believed to be her wish. Nevertheless, it is worth considering what the effect may have been if large numbers of children were in fact misled into producing false positives.

Considering the possible effect on the SP group, first of all, it is notable that any false positives among their results would have led to an underestimation of SP idiom comprehension. In other words, the SP children would have emerged with an even higher proportion of "appropriate" items than that reported in the results of this study. Thus, the conclusion that the SP group demonstrated significantly more items "appropriately" than "inappropriately" would be strengthened.

If we assume that false positives did result on the play task and, by extension, on the definition task, what would the effect have been on the group differences? At its most powerful, this factor might have led to all instances of MJ, MI and LD "literal" codes for "on the spot" being false positives. If this were true (and it is unlikely that it is), the effect would have been to reduce the difference between the SP group and the MJ, MI and LD groups, respectively, with regard to literality and overall "inappropriacy". In order that the effect of this "worst case scenario" on the conclusions of this study could be assessed, Mann-Whitney Test statistics were calculated to examine the differences between groups on the codes "literal", "inferred literal" and "inappropriate" when the data for "on the spot" is excluded. These between-groups differences (for both tasks) were then compared with those computed for the full data set.
On the play task, the comparison of group differences on each of the codes "literal", "inferred literal" and "inappropriate" revealed the same significance status for all but one pairwise difference. The difference between the SP and MJ groups for the code "literal" was not significant when the full data set was included, but did reach significance when "on the spot" was removed (p = 0.0147). No other changes in significance status emerge, but the probability values for group differences were almost always lower on the amended data set than on the full data set. On the definition task, the significance status is not different for any group across the two sets of data. As on the play task, however, the probability differences are mostly smaller for the amended data set than for the full data set.

It can therefore be concluded that the "worst case scenario" would make the differences between groups on the codes "literal", "inferred literal" and "inappropriate" appear smaller than they really were. The conclusion that the SP group is not significantly more literal than any of the three other groups would have to be reduced to the assertion that the SP group was not significantly more literal than the MI or LD groups. It would still be legitimate to conclude that the SP group was significantly more "inappropriate" than were each of the other three groups. Given that the absolute incidence of "literal" and "inferred literal" scores in each group was very small anyway, a "green dot factor", even at its most powerful, could not have exerted a great impact on the outcome of this study.

Beyond the immediate context of the play task, other factors relating to idiom usage may partially account for the between-idioms differences found in this study. As noted above, idioms that more frequently occur in their figurative than literal senses may be more likely to be interpreted idiomatically. To some extent, this was borne out in this study. "Gave a hand", "under lock and key" and "kept and eye on" are all, impressionistically, far more likely to be used idiomatically
than literally\(^1\). Each of these idioms was scored "appropriate" by at least two-thirds of each group. It is, however, difficult to draw any conclusions from this finding. Firstly, idioms which are as likely to be encountered literally as idiomatically were also well understood. "Took off", "dropped in" and "went round", for example, were also scored "appropriate" by over two-thirds of each group\(^2\). Secondly, the relatively poorer performance of some other idioms may have stemmed from their not being known to these children, rather than from their being equally likely to occur literally as idiomatically. Thirdly, the feasibility of an account based on children's use of previous language experience to predict probable meaning is questionable with regard to the SP group. In this group at least, it is unlikely that this pattern of between-idioms differences derived from probability judgements.

A more plausible account of the relationship between idioms derives from a relative familiarity perspective. It may simply be the case that these and other idioms on which the subjects all performed well were more familiar to these children than were those on which they performed less well. This hypothesis is consistent with the conclusions of Strand and Fraser (1979) that the idioms acquired earliest by young children are those which are used in interaction with young children and are useful to them. This proposal fails, however, to explain the consistency across groups in the relationship between idioms, since it cannot be that all children acquire idioms in the same order.

The likeliest explanation of these between-idioms results is that the particular structural and contextual features associated with each of these idioms combined to produce this pattern of performance. Given that phrasal verb idioms are evident in very young children's language (as illustrated above) and that they

\(^1\)"Ran over" is also more likely to occur idiomatically than literally. Although it did not rank very highly on the play task, it may nevertheless have been well understood, as discussed above.

\(^2\)These are commonly used literally, as in "he took the book off the table", "she dropped the soap in the bath", and "the thoroughfare was closed so she went round the building".
seem to be used in interaction with children (at least in the classroom) more often than are other structural types, it is likely that these are the first idioms acquired by young children. Thus far, this is consistent with phrasal verb idioms having been among those best demonstrated on the play task. Why some phrasal verb idioms were consistently well demonstrated whereas others attracted fewer "appropriate" scores may be explained by the fact that the former were all frozen idioms, were all presented in an uninterrupted syntactic form and were not subject to coding difficulty. The clause and adverbial phrase idioms included were probably among those frequently used in interaction with young children. As such, they too stood a good chance of being understood by many of these children in each group. Those that were presented within a strong idiomatic context, were more likely to be encountered idiomatically than literally and were not subject to coding difficulties, were most likely to attract a high number of "appropriate" scores within each group. Meeting all of these criteria, "gave a hand" emerged as the best understood non-phrasal verb idiom in three of the four groups. In contrast, "on the spot" failed to meet two of these three criteria and thus consistently attracted fewest "appropriate" scores of all the non-phrasal verb idioms.

In the absence of detailed reports in the literature of between-idioms differences in context, it is not possible to validate this theory of idiom differences. Future research might further explore the extent to which between-idioms differences are consistent across groups of children and confirm or refute the

---

1 Of course, children are exposed to idioms in many other settings as well. As argued above, however, it is more likely that parents and carers use a greater number of phrasal verb idioms than other types in interaction with young children.

2 Although children are exposed to a great number of clause idioms and adverbial phrase idioms on television, this is arguably a poorer idiom learning environment. In the classroom, the child is able to ask the meaning or indicate lack of comprehension non-verbally and thereby obtain an explicit definition. Unless the child is watching with someone willing and able to explain an unfamiliar idiom, the only channel of learning from television is inference. Of course, once an idiom has been acquired or its meaning inferred, encountering it on television will provide an opportunity for the reinforcement of learning.

3 Strand and Fraser (1979) do present group results for each of their idioms. Unfortunately, these idioms were presented to the subjects in isolation and therefore cannot be considered in terms of contextual variables.
10. BETWEEN-TASKS DIFFERENCES

Between-tasks difference results were found to be highly consistent with the predictions of this study. All four groups scored more "appropriate" items and fewer "inappropriate" items on the play task than on the definition task. With the exception of the MJ "appropriate" scores, these differences were all significant.

The discrepancy was smallest for the MJ group who scored "appropriate" for 81.7% of their play task items and for 75.6% of their definition task items (difference n.s). Their "inappropriate" score rose from 8.3% on the play task to 15% on the definition task (p = 0.0486). A larger discrepancy was found for the MI group who scored "appropriate" for 76.1% on the play task and 62.8% on the definition task (p = 0.0299). On the play task, the MI group scored 9.4% "inappropriate" as opposed to 31.1% on the definition task (p = 0.0009). Greater still was the between-tasks discrepancy of the SP group. On the play task, this group scored 60.9% "appropriate", but this fell to 48.2% on the definition task (p = 0.0023). Their "inappropriate" score rose from 25% on the play task to 42.4% on the definition task (p = 0.0001). The largest discrepancy of all was that associated with the LD group. Having scored 74.4% "appropriate" on the play task, they went on to score only 46.7% "appropriate" on the definition task (p = 0.0005). Their "inappropriate" scores increased from 10% on the play task to as much as 45% on the definition task (p = 0.0004). Both of these differences are highly significant.

Clearly, these results support the hypotheses that "the number of "appropriate" items scored by any group on the play task (would) be greater than the number scored by that group on the
definition task" (Hypothesis (xi)), and that "the number of "inappropriate" items scored by any group on the play task would be smaller than the number scored by that group on the definition task" (Hypothesis (xii)). Furthermore, as predicted\(^1\), the MJ, MI and LD groups all scored significantly more "appropriate" than "inappropriate" items on the play task, but this gap narrowed for all the groups on the definition task. Indeed, for the LD and SP groups, respectively, there was barely any difference between "appropriate" and "inappropriate" scores on the definition task (LD "appropriate" vs "inappropriate": play task 74.4% vs 10%, \(p = 0.0004\); definition task 46.7% vs 45%, n.s.; SP "appropriate" vs "inappropriate": play task 60.9% vs 25%, \(p = 0.0001\); definition task 48.2% vs 42.4%, n.s.).

Among the mainstream children, higher "inappropriate" scores on the definition task did not generally stem from the groups' scoring more items "literal", "inferred literal" or "fuzzy". (The MI "inferred literal" score did increase but this difference did not reach significance.) Their increase in "inappropriate" scores derived chiefly from "repeat", "narrative" and "repeat-narrative" items (MJ: 0.6%, 3.9% and 5.6%, respectively; MI: 1.7%, 1.1% and 14.4%, respectively). The definition task codes "repeat", "narrative" and "repeat-narrative" do in fact equate to "fuzzy" action types (a) and (b) on the play task\(^2\). Thus, these response types are not peculiar to the definition task. However, the proportion of play task items that corresponded to "fuzzy" action types (a) and (b) was very much smaller than the proportion of items scored "repeat", "narrative" and "repeat-narrative" on the definition task. Coupled with the finding that the number of "repeat-narrative" items scored by the MI group was significantly greater than that scored by the MJ group (\(p =

\(^1\)Hypothesis (xiii): "The MI and MJ groups will both score more "appropriate" items than "inappropriate" items on both tasks."

Hypothesis (xiv): "The LD group will score more "appropriate" items than "inappropriate" items on the play task."

\(^2\)Fuzzy actions of type (a) represented part of the sentence but failed to demonstrate the idiomatic, literal or any other figurative sense of the expression. Type (b) actions represented a part of the ensuing narrative but failed to demonstrate the idiomatic, literal or any other figurative sense of the expression.
0.0120), this suggests that these responses reflected difficulty with the act of defining, itself.

Indeed, Litowitz (1977) includes repetition of the stimulus item as characteristic of immature definitional skill. Like many other authors, Litowitz observes that the quality of children's definitions improves gradually in line with linguistic and metalinguistic development (Johnson & Anglin, 1995; McGhee-Bidlak, 1991; Snow, 1990; Storck & Looft, 1973 all cited in Nippold, 1995; Wehren, DeLisi & Arnold, 1981 cited in Menyuk, 1988). Johnson and Anglin demonstrate further that some items are more difficult to define than others. In particular, the mastery of verb definition appears to occur later than that of noun definition. Notably, most of the idioms in this study function as verbs. To this author's knowledge, Johnson & Anglin are the only authors who have included idioms in their definition sample. The authors note that the majority of their grade 1 (mean age 6;8) subjects were unable to demonstrate knowledge of any idioms expressively. The 6- and 7-year old MI children in the present study proved themselves capable of producing many adequate definitions. The fall-off in their "appropriate" scores and the evidence of immature definitional skills in their data are however, consistent with the conclusion that definition tasks underestimate idiom comprehension in normally-developing children (Ackerman, 1982; Cacciari & Levorato, 1989; Gibbs, 1987; Nippold, 1995; Nippold & Rudzinski, 1993; Nippold & Taylor, 1995; Prinz, 1983).

Even greater difficulty with paraphrase tasks has been reported for SLI children. Hoar (1977 cited in Fey & Leonard, 1983) found that SLI children in grades 1 to 7 (6 to 12 years old) produced more errors than normal-language peers on a paraphrase task. In contrast to their normal peers, the SLI children frequently repeated the sentences, inverted sentence constituents

---

1 Other manifestations of immature definitional skills among the MI children and, to a lesser extent, among the MJ children, included very concrete definitions closely tied to personal experience (Litowitz, 1977; Johnson & Anglin, 1995).
inappropriately or replaced words with their antonyms. The present study supports the finding that LD children have enormous difficulty in coping with the demands of a definition task. In fact, as noted above, the impact of this task on the LD group was greater than that on the SP group. This finding is consistent with characterisations of children with semantic-pragmatic difficulties as having relatively little difficulty with language form (Conti-Ramsden & Gunn, 1986; McTear, 1985; Rapin & Allen, 1983, 1987; Vance & Wells, 1994).

The more dramatic task effect on the LD than SP children has important implications for the drawing of conclusions from the definition task. On the play task, significant differences between the SP and LD groups were evident ("appropriate" score difference: $p = 0.0091$; "inappropriate" score difference: $p = 0.0015$). On the definition task, a comparison of their overall "appropriate" and "inappropriate" scores revealed almost identical scores ("appropriate": SP 48.2% vs LD 46.7%, n.s.; "inappropriate": SP 42.4% vs LD 45%, n.s.)\(^1\)\(^2\). That the definition task masks a significant between-groups difference militates against the use of definition tasks in formal clinical tests such as the Test of Language Competence-Expanded (Wiig & Secord, 1989) and the Fullerton Language Test for Adolescents (Thorum, 1986).

Not only does the definition task mask the difference between the SP and LD groups, but, as evidenced above, it also grossly underestimates the considerable idiom ability revealed in the SP group on the play task. When coding definition data, the author

---

\(^1\) Interestingly, a discriminant analysis of the definition task broad code data was still able to assign LD and SP children to the correct groups with a high degree of accuracy. Although the overall broad code results are very similar, slightly different patterns are evident in an idiom-by-idiom analysis.

\(^2\) The LD definition task scores do not differ significantly from those of the "SPD" or "ASP" subgroup, either. However, the performance of the "SPD" subgroup was slightly better than that of the LD group, whereas the performance of the "ASP" subgroup was slightly worse. The relationship between the subgroups mirrors that found on the play task. Interestingly, however, the between-tasks discrepancy for the "SPD" group was more significant than was that for the "ASP" group (between-tasks discrepancy for "appropriate" scores: "SPD" $p = 0.0178$, "ASP" n.s.; between-tasks discrepancy for "inappropriate" scores: "SPD" $p = 0.0007$, "ASP" $p = 0.126$). This pattern is consistent with Bishop's (1989) characterisation of "semantic pragmatic disorder" in terms of greater language difficulties than typically occur in Asperger syndrome.
often suspected that a given SP child knew the meaning of a particular idiom but she was unable to score the definition "appropriate" since it fell short of the criteria for this code. Such intuition was often proved correct when play and definition task results were subsequently compared. Examination of the SP definition data revealed a range of characteristic difficulties which, to varying degrees, functioned to reduce the likelihood of true comprehension being recognised. These included tangential responses, overly specific replies to questions, topic drift, obsessional topics of conversation, unestablished referents and word-finding difficulties. As illustrated by the following examples, these could all result in a given definition failing to meet the criteria for an "idiomatic" and/or "inferred idiomatic" code, even if the child actually understood the idiom.

Example 1: Topic drift

A: He put him in a cell and kept him under lock and key. What do you think "under lock and key" means?
SP3: under lock and key means you go into jail and if you be good you can come out but if you been still¹ naughty you can stay in you know
A: uhu
SP3: but if you be sensible you can go out of jail if you be sensible
A: so what does "under lock and key" mean?
SP3: means um have it in your hands and and put it and turn it and and lock it and and if anyone bes (/biːz/) naughty they get into the police station do they?

As evidenced by the eventual definition in the child's third turn in this interaction ("put it and turn it and and lock it"), this child did understand the idiom. However, his first two turns would not have yielded an "appropriate" score. It appears as though he is going to provide an "appropriate" response, but he drifts into a related topic and fails to return to the idiom until his attention is brought back to the task at hand by the author's prompt, "So what does 'under lock and key' mean?"

¹ Bold print reflects stressed words.

308
Example 2: Obsessional topics of conversation

A: The bike rider didn't see Rover and ran over him. What do you think "ran over" means?
SP23: (silence) um I don't know
A: okay
SP23: sort of / it was in the yard when the (car) was coming and why / you know / tha- and I've got a pattisepholosorous which I hadn't got

It is not entirely clear whether or not this child does know the meaning of "ran over". On the play task, this item scored "ambiguous-idiomatic/literal". Here he claims not to know the answer but then continues to produce what might be the beginning of an "appropriate" definition. The shift into his favourite conversational topic (dinosaurs) may represent an attempt to divert the conversation to a more comfortable topic or to take control of the conversation to avoid being asked questions.

Example 3: Giving a more specific answer than is required

A: Sally turned the bike rider over to the policeman. What do you think "turned over" means?
SP5: that means um she said this motor / um bike rider ran over my dog og and took off without even bothering to stop that's a criminal offence
A: uhu
SP5: and so please could you lock him up

While describing the process of turning the bike rider over to the policeman, this response nevertheless fails to define the critical elements of the idiom "turned over". It is unclear to what extent the overly specific and concrete response reflects definitional immaturity (Johnson & Anglin, 1995; Litowitz, 1977) rather than a pragmatic problem.

Example 4: Unestablished referent

A: The bike rider didn't see Rover and ran over him. What do you think "ran over" means?
SP9: if something's on the road it flattens it
Given that many of the SP children used personal pronouns interchangeably, it is impossible to be certain whether the subject of the verb, "flattens" refers to a person or a vehicle. The use of the word "flattens" does, however, suggest that the subject must be something heavier than a person.

Example 5: Word finding difficulties

A: Sally gave her mum a hand to carry Rover to the white car. What do you think "gave a hand" means?
SP2: hand means you / like this (holds arms as though cradling a baby)
A: can you say it?
SP2: cuddle (it means) take her to the car

Litowitz (1977) characterises gesture and mime as the earliest form of definition. The fact that this child displayed more mature forms of definition with respect to other idioms suggests that this action is more likely to reflect word finding difficulties1.

In addition to having a variable effect on different populations, the definition task also had a variable degree of impact on different idioms. Between-tasks results for individual idioms suggest that some idioms were inherently more difficult to define than others. In particular, "went round" and "got on to" were very poorly defined despite having attracted high scores on the play task. ("Went round" ranked in the top four idioms for each group on the play task. It ranked in twelfth place (out of twelve) for each of the four groups on the definition task. A similar pattern of results was found for "got on to".) It is notable that both of these idioms involve progression. Difficulty with respect to temporal relations has been reported in children with semantic-pragmatic difficulties (McConachie & Jones, 1991; McTear & Conti-Ramsden, 1992, Smedley, 1989), but the fact that all four groups defined these idioms poorly

\[1\] Although 46.2% of the SP group was noted on the Checklist to display word-finding difficulties, as many as 80% of the LD children also scored positive for this feature.
suggests that this relates to a more general difficulty. It may be that the need for these idioms to be defined in more specific terms than some other idioms accounted for their low levels of "appropriate" scores. It was further evident that many children defined these idioms in terms of other idioms which could also be interpreted literally or idiomatically. For example, "got on to" was defined by some subjects as "moved on". This expression may mean "progressed" or it may mean "moved up onto". Since this definition might represent either the literal or the idiomatic sense of "got on to", it was coded "ambiguous-idiomatic/literal". The lack of correspondence between ease of comprehension and ease of definition was reflected in the finding that there was no correlation of ranks across tasks (Spearman's Rank Correlation Coefficient revealed the following correlations across tasks: MJ: rho = 0.044, n.s.; MI: rho = 0.396, n.s.; LD: rho = 0.148, n.s.; SP: rho = 0.12, n.s.).

While the fall-off in performance on the definition task almost certainly related to the increased linguistic and metalinguistic demands of the task relative to the play task, it is also possible that fatigue played some role in producing this result. This is particularly likely for the SP and LD children, since their sessions were typically longer than those of the mainstream children. Interestingly, order effects might, theoretically, lead us to anticipate an advantage for the definition task. As noted above, the children heard the idioms in context twice between the time that they acted them out and defined them. Furthermore, the idioms had already been submitted to processing in the preceding play task. Levorato (1993) suggests that this might enhance children's performance on an ensuing idiom comprehension task. Further research demonstrates that the process of story enactment with props might itself facilitate story comprehension and retention (Pellegrini & Galda, 1982). Indeed, these children had not only acted out the idioms before defining them, but had also watched their enactment on video. A further potentially facilitative feature of the definition task was that the item of interest was identified. It was noted above
that some "fuzzy" items on the play task may have stemmed from the idioms not representing the most salient aspect of the sentence to the children concerned. This factor could not have applied to the definition task.

Clearly it is difficult to assess whether and to what extent such factors may have facilitated the demonstration of idiom comprehension on the definition task. What is evident, however, is that if these factors did enhance performance, they did not compensate for the difficulty of the task. The between-tasks results strongly indicate that the findings of idiom definition tasks should be interpreted with a high degree of caution. Nevertheless, if interpreted with care and used as an adjunct to a more valid measure, definition tasks may offer valuable additional insight into a child's ability and disability with regard to idiom comprehension.

Pioneered in this project, the idiom play task has emerged as a potentially valuable assessment tool for research and clinical practice. Certainly, it suffers some limitations. It is notable that not all idioms are amenable to testing in this format. The coding of play data is an arduous task and has been found to give rise to a higher degree of ambiguity than does the definition task. In some instances, coding difficulty has produced false positive results, although the clear cases of false positives on the play task are considerably fewer than were found on the definition task. While arguably less restrictive than previous methodologies, the play task is still unsuitable for assessing children who cannot cope with its demands (as

1 Future use of the play task might benefit from coding directly from video. Not only would this eliminate the transcription load, but it would also reduce the subjectivity of the coding process. Ideally, if the play task is being used to assess groups of children for research purposes, the video clips for each idiom should be spliced so that coding can be executed by idiom rather than by child.

2 The replacement of idioms such as "ran over" would, however, greatly reduce the level of ambiguity on the play task.

3 Of all the play task items scored "inappropriate", only 4.3% corresponded to "appropriate" definition task scores. Of the definition task items scored "inappropriate", 20.3% had "appropriate" play task counterparts. In the LD group, two thirds (67%) of "inappropriate" items on the definition task corresponded to "appropriate" play task scores.
Nevertheless, as compared with the definition task, the play task has been revealed to represent a more valid measurement of idiom comprehension in normally-developing and, particularly, in LD and SP children. It is furthermore a more sensitive tool in that it has the potential to tap partial semantic knowledge which does not yet allow the child to explain the meaning of an idiom but may be sufficient for the child to interpret it passively. By virtue of not making expressive or metalinguistic demands, the play task is further able to tap comprehension more directly than are conventional methodologies. The potential of this task to tap the ceiling of a child's idiom comprehension ability is enhanced by its being a non-threatening and intrinsically rewarding procedure. While providing a context which aims to optimise comprehension, the play task does not provide any answers for the child. Unlike multiple choice tasks, this procedure ensures that the child acts from his or her own resources. As such, it also reduces the effects of guessing and response bias. While as yet in its infancy, the play task has emerged as a potentially valuable assessment tool for research and clinical practice.
CHAPTER 7: CONCLUSIONS AND IMPLICATIONS

1. CONCLUSIONS OF THE STUDY

Consistent with the first of the main hypotheses outlined in Chapter 3, the results of this study confirm that the SP group was less able to demonstrate the idiomatic meaning of twelve common idioms on the play task than were the MJ, MI or LD groups. As predicted, the SP group scored significantly fewer "appropriate" items and significantly more "inappropriate" items than did any of the other three groups. Critically, however, the SP group was not found to be more literal in their interpretations than were the other children. None of the SP children were found to be entirely, or even predominantly, literal in their interpretations. Their relatively high levels of "inappropriate" scores derived primarily from "fuzzy" demonstrations. These may reflect an awareness of the inappropriateness of the literal meaning in the absence of adequate idiom vocabulary. Alternatively, such actions may be underpinned by a difficulty in retrieving or selecting the idiomatic meaning in context. In at least some instances, "fuzzy" items may also relate to weakness in areas of skill not directly associated with idiom comprehension, or to inherent coding difficulties.

Contrary to the second main hypothesis and despite relative weakness on the play task, the SP group was found to have scored significantly more items "appropriate" than "inappropriate" on the play task. It is likely that the inclusion only of concrete idioms (demonstrable with props) and the visual support provided by the play set facilitated idiom comprehension in this group. Nevertheless, the SP group did evidence considerable idiom comprehension ability.

While as a whole, the SP group displayed significantly poorer performance than did any of the other three groups, considerable
within-SP group variance was evident. Some SP children exhibited no comprehension impairment with regard to these idioms. Others had considerable difficulty in demonstrating "appropriate" idiom interpretation. Age differences accounted for a small proportion of this variance, but age was nevertheless a weak predictor of play task performance. Appreciable differences between the "SPD" and "ASP" subgroups on the play task and on the Checklist suggested that different patterns of strength and weakness with respect to "critical" pragmatic and cognitive skills accounted for much of the variance in play task performance. Weakness in idiom comprehension and in the skills argued to underpin it were associated to a greater extent with the "ASP" subgroup than with the "SPD" subgroup. Relative ability and disability were, however, evidenced within both of these diagnostic groups.

In addition to variance in overall play task scores, individual SP children also appeared to differ with regard to the nature of their idiom difficulties. Children whose levels of "appropriate" interpretation were superficially similar, could nevertheless be argued to differ with regard to the sites and sources of their idiom comprehension breakdown.

While both the play task and the definition task were found to be useful sources of information in idiom comprehension assessment, the play task emerged as being a considerably better indicator of a child's ceiling of ability. Between-tasks results supported the third main hypothesis. The SP group was found to be significantly less able to demonstrate idiom comprehension on the definition task than on the play task. Poorer performance on the definition task than on the play task was also found for each of the other three groups.

The findings of the study with regard to the mainstream and LD groups were consistent with the literature. A trend for the older MJ group to score more items "appropriate" than did the younger MI group was observed, although the difference was not significant. This may reflect the commonality of the idioms
investigated. The performance of the LD group was found to be significantly poorer than that of the MJ group, but was not significantly different from that of the MI group. These results are consistent with the conclusions of previous studies that the idiom comprehension of language-disordered children in this age range tends to lag behind that of normally-developing children by two- to three years.

Between-idioms results suggest that the inherent properties of idioms in interaction with the contexts in which they are encountered influence the ease with which specific idioms are interpreted in context. This combination may also affect the rate at which different idioms are acquired.

2. IMPLICATIONS FOR THE CHARACTERISATION OF CHILDREN WITH SEMANTIC-PRAGMATIC DIFFICULTIES

The results of this study indicate that children with semantic-pragmatic difficulties vary considerably with regard to idiom comprehension ability and disability. This variation probably reflects more fundamental heterogeneity in relation to a range of skills underpinning idiom comprehension. These include flexibility of thought, theory of mind, attention to context, prosody and overall coherence, and the ability to integrate world knowledge and current contextual information to guide inferencing.

While the results of this study cannot necessarily be generalized, it is evident that not all children with semantic-pragmatic difficulties are impaired in their comprehension of (at least these) common, concrete idioms. Indeed, even those children who appear to be impaired in idiom comprehension may nevertheless display some degree of idiom competence. It is further evident that idiom comprehension difficulty does not necessarily imply fundamental literality. Such difficulty may
relate to problems of acquiring new idioms and/or to difficulty in using context to trigger the retrieval and guide the selection of learned idiomatic meaning.

It thus appears that the characterisation of children with semantic-pragmatic difficulties as predominantly literal needs to be revised. At a very general level, it may be said that children with semantic-pragmatic difficulties tend to have mild to severe degrees of difficulty in interpreting idioms in context. However, rather than seeking a blanket characterisation of an essentially heterogeneous group, it may be more useful to speak of idiom comprehension impairment as a secondary manifestation of semantic and/or pragmatic difficulties. Since the group of children who have these difficulties probably vary in symptomatology and aetiology, it follows that they will also vary with respect to idiom comprehension ability and the reasons for its breakdown.

3. IMPLICATIONS FOR DIFFERENTIAL DIAGNOSIS

Despite considerable variability, the results of the play task support a distinction based on idiom disability between children considered to have semantic and pragmatic difficulties and those whose language difficulties are thought not to be primarily of a semantic and/or pragmatic nature. However, the difference appears to be one of degree rather than of ability vs disability. It may also be that the sites and sources of breakdown differ across these groups.

In addition to discriminating SP and LD groups, the play task has also been found to discriminate between children said to have "semantic-pragmatic disorder/difficulties" and those diagnosed with Asperger syndrome or high-functioning autism. Given that there was considerable variance within each of these groups, the results need to be interpreted with caution. However, these
findings do suggest that there is some consistency with respect to the basis on which the label "semantic-pragmatic disorder/difficulties" as opposed to "Asperger syndrome"/"high-functioning autism"/"autism" is being assigned by clinicians. The results are further indicative of some meaningful difference between these two diagnostic groups.

4. IMPLICATIONS FOR TEACHING AND THERAPY

At the outset of this report, it was proposed that the recommendation to teachers to avoid the use of idioms and other forms of non-literal language in interaction with language-disordered children may be both unfeasible and undesirable. Evidence from teacher-language samples and discussion with teachers support the view that such advice would be difficult to heed. The results of this study suggest further that children with semantic-pragmatic and other language difficulties may have a reduced ability to acquire the concept of idiomaticity, to learn new idioms and to practice idiom comprehension. As such, it may be particularly important for these children that parents, teachers and clinicians do use idioms, are aware of their own use of idioms, monitor comprehension, reinforce comprehension and provide explicit definitions for unfamiliar expressions.

As argued above, idiom comprehension involves the integration of both semantic and pragmatic knowledge. As such, general therapeutic intervention aimed at increasing flexibility of thought, attention to coherence, use of context, application of world knowledge and inferencing skills will inevitably have a knock-on effect on idiom comprehension. Development of comprehension monitoring and help-seeking strategies within such children will also promote the triggering of idiomatic meanings and the learning of new idioms.

Given that idiom comprehension is at least as much a pragmatic
task as it is semantic, rote learning of idioms on its own is unlikely to have a dramatic impact on the interpretative competence of a child with pragmatic disability. Nevertheless, specific attention to idioms may be valuable, particularly if it focuses on idioms very commonly used in the child's environment, incorporates generalization of learning across contexts and focuses on the entire process of interpreting idioms in context (see Appendix K).

The results of this study suggest that, in many instances, children with semantic-pragmatic difficulties in the age range 6;6 to 11;6 will already have grasped the fundamental concept of idiomaticity. Baseline assessments may reveal varying degrees of idiom ability which would provide a vital foothold for therapeutic intervention. In the light of the finding that definition tasks underestimate idiom comprehension, the usefulness of current idiom comprehension tests as a baseline for therapy needs to be questioned. The play task, in its current form or with different idioms embedded, may prove a useful alternative or adjunct for assessing the ceiling of a child's idiom comprehension and for teaching idioms (see Appendix K). Particularly if combined with definition task data and additional information about the child's general functioning with regard to social interaction, imagination and communication, the play task may help to identify sites and sources of idiom comprehension breakdown in individual children. With idioms drawn from the child's own linguistic environment, a modified version of the play task might also be used with normally-developing children to establish norms for the specific idioms included. Such information would facilitate the evaluation of a given child's ability or disability relative to the norm.

It is envisaged that such information would constitute a starting point in the assessment of idiom comprehension in children with semantic-pragmatic difficulties. Once a picture of the child's ceiling of ability has been obtained, his/her comprehension in less facilitative contexts and with respect to more abstract and
6. IMPLICATIONS FOR RESEARCH

While the aim of this study was to investigate idiom comprehension in children with semantic-pragmatic difficulties, the results have some important implications for idiom comprehension research in general. The examination of idiom comprehension impairment leads us to conclude that the task of idiom comprehension is at least as much pragmatic as it is semantic. This finding strongly supports the trend in idiom comprehension research towards investigating children's ability to interpret idioms in context.

The results of this study further confirm previous findings as to the weakness of definition tasks as measures of idiom comprehension. In particular, young children and children with expressive, metalinguistic, semantic or pragmatic difficulties have been found to be seriously disadvantaged by the demands of a definition task. The validity of research methodologies based on definition tasks is therefore questionable. Nevertheless, as an adjunct to a more appropriate assessment tool, a definition task may contribute valuable information and help to clarify points of ambiguity that may arise from other procedures such as the play task.

While subject to some limitations, the idiom play task has been shown to represent a valuable new tool for the assessment of idiom comprehension in both clinical and normally-developing populations. The literature indicates that the use of appropriate methodologies may reveal idiom ability in children previously considered to be unable to interpret expressions idiomatically. The results of the MI group on the play task suggest that investigation of children younger than 6 years of age using the play task may well uncover idiom comprehension.
ability in very young children. Such research may, furthermore, shed light on early patterns of learning with regard to different types of idioms.

As an arguably more valid measure of idiom comprehension in children with semantic-pragmatic difficulties, the play task also opens the way to more detailed research on idiom ability in this population. Future research might focus on different types of idioms and their impact on larger groups of children considered to have semantic-pragmatic difficulties. Given that teachers and clinicians must necessarily be selective with regard to which idioms are targeted in intervention, further insight into the differential impact of different idiom subtypes and different contexts would be valuable. In the light of the significant, albeit weak, SP age-performance correlations found in this study, it would also be interesting to trace the development of idiom comprehension ability in a range of children considered to have semantic-pragmatic difficulties. Such research would clearly need to take account of therapeutic input and levels of idiomatic exposure, among other variables.

Beyond idiom comprehension, investigation of idiom usage by children considered to have semantic-pragmatic difficulties may well offer a valuable additional source of information.

As evidenced by this study, systematic investigation of idiom comprehension disability is valuable not only for its clinical application, but also for the insight it might afford into idiom acquisition and processing in the normally-developing population.
REFERENCES


322


335


APPENDIX A: KERBEL, GRUNWELL AND GRUNDY (1996)

The paper below has been accepted for publication in the European Journal of Disorders of Communication, Vol. 31 (1), 65-75.
A PLAY-BASED METHODOLOGY FOR ASSESSING IDIOM COMPREHENSION IN CHILDREN WITH SEMANTIC-PRAGMATIC DIFFICULTIES

DEBRA KERBEL, PAM GRUNWELL and KIM GRUNDY

Department of Human Communication
De Montfort University, Leicester

Keywords: idiom comprehension, semantic-pragmatic difficulties, language disorders, assessment
ABSTRACT

In response to the lack of suitable materials for assessing idiom comprehension in children with semantic-pragmatic difficulties, a new, play-based methodology is presented. The child listens to a tape-recorded story into which are embedded a range of idiomatic expressions. The child then acts out the story as it is played again, sentence by sentence. The play is video-recorded and transcribed. Actions for each idiom are categorised and are then amenable to within-child, between child and between-group analyses. The results of a pilot project indicate that this methodology overcomes many of the problems inherent in assessing children identified as having semantic-pragmatic difficulties.
This paper presents a new, play-based methodology developed for the investigation of idiomatic comprehension in children identified as having semantic-pragmatic difficulties (SPD). One of the central features characterising this group is a significant difficulty in using and understanding all forms of non-literal language. The tendency among SPD children to over-literalise is manifest in rigid concept boundaries, in the failure to grasp the intended meaning of indirect speech acts, and in the misinterpretation of humour, irony, sarcasm, metaphor, similes, proverbs and idioms (Bishop & Adams, 1989; Smedley, 1989; McTear & Conti-Ramsden, 1992).

In the absence of assessment materials appropriate to the investigation of SPD children, the development of a new methodology was considered necessary. This paper presents a critique of existing research and assessment methodologies and examines the manner in which idioms have been defined in these materials. It then outlines the definition of idioms applied in this study and the process whereby items were selected for the new procedure. The development of the play-based methodology is reported and the new procedure, designed to overcome the problems inherent in conventional methodologies is described and discussed.

A CRITIQUE OF EXISTING RESEARCH AND ASSESSMENT METHODOLOGIES

Studies of idiom comprehension in children have, to date, drawn on two basic methodologies: definition tasks and multiple choice tasks. Idioms are presented in isolation, in sentences or in paragraphs. The subjects are then required to define the expressions (Ackerman, 1982; Prinz, 1983; Gibbs, 1987, 1991; Nippold & Martin, 1989) and/or to select the most appropriate meaning from a written (Cacciari & Levorato, 1989; Levorato & Cacciari, 1992), oral (Ackerman, 1982; Gibbs, 1987, 1991) or pictorial multiple choice (Lodge & Leach, 1975; Strand & Fraser, 1979; Prinz, 1983; Brinton, Fujiki & Mackey, 1985; Ezell &
Goldstein, 1991; Vance & Wells, 1994).

Formal tests of idiom comprehension have, similarly, employed one or a combination of these tasks. The idioms subtest of the Test of Language Competence - Expanded Edition (TLC-E) (Wiig & Secord, 1988) presents a series of idioms in two-sentence contexts. The child is required firstly to define the idiom and, secondly, to select its meaning from a multiple choice. Children aged 5;0 - 9;11 years choose from four line drawings, while those aged 9;0 - 18;11 years select from four written options.

The TLC-E is designed "to identify children ... who have not acquired the expected levels of metalinguistic competence in semantics, syntax and/or pragmatics" (p. 1). The methodologies employed in its idioms subtest (one of four subtests) serve, however, to weaken its validity with respect to SPD children. The definition task requires subjects to use expressive skills characteristically disordered in SPD children (Bishop & Adams, 1989; Smedley, 1989). The pictorial multiple choice task makes demands on a child's ability to draw logical inferences, interpret non-verbal communication, identify salient features and integrate information. Difficulties in these areas are characteristic of children with SPD.

The verbal multiple choice gives rise to further weakness in that the correct option corresponding to each target idiom is itself always idiomatic. For example, the correct meaning for the sentence "Maybe we should stew over that" is "Let's toss that around some more". In some instances, all of the options include idioms. Clearly, making the correct choice depends on an understanding not just of the target idiom, but of several others as well. In this context, incorrect choices do not allow for any clear conclusions to be drawn.

Given that its methodology disadvantages SPD children, along with the potentially confounding variables of guessing, response bias and ordering effects associated with multiple choice tasks, the
TLC-E emerges as an inappropriate instrument for investigating these children. Its validity with British children is further questionable in view of its use of American idioms.

The idiom subtest of the Fullerton Language Test for Adolescents (Thorum, 1986) (FLTA) is equally unsuitable. Serious weaknesses in this test have been identified by several authors. Firstly, the FLTA divides scores for each age range into those that reach competency level, instruction level and frustration level. Using the FLTA's criteria, Stephens & Montgomery (1985) and Lieberman, Heffron, West, Hutchinson & Swem (1987) reveal that most normal subjects fail to attain the competency level. Indeed, even among university SLT students, surprisingly few achieved competency scores (Brasseur & Jiminez, 1989).

Secondly, Thorum, himself, found that only three out of the 20 items were correctly defined by 80% of the norming sample. This low level of accurate definitions was replicated with university students (Brasseur & Jiminez, 1989).

The validity of the FLTA is, thirdly, undermined by the surprising absence of an age-related progression in ability (Stephens & Montgomery, 1985). Thorum (1980) asserts that the test is designed not as a developmental measure but only to discriminate between language-impaired and language-normal students. Unfortunately, as Stephens and Montgomery observe, this claim for predictive validity is not supported by any published data.

In the light of these important weaknesses, results yielded by this test for any population need be treated with great caution. It is further unclear to what extent the results reflect a child's ability to recognise idioms in real life situations since the procedure begins with an explanation of the nature of idioms and then subjects are specifically instructed to explain the idiomatic meaning of each item. Its suitability for use with SPD children is further undermined by its reliance on a verbal
expression task.

In contrast to the TLC-E and FLTA, the Conley-Vernon Idiom Test (Conley, 1976) makes no expressive demands on its subjects. Sentences are presented with a final idiom omitted and the child fills in the blank from a choice of five responses. Unfortunately, the test's validity is undermined by the fact that up to three out of the five options are syntactically inappropriate and by the inclusion of more than one plausible answer for a given item. In addition, as many as four of the five options are themselves idiomatic expressions, giving rise to the same confounding variable as in the TLC-E.

DEFINING IDIOMS

Given the inadequacies of existing assessments, it was necessary not only to devise an original investigation procedure, but also to define and identify appropriate idioms for inclusion in the procedure.

By definition, an idiom is a polylexemic expression, the meaning of which cannot be deduced from the meaning of the individual words. Like words, however, idioms have fixed, conventional meanings entered in dictionaries (Allen, 1986). For the purposes of this study, an expression is judged to be an idiom if it is included in the Oxford Dictionary of Current Idiomatic English (Vol. 1: Cowie & Mackin, 1993 and Vol. 2: Cowie, Mackin & McCaig, 1985).

Previous research studies and assessment tools have defined idioms in the same manner. In fact, dictionaries have constituted the sole source of idioms for most research and assessment materials. The practice of drawing idioms from dictionaries rather than from real life gives rise however to a potentially artificial assessment. Indeed there is no way of knowing whether the expressions included are ones to which
children are generally exposed.

Assessment and research tools have, for the most part, focused exclusively on the type of idiom exemplified by the expressions "spill the beans" and "kick the bucket". Such materials are representative of only the "pure idiom" subtype, failing to assess comprehension of other subtypes on what is most accurately viewed as a spectrum of idiomaticity (Cowie et al., 1985). The idiom collection procedure used in this study (described below) in fact found that most idioms to which children are exposed in classroom settings are phrasal verbs (e.g. "run over", "carry on"). Significantly, some researchers appear to include phrasal verb idioms inadvertently in their contextualising sentences, thus weakening the power of their methodologies (Gibbs, 1987; Nippold & Martin, 1989; Vance and Wells, 1994). The procedure employed by Nippold and Martin, for example, included the following item: "Peter was bragging about all the fish he had caught. When he found out that his sister had caught more, it took the wind out of his sails" (our italics) (1989: 66).

THE PLAY-BASED ASSESSMENT PROCEDURE

In view of the problems associated with conventional methodologies, the present study sought to develop an alternative approach capable of yielding valid data with regard to SPD children. The procedure requires the child to listen to a tape-recorded story into which a range of idioms are embedded. Previous studies have indicated that idiom comprehension is facilitated when expressions are presented in context (Brinton et al., 1985; Cacciari & Levorato, 1989; Nippold & Martin, 1989; Abkarian, Jones & West, 1990). After listening to the story once, the child hears it again and acts it out, sentence by sentence, using the play set and props provided. The play and the story are recorded on video and subsequently analyzed.
(i) MATERIALS

Given that the findings of this procedure have potentially significant educational and therapeutic implications, it was considered essential that the idioms included be ones to which children are exposed. To this end, 2200 idioms were extracted from 31 hours of children's British television and 21 hours of routine classroom teaching in mainstream and language schools (Kerbel, Grunwell & Grundy, in preparation).

In order to ensure the validity of the procedure, a set of eligibility criteria had to be met before an idiom could be considered for inclusion. It was necessary, first of all, that the idioms had both a figurative and literal sense and, secondly, that both of these meanings were clearly demonstrable with the available props. These criteria ensured that the play task would allow for the child's literalization strategy or idiomatization strategy to be demonstrated. For this reason, idioms were excluded if they could only be indirectly demonstrated (e.g. "be on top of the world", "change one's mind"), had closely related literal and figurative meanings which might not have been distinguishable when demonstrated (e.g. "be sick", "go to bed", "keep one's eyes open"), or if their figurative sense was difficult to demonstrate (e.g. "run out").

It was further necessary that no expressive language on the part of the child be required. As noted above, children with SPD are characterised by numerous expressive language problems which might distort the outcome of a task relying on verbal expression. Thus, all idioms which required verbal expression or cessation of speech in the play task were disqualified (e.g. "you can say that again", "shut up"). Similarly, idioms of which young children could not be assumed to have experience were excluded (e.g. "fruit machine"; "shoot up"; "the silver screen").

Drawing from those confirmed as suitable, 16 idioms were embedded into a 400-word story. The language of the story was
deliberately simple and unambiguous. Following a pilot of the procedure using mainstream and SPD children, the story was partially modified. Four idioms were removed and some recontextualization took place. These amendments ensured that all of the idioms would be understood by the majority of 6-year old, mainstream children and that the literal and idiomatic actions corresponding to each idiom could be clearly distinguished by the tester. The story was reduced to 277 words and 1.5 minutes in length. (A transcript of the story is included in Appendix A.)

In the interests of standardisation, the children listen to a tape-recording of the story read by a received pronunciation speaker. This was selected as the accent least likely to bias any particular children. The reader took care not to stress the idiomatic expressions unnaturally.

(ii) **PLAY SET**

A play set, measuring 78 cm x 61 cm, was constructed from plywood. This includes the buildings and road referred to in the story. In order to maximize the camcorder's view, the buildings are constructed so as to have two walls, two open sides and an open roof.

Additional props, scaled to size, correspond to those mentioned in the story. These are placed in fixed positions on the play set. The characters of the story are represented by Playmobil people. Their realism, size and the mobility of their limbs and heads make them suitable for the acting out of the story by junior school children.

**FIGURE 1 (PHOTOGRAPH OF PLAY SET) HERE**
(iii) ADMINISTRATION

Following a brief conversation, the child is told, "Today we are going to make our own video". The tester then identifies each of the characters in the story and indicates all the locations for the actions. The child is then told

On this tape, I have a story about Sally and Rover and all of these people. First we'll listen to the whole story. Then we'll listen to it again and I'd like you to make all of these people do whatever the story says. First, let's listen to the whole story. You just need to listen.

Once the child has heard the story, the tester says, "That's the whole story. Now we'll listen to the story again. This time, make the people do whatever the story says". The tester then gives the child a simple example using the characters of the story and demonstrates how it would be enacted using the props. The child is required to act out two consecutive trials correctly before the procedure continues. This ensures that the child has grasped the nature of the task. The number of trials required and the nature of the child's response are noted.

Once the child has achieved two consecutive correct trials, the tester asks, "Do you understand what I want you to do?" If the child seems uncertain, the procedure is explained further and trials continue. If the child indicates understanding, the tester reminds him or her what s/he needs to do and stresses that the child does "not need to rush" since the tape will be stopped "after each bit of the story". The story is then played. The entire procedure is filmed on video.

During the task, the tester remains silent except in the
following conditions. If the child repeats all or part of the idiom or the sentence in which it is contained, the tester says, "You do it". If the child asks how to act something or suggests an action, the tester says "Do whatever the story says". If the child fails to act or queries the meaning of an idiom, the tester says "Let's listen again and then you do it". The sentence is then replayed.

(iv) TRANSCRIPTION AND ANALYSIS

As soon as possible following the session, the video tape is transcribed. The transcription takes the form of a description of the position of relevant props prior to each sentence, as well as a detailed description of the child's actions from the time the sentence is heard to the time the next sentence begins. Any verbal or non-verbal communication is included. Responses corresponding to each idiom in turn is collated.

Once transcribed, the actions for each item are categorised in accordance with the categories described in Table 1. For each idiom, criteria for an acceptable idiomatic demonstration and an acceptable literal demonstration have been devised. These represent the polar extremes of the continuum for that idiom. Following Zadeh (1965), the cline of the continuum is divided into a "fuzzy-idiomatic" and a "fuzzy-literal" category.

Once categorised, the results are examined for evidence of
patterns of understanding and response types corresponding to different types of idioms. These patterns can be compared within the child, within the population group or between groups.

**TABLE 1: CATEGORIES FOR ANALYSIS**

* These examples, drawn from pilot data, all relate to the sentence, "The bike rider stopped on the spot".

1. **IDIOMATIC**
   
The action clearly demonstrates the target idiomatic sense of the expression. This may involve a series of actions or merely one action.
   (* e.g. the child immediately stops the bike rider wherever he is)

2. **FUZZY-IDIOMATIC**
   
The action indicates that the child has responded to the expression in its idiomatic sense but is not a perfect demonstration of this meaning.
   (e.g. the bike rider having toppled, the child lifts him and places him where he has been lying)

3. **FUZZY-LITERAL**
   
The action indicates that the child has responded to the expression in its literal sense but is not a perfect demonstration of this meaning.
   (e.g. the child places the bike rider on the spot where the accident previously occurred)

4. **LITERAL**
   
The action clearly demonstrates the literal sense of the expression.
   (e.g. the child places the bike rider on the green dot stuck onto the play set)

5. **AMBIGUOUS**
   
The action could be interpreted as a demonstration of either an idiomatic or literal understanding of the expression. It cannot therefore be confidently coded as one or the other.
   (e.g. the bike rider is on the green dot at the moment the child hears the sentence and the child immediately stops him there)

6. **FUZZY**
   
The action cannot be interpreted or does not directly relate to the relevant expression.
   (e.g. the child moves the already-stationary bike rider to another location)

7. **APPROPRIATE NO ACTION**
   
Prior to the expression, the props are already arranged in accordance with the idiomatic sense of the expression. The child appropriately does not make any alterations in response to the expression.
   (e.g. the bike rider is already stationary and the child does not move him)

8. **INAPPROPRIATE NO ACTION**
   
The child does not act on the props at all when it would be appropriate to do so in response to the expression.

9. **TECHNICAL ERROR**
   
The child's action was either not filmed or was interrupted by the premature playing of the next part of the story.
DISCUSSION

The methodology presented in this paper was piloted on five mainstream children aged 6;2 - 6;7 years, five mainstream children aged 9;2 - 10;8 years, and five children aged 6;8 - 9;3 who had been identified by SLTs as having SPD.

The results of the pilot stage of this project support the play-based procedure's ability to overcome the difficulties associated with assessing SPD children. Vance and Wells's (1994) assert that "(as) limitations of play are characteristic of children with SPD (Bishop and Rosenbloom, 1987) any procedure involving acting-out with toys might prejudice their performance" (p. 31). In fact, the SPD children assessed in the pilot project coped well with the demands of the procedure. They were all noted to be scanning the set during the first run of the story, apparently identifying the relevant people and props as the story progressed. The number of trials required ranged from two to six. Following these, the children all appeared to have grasped the nature of the task, acting out all non-idiomatic language appropriately. No child displayed any difficulty in manipulating the toys. They all appeared to enjoy the task and displayed neither restlessness nor any reluctance.

The necessity for strict eligibility criteria inevitably restricts the range of idioms whose comprehension can be assessed by this procedure. Certainly this is a limitation. It does, however, appear to be outweighed by a number of important advantages over conventional assessment measures. Significantly, this methodology enables the child to display his or her passive comprehension of idioms without drawing on expressive language or metalinguistic skills. The physical presence of the play set and people provide concrete visual support which may enhance otherwise reduced attention and listening abilities. The fact that all the children scanned the set appropriately during the first run of the story lends
support to this view. It was hoped that this, together with the repetition of the story, might assist the child to make better use of the idiomatic-biasing context. These features increase the likelihood of tapping the child's ceiling of ability.

The use of this format further ensures that the child acts from his or her own resources rather than selecting from a choice of options. This reduces the potential interference of guessing and response bias, and eliminates ordering effects.

While the task is in fact a means of investigating idiom comprehension, it appears to the child to focus on the making of a video and playing with toys associated with recreation. It is thus intrinsically rewarding to the child and greatly reduces the formality characteristic of paper-and-pencil tests. The focus on making a video as opposed to testing, and the reduced necessity for eye contact and social interaction function further to make the situation less threatening. In contrast to many other tests, this task has the added advantage of not allowing the child to fail. This is an important feature and one which helps to sustain motivation in children who typically have extensive experience of failure.

In addition to its value as a research methodology, the play-based procedure could be usefully applied in clinical settings. In this context, it is envisaged that the child's performance would be scored during the procedure with the video recording being used to check these results. A play-based methodology has further potential for assessing children's comprehension of, for example, phonemic distinctions and syntactic constructions.

Following encouraging pilot results, the procedure is currently being employed in the second stage of the research investigation. The results of this study will be reported in a future publication.
ACKNOWLEDGEMENTS

The authors wish to thank Jude Balding for the video-recording; David Kerbel for the construction of the play set; and the staff and pupils of Cherryleas Assessment Centre, Dawn House School, Marriott Primary School, Sherard County Primary School, St Mary's Fields Primary School, Thurnby Lodge Primary School and Uffculme School. This work forms part of a PhD project undertaken by the first author in the Department of Human Communication, De Montfort University, Leicester and supported by De Montfort University.

POST SCRIPT

Following completion of this pilot study, it was discovered that a similar methodology was used by Vosniadou, Ortony, Reynolds and Wilson (1984) for investigating normally-developing children's understanding of metaphors. This investigation required children to act out metaphors using a playset. The play-based methodology is therefore not entirely novel, but its use in investigating idiom comprehension and in the clinical context is innovatory and has considerable potential for further development.

REFERENCES


356
One day, Sally and Rover were playing in the garden. Suddenly, Rover ran into the road. A green motorbike was driving on the road. The bike rider didn't see Rover and ran over him. The bike rider took off without stopping. Rover lay in the road. Sally's mum ran to Rover. Sally gave her mum a hand to carry Rover to the white car. They put him inside. Sally's mum drove to the vet's office. She carried Rover into the room. She put Rover on the floor. There was a heap of bandages on the vet's table. The vet moved them up so that there was enough space for Rover. Sally's mum put Rover on the table.

The vet examined Rover's head. Then he got on to Rover's legs. Rover's leg was broken.

Sally and her mum left the vet's office. They saw the bike rider walking in the road. Sally chased him. The bike rider started running. Suddenly the bike rider stopped on the spot. There was a policeman in the road. Sally turned the bike rider over to the policeman. The policeman took the bike rider to the police station. He put him in a cell and kept him under lock and key. The bike rider wanted his motorbike at the police station. The policeman went to pick it up.

The next day, Sally's mum drove to the vet's office. She brought Rover home. Sally put Rover into his basket. Sally kept an eye on him all day.

That night, the policeman dropped in to visit Sally and Rover. He patted Rover. Rover was very happy. Rover went round everyone and licked them all.
APPENDIX B: DETAILS OF SUBJECTS

Tables 65 and 66 below present the personal details pertaining to the mainstream subjects in this study. These tables also include the scores achieved by these children on the TROG (Bishop, 1983) which was administered as a pretest to all of the mainstream subjects.

### TABLE 65: MI CHILDREN - PERSONAL DETAILS AND TROG SCORES

<table>
<thead>
<tr>
<th>SUBJECT</th>
<th>GENDER</th>
<th>AGE</th>
<th>TROG BLOCKS PASSED</th>
<th>TROG CENTILE</th>
</tr>
</thead>
<tbody>
<tr>
<td>M11</td>
<td>M</td>
<td>6;11</td>
<td>16</td>
<td>75</td>
</tr>
<tr>
<td>M12</td>
<td>F</td>
<td>7;0</td>
<td>17</td>
<td>75</td>
</tr>
<tr>
<td>M13</td>
<td>M</td>
<td>7;3</td>
<td>18</td>
<td>90</td>
</tr>
<tr>
<td>M14</td>
<td>F</td>
<td>7;0</td>
<td>14</td>
<td>62</td>
</tr>
<tr>
<td>M15</td>
<td>F</td>
<td>6;8</td>
<td>18</td>
<td>95</td>
</tr>
<tr>
<td>M16</td>
<td>M</td>
<td>6;9</td>
<td>18</td>
<td>95</td>
</tr>
<tr>
<td>M17</td>
<td>M</td>
<td>6;10</td>
<td>18</td>
<td>95</td>
</tr>
<tr>
<td>M18</td>
<td>M</td>
<td>7;6</td>
<td>12</td>
<td>20</td>
</tr>
<tr>
<td>M19</td>
<td>F</td>
<td>6;6</td>
<td>15</td>
<td>62</td>
</tr>
<tr>
<td>M110</td>
<td>M</td>
<td>6;7</td>
<td>19</td>
<td>99</td>
</tr>
<tr>
<td>M111</td>
<td>M</td>
<td>6;7</td>
<td>14</td>
<td>50</td>
</tr>
<tr>
<td>M112</td>
<td>M</td>
<td>7;0</td>
<td>12</td>
<td>20</td>
</tr>
<tr>
<td>M113</td>
<td>M</td>
<td>7;6</td>
<td>13</td>
<td>25</td>
</tr>
<tr>
<td>M114</td>
<td>F</td>
<td>6;8</td>
<td>14</td>
<td>50</td>
</tr>
<tr>
<td>M115</td>
<td>F</td>
<td>6;11</td>
<td>16</td>
<td>75</td>
</tr>
</tbody>
</table>
### TABLE 66: MJ CHILDREN – PERSONAL DETAILS AND TROG SCORES

<table>
<thead>
<tr>
<th>CHILD</th>
<th>GENDER</th>
<th>AGE</th>
<th>TROG BLOCKS PASSED</th>
<th>TROG CENTILE</th>
</tr>
</thead>
<tbody>
<tr>
<td>MJ1</td>
<td>F</td>
<td>10;6</td>
<td>16</td>
<td>25</td>
</tr>
<tr>
<td>MJ2</td>
<td>F</td>
<td>11;2</td>
<td>19</td>
<td>95</td>
</tr>
<tr>
<td>MJ3</td>
<td>M</td>
<td>10;9</td>
<td>19</td>
<td>95</td>
</tr>
<tr>
<td>MJ4</td>
<td>M</td>
<td>10;6</td>
<td>17</td>
<td>38</td>
</tr>
<tr>
<td>MJ5</td>
<td>F</td>
<td>11;0</td>
<td>19</td>
<td>95</td>
</tr>
<tr>
<td>MJ6</td>
<td>F</td>
<td>10;11</td>
<td>19</td>
<td>70</td>
</tr>
<tr>
<td>MJ7</td>
<td>M</td>
<td>10;7</td>
<td>19</td>
<td>95</td>
</tr>
<tr>
<td>MJ8</td>
<td>F</td>
<td>10;7</td>
<td>19</td>
<td>95</td>
</tr>
<tr>
<td>MJ9</td>
<td>M</td>
<td>10;6</td>
<td>19</td>
<td>25</td>
</tr>
<tr>
<td>MJ10</td>
<td>F</td>
<td>10;10</td>
<td>19</td>
<td>70</td>
</tr>
<tr>
<td>MJ11</td>
<td>M</td>
<td>11;6</td>
<td>19</td>
<td>70</td>
</tr>
<tr>
<td>MJ12</td>
<td>F</td>
<td>10;9</td>
<td>19</td>
<td>95</td>
</tr>
<tr>
<td>MJ13</td>
<td>F</td>
<td>10;9</td>
<td>20</td>
<td>99</td>
</tr>
<tr>
<td>MJ14</td>
<td>F</td>
<td>10;6</td>
<td>19</td>
<td>70</td>
</tr>
<tr>
<td>MJ15</td>
<td>F</td>
<td>10;6</td>
<td>19</td>
<td>95</td>
</tr>
</tbody>
</table>

The TROG scores above indicate that the mainstream children who participated in this study encompassed a range of comprehension ability. As noted in Chapter 4, children who scored below -1 standard deviation\(^1\) were eliminated from the study.

---

\(^1\)In order to be included in the study, children aged 6;6 - 6;11 thus had to pass at least ten blocks, while those aged 7;0 - 7;6 had to pass eleven blocks. Children aged 10;0 - 10;6 had to pass 15 blocks and those aged 11-0 - 11;6 had to pass 16 blocks.
Table 67 below presents the personal details pertaining to the LD subjects in this study. It also presents the diagnosis or description of each child's difficulties, as provided by SLTs.

### TABLE 67: LD CHILDREN - PERSONAL DETAILS

<table>
<thead>
<tr>
<th>SUBJECT</th>
<th>GENDER</th>
<th>AGE</th>
<th>DIAGNOSIS / DESCRIPTION OF PROBLEMS PROVIDED BY SLT</th>
</tr>
</thead>
<tbody>
<tr>
<td>LD1</td>
<td>F</td>
<td>10;4</td>
<td>Expressive language difficulties</td>
</tr>
<tr>
<td>LD2</td>
<td>M</td>
<td>10;1</td>
<td>High-level comprehension and expression difficulties, and phonological awareness problems</td>
</tr>
<tr>
<td>LD3</td>
<td>M</td>
<td>10;4</td>
<td>Language disorder</td>
</tr>
<tr>
<td>LD4</td>
<td>M</td>
<td>10;7</td>
<td>Weak auditory memory and semantic disorder affecting all levels of expressive language and language use</td>
</tr>
<tr>
<td>LD5</td>
<td>M</td>
<td>8;7</td>
<td>Moderate receptive and expressive language disorder</td>
</tr>
<tr>
<td>LD6</td>
<td>M</td>
<td>8;8</td>
<td>Phonological and language disorder</td>
</tr>
<tr>
<td>LD7</td>
<td>M</td>
<td>10;5</td>
<td>Global language disorder (very slow at processing)</td>
</tr>
<tr>
<td>LD8</td>
<td>M</td>
<td>8;2</td>
<td>Phonological impairment and dysfluency (stammer)</td>
</tr>
<tr>
<td>LD9</td>
<td>M</td>
<td>9;8</td>
<td>Receptive and expressive impairment with some pragmatic problems</td>
</tr>
<tr>
<td>LD10</td>
<td>M</td>
<td>9;7</td>
<td>Phonological impairment</td>
</tr>
<tr>
<td>LD11</td>
<td>M</td>
<td>11;2</td>
<td>Receptive and expressive language disorder</td>
</tr>
<tr>
<td>LD12</td>
<td>M</td>
<td>10;10</td>
<td>Receptive and expressive language disorder (dyspraxia)</td>
</tr>
<tr>
<td>LD13</td>
<td>F</td>
<td>10;11</td>
<td>Difficulty processing abstract information; comprehension difficulties; central organisation difficulties</td>
</tr>
<tr>
<td>LD14</td>
<td>M</td>
<td>8;6</td>
<td>Poor listening skills; poor high-level comprehension and organisation of output; some semantic problems</td>
</tr>
<tr>
<td>LD15</td>
<td>F</td>
<td>10;8</td>
<td>Language disorder</td>
</tr>
</tbody>
</table>

As noted in Chapters 2 and 4, LD subjects were identified by SLTs and had been diagnosed prior to this study. Language test
scores, provided by the children's SLTs, are included for information in Tables 68, 69 and 70 below. "Chronological age" is abbreviated as "CA" and "age equivalent" is abbreviated as "AE". The names of tests are abbreviated as follows:-

<table>
<thead>
<tr>
<th>Test Code</th>
<th>Test Name</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>BOEHM TBC</td>
<td>Boehm Test of Basic Concepts (Boehm, 1986)</td>
<td></td>
</tr>
<tr>
<td>BPVS</td>
<td>British Picture Vocabulary Scale (Dunn, Dunn, Whetton &amp; Pintillie, 1982)</td>
<td></td>
</tr>
<tr>
<td>BRACKEN BSC</td>
<td>Bracken Basic Concepts Scale (Bracken, 1984)</td>
<td></td>
</tr>
<tr>
<td>BUS STORY</td>
<td>The Bus Story: a Test of Continuous Speech (Renfrew, 1969)</td>
<td></td>
</tr>
<tr>
<td>CELF</td>
<td>Clinical Evaluation of Language Function (Semel &amp; Wiig, 1981)</td>
<td></td>
</tr>
<tr>
<td>EAT</td>
<td>Edinburgh Articulation Test (Anthony, Bogle, Ingram &amp; Mclsaac, 1971)</td>
<td></td>
</tr>
<tr>
<td>GERMAN TWF</td>
<td>Test of Word Finding (German, 1986)</td>
<td></td>
</tr>
<tr>
<td>RENFREW APT</td>
<td>Action Picture Test (Renfrew, 1988)</td>
<td></td>
</tr>
<tr>
<td>TACL</td>
<td>Test for Auditory Comprehension of Language (Carrow, 1973)</td>
<td></td>
</tr>
<tr>
<td>TROG</td>
<td>Test for Reception of Grammar (Bishop, 1983)</td>
<td></td>
</tr>
<tr>
<td>TOKEN TEST</td>
<td>The Token Test for Children (DiSimoni, 1978)</td>
<td></td>
</tr>
<tr>
<td>WFVS:</td>
<td>Word-finding Vocabulary Scale (Renfrew, 1977)</td>
<td></td>
</tr>
<tr>
<td>TEST</td>
<td>LD1 (10;4)</td>
<td>LD2 (10;1)</td>
</tr>
<tr>
<td>------------</td>
<td>------------</td>
<td>------------</td>
</tr>
<tr>
<td>TROG</td>
<td>CA 9;7</td>
<td>CA 9;0</td>
</tr>
<tr>
<td></td>
<td>AE 5;6</td>
<td>AE 6;0</td>
</tr>
<tr>
<td>TOKEN TEST</td>
<td>CA 8;7</td>
<td>CA 9;0</td>
</tr>
<tr>
<td>RESPONSE</td>
<td>CA 9;7</td>
<td>CA 9;0</td>
</tr>
<tr>
<td></td>
<td>AE 8;4</td>
<td>AE 8;4</td>
</tr>
<tr>
<td></td>
<td>Centile 26</td>
<td>Centile 34</td>
</tr>
<tr>
<td>BPVS</td>
<td>CA 9;0</td>
<td>CA 9;0</td>
</tr>
<tr>
<td></td>
<td>AE 8;4</td>
<td>AE 8;4</td>
</tr>
<tr>
<td></td>
<td>Centile 26</td>
<td>Centile 34</td>
</tr>
<tr>
<td>BRACKEN</td>
<td>CA 9;7</td>
<td>CA 9;1</td>
</tr>
<tr>
<td>BCS</td>
<td>Info age 8;3</td>
<td>Info age 8;1</td>
</tr>
<tr>
<td>STORY</td>
<td>CA 9;7</td>
<td>CA 9;1</td>
</tr>
<tr>
<td></td>
<td>Info age 8;3</td>
<td>Info age 8;1</td>
</tr>
<tr>
<td>RENFREW</td>
<td>CA 9;7</td>
<td>CA 9;11</td>
</tr>
<tr>
<td>APT</td>
<td>AE 6;10</td>
<td>AE 6;3</td>
</tr>
<tr>
<td></td>
<td>Receptive lg 74</td>
<td>Receptive lg 74</td>
</tr>
<tr>
<td>CELF</td>
<td>CA 9;7</td>
<td>CA 9;11</td>
</tr>
<tr>
<td></td>
<td>AE 6;10</td>
<td>AE 6;3</td>
</tr>
<tr>
<td></td>
<td>Receptive lg 74</td>
<td>Receptive lg 74</td>
</tr>
<tr>
<td></td>
<td>Expres'sive lg 73</td>
<td>Expres'sive lg 74</td>
</tr>
<tr>
<td>WFVS</td>
<td>CA 9;7</td>
<td>CA 8;8</td>
</tr>
<tr>
<td></td>
<td>AE 7;0</td>
<td>AE 5;6</td>
</tr>
<tr>
<td>TEST</td>
<td>LD6 (8;8)</td>
<td>LD7 (10;5)</td>
</tr>
<tr>
<td>------</td>
<td>-----------</td>
<td>------------</td>
</tr>
<tr>
<td>TROG</td>
<td>CA 8;7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AE 8;0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Centile 38</td>
<td></td>
</tr>
<tr>
<td>BPVS</td>
<td>CA 7;8</td>
<td>CA 9;7</td>
</tr>
<tr>
<td></td>
<td>AE 6;3</td>
<td>AE 7;10</td>
</tr>
<tr>
<td></td>
<td>Centile 16</td>
<td>Centile 18</td>
</tr>
<tr>
<td>BOEM</td>
<td>CA 9;7</td>
<td>CA 9;0</td>
</tr>
<tr>
<td>TBS</td>
<td>Score 47/50</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BRACKEN</td>
<td>CA 7;8</td>
<td>CA 9;1</td>
</tr>
<tr>
<td>BCS</td>
<td>AE 4;6</td>
<td>AE 7;10 ceiling</td>
</tr>
<tr>
<td></td>
<td>Centile &lt;1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RENFREW</td>
<td>CA 7;8</td>
<td>CA 8;0</td>
</tr>
<tr>
<td>APT</td>
<td>Info age &lt;3;6</td>
<td>Info age 8;3</td>
</tr>
<tr>
<td></td>
<td>Gram age &lt;3;6</td>
<td>Gram age 8;3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CELF</td>
<td>CA 9;7</td>
<td>CA 8;0</td>
</tr>
<tr>
<td></td>
<td>AE 5;6</td>
<td>AE 7;6</td>
</tr>
<tr>
<td></td>
<td>Receptive Lg 54</td>
<td>Receptive Lg 95</td>
</tr>
<tr>
<td></td>
<td>Expres've Lg 50</td>
<td>Expres've Lg 84</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GERMAN</td>
<td>CA 9;7</td>
<td>CA 8;0</td>
</tr>
<tr>
<td>TWF</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CA 7;3</td>
<td>Centile 34</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WFVS</td>
<td>CA 8;7</td>
<td>CA 9;8</td>
</tr>
<tr>
<td></td>
<td>No. correct 32</td>
<td>AE 7;0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EAT</td>
<td>CA 7;8</td>
<td>CA 9;9</td>
</tr>
<tr>
<td></td>
<td>AE &lt;3;0</td>
<td>AE 5;5</td>
</tr>
</tbody>
</table>

363
<table>
<thead>
<tr>
<th>TEST</th>
<th>LD11 (11;2)</th>
<th>LD12 (10;10)</th>
<th>LD13 (10;11)</th>
<th>LD14 (8;6)</th>
<th>LD15 (10;8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TROG</td>
<td>CA 10;4</td>
<td>CA 10;6</td>
<td>CA 10;1</td>
<td>CA 8;6</td>
<td>CA 9;9</td>
</tr>
<tr>
<td></td>
<td>AE 8;0</td>
<td>AE 10;0</td>
<td>AE 5;4</td>
<td>AE 10;0</td>
<td>AE 5;9</td>
</tr>
<tr>
<td></td>
<td>Centile 5</td>
<td>Centile 10</td>
<td>Centile &lt;1</td>
<td>Centile 75</td>
<td>Centile 3</td>
</tr>
<tr>
<td>TACL</td>
<td></td>
<td>CA 10;1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>AE 7;0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BPVS</td>
<td>CA 10;4</td>
<td>CA 10;6</td>
<td>CA 10;1</td>
<td>CA 9;9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AE 7;2</td>
<td>AE 10;4</td>
<td>AE 5;6</td>
<td>AE 6;4</td>
<td></td>
</tr>
<tr>
<td>BOEHM</td>
<td></td>
<td></td>
<td></td>
<td>CA 9;9</td>
<td>CA 9;9</td>
</tr>
<tr>
<td>TBC</td>
<td></td>
<td></td>
<td></td>
<td>Score 37/50</td>
<td></td>
</tr>
<tr>
<td>BRACKEN</td>
<td></td>
<td></td>
<td>CA 8;6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BCS</td>
<td></td>
<td></td>
<td>AE 7;10</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Centile 61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CELF</td>
<td>CA 10;4</td>
<td>CA 10;6</td>
<td>CA 8;6</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AE 5;8</td>
<td>AE 10;6</td>
<td>AE 6;1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Receptive lg 101</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Expressive lg 97</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 71 below presents the personal details pertaining to each of the SP subjects in this study. It also includes the diagnosis or description of each child's difficulties as provided by SLTs or Autism Outreach Service teachers.

<table>
<thead>
<tr>
<th>SUBJECT</th>
<th>GENDER</th>
<th>AGE</th>
<th>DIAGNOSIS / DESCRIPTION OF PROBLEMS PROVIDED BY SLT OR AUTISM OUTREACH SERVICE TEACHER</th>
</tr>
</thead>
<tbody>
<tr>
<td>SP1</td>
<td>M</td>
<td>7;2</td>
<td>Semantic-pragmatic difficulties</td>
</tr>
<tr>
<td>SP2</td>
<td>M</td>
<td>10;9</td>
<td>Semantic-pragmatic disorder</td>
</tr>
<tr>
<td>SP3</td>
<td>M</td>
<td>10;1</td>
<td>Asperger Syndrome / high-functioning autism</td>
</tr>
<tr>
<td>SP4</td>
<td>M</td>
<td>11;3</td>
<td>Semantic-pragmatic disorder</td>
</tr>
<tr>
<td>SP5</td>
<td>M</td>
<td>9;5</td>
<td>Semantic-pragmatic disorder</td>
</tr>
<tr>
<td>SP6</td>
<td>M</td>
<td>10;2</td>
<td>Semantic-pragmatic difficulties</td>
</tr>
<tr>
<td>SP7</td>
<td>M</td>
<td>8;4</td>
<td>Semantic-pragmatic difficulties</td>
</tr>
<tr>
<td>SP8</td>
<td>M</td>
<td>9;1</td>
<td>Specific language impairment (semantic/pragmatic)</td>
</tr>
<tr>
<td>SP9</td>
<td>M</td>
<td>9;7</td>
<td>Specific language impairment (semantic/pragmatic)</td>
</tr>
<tr>
<td>SP10</td>
<td>M</td>
<td>11;3</td>
<td>Semantic-pragmatic difficulties</td>
</tr>
<tr>
<td>SP11</td>
<td>M</td>
<td>9;6</td>
<td>Semantic-pragmatic difficulties</td>
</tr>
<tr>
<td>SP12</td>
<td>M</td>
<td>8;0</td>
<td>Pragmatic difficulty</td>
</tr>
<tr>
<td>SP13</td>
<td>M</td>
<td>10;4</td>
<td>Semantic-pragmatic disorder</td>
</tr>
<tr>
<td>SP14</td>
<td>M</td>
<td>8;9</td>
<td>Semantic-pragmatic difficulties</td>
</tr>
<tr>
<td>SP15</td>
<td>M</td>
<td>10;7</td>
<td>Semantic &amp; pragmatic difficulties</td>
</tr>
<tr>
<td>SP16</td>
<td>M</td>
<td>11;2</td>
<td>Semantic &amp; pragmatic difficulties</td>
</tr>
<tr>
<td>SP17</td>
<td>M</td>
<td>10;8</td>
<td>Semantic-pragmatic disorder</td>
</tr>
<tr>
<td>SP18</td>
<td>M</td>
<td>11;1</td>
<td>Semantic-pragmatic disorder affecting reception &amp; expression</td>
</tr>
<tr>
<td>SP19</td>
<td>M</td>
<td>10;2</td>
<td>High-functioning autism (query Pathological Demand-Avoidance Syndrome)</td>
</tr>
<tr>
<td>SP20</td>
<td>M</td>
<td>7;10</td>
<td>Asperger Syndrome</td>
</tr>
<tr>
<td>SP21</td>
<td>M</td>
<td>8;7</td>
<td>Asperger Syndrome</td>
</tr>
<tr>
<td>SP22</td>
<td>F</td>
<td>9;0</td>
<td>High-functioning autism / Asperger syndrome</td>
</tr>
<tr>
<td>SP23</td>
<td>M</td>
<td>6;5</td>
<td>Asperger syndrome</td>
</tr>
<tr>
<td>SP24</td>
<td>M</td>
<td>8;5</td>
<td>Asperger syndrome</td>
</tr>
<tr>
<td>SP25</td>
<td>F</td>
<td>10;6</td>
<td>Asperger syndrome</td>
</tr>
<tr>
<td>SP26</td>
<td>M</td>
<td>7;1</td>
<td>High-functioning autism / Asperger syndrome</td>
</tr>
</tbody>
</table>
APPENDIX C: THE CHECKLIST

1. THE CHECKLIST

The Checklist comprised seven sections labelled A to G. Sections A to F were informed primarily by Wing's (1991) continua relating to the central features of autism. These describe severe to mild manifestations of impairment in the areas of social interaction, social communication and social imagination. They also encompass repetitive patterns of behaviour, unusual responses to sensory stimuli and unusual movements. Section G comprised a list of language features associated with semantic-pragmatic difficulties in the literature. This section was informed chiefly by Andersen-Wood and Smith (in press).

Sections A, B, C and D each required the SLT or teacher to circle the description most typical of the child concerned. Sections E, F and G involved the circling of all features considered to be characteristic of the child.

A checklist was completed for each child in the LD and SP groups by the child's SLT, Autism Outreach Service teacher or a school teacher who knew the child well. While an effort was made to exemplify potentially confusing terminology on the Checklist, it should nevertheless be borne in mind that individual SLTs and teachers may have differed in their interpretations of the features listed. The thresholds applied by these professionals with respect to individual features may also have varied. It is further possible that the behaviour typical of a particular child in interaction with his/her SLT or teacher may differ from that in other contexts. It is therefore necessary to interpret the results of the Checklist cautiously.

A copy of the Checklist follows.
CHILD'S NAME:

CURRENT DIAGNOSIS (IF KNOWN):

YOUR NAME: RELATIONSHIP TO CHILD:

FOR EACH OF THE FOLLOWING, PLEASE CIRCLE THE ONE DESCRIPTION WHICH BEST SUITS THIS CHILD. WHERE EXAMPLES ARE GIVEN, PLEASE UNDERLINE THOSE WHICH APPLY TO THIS CHILD.

A. Social interaction

1. aloof and indifferent
2. approaches for physical needs only
3. passively accepts approaches
4. makes bizarre approaches
5. none of the above

B. Social communication

1. no communication
2. communicates needs only
3. replies if approached by someone else
4. spontaneous, one-sided communication
5. none of the above

C. Unusual responses to sensory stimuli (eg. over-sensitive to sound; fascinated by lights, touches and tastes; smells, objects or people; indifferent to pain or cold)

1. very marked
2. marked
3. occasional
4. absent

D. Movements (eg. flaps, jumps, rocks, tip-toe walks, postures)

1. very marked
2. marked
3. occasional
4. absent
PLEASE CIRCLE ALL THE DESCRIPTIONS BELOW WHICH APPLY TO THIS CHILD. UNDERLINE APPLICABLE EXAMPLES OR ADD YOUR OWN.

E. Social imagination

1. no imagination
2. copies others in a mechanical way
3. uses dolls or other toys correctly but in a limited, uncreative, repetitive way
4. acts out one/very few themes (eg. Batman) repetitively
5. uses other children as "mechanical aids" rather than as partners in play

F. Repetitive pattern of self-chosen activities

1. simple, bodily-directed activities (eg. face slapping, self-injury)
2. simple, object-directed activities (eg. switching taps or lights on and off, spinning objects)
3. activities involving complex routines, manipulation of objects or movements (eg. bedtime ritual, lining up objects, attachment to objects, whole-body movements)
4. verbal or abstract activities (eg. obsessive interest in time-tables or movements of the planets, repetitive questioning)

G. Language

1. no language
2. limited language - mostly echolalic
3. abnormal prosody (eg. inappropriate or unusual volume, stress, rhythm, intonation or speed of speech)
4. long-winded
5. pedantic
6. overly formal
7. dominates conversations
8. very passive in conversations
9. poor turn-taking in conversation (eg. doesn't take his/her turn, does not allow others to have a turn)
10. answers questions tangentially
11. asks questions repetitively
12. does not ask for clarification when not understanding
13. does not give adequate clarification of own meaning when asked to
14. gives more information than is necessary
15. gives less information than is necessary
16. seems to lack awareness of what conversational partner can or can't be expected to know about
17. pays attention to irrelevant, non-salient details
18. difficulty maintaining a conversational topic
19. tends to end conversations/interactions very abruptly
20. sticks obsessively to one/few conversational topics
21. has rigid concept boundaries
22. uses jargon
23. shows signs of word-finding difficulties (eg. hesitates, makes up words, talks around a word, makes false starts)
24. overuses clichés or stereotyped utterances
25. perseverates
26. uses full sentence forms rather than ellipsis (eg. Adult: "Have you brought your book?" Child: "I have brought my book", rather than just "Yes")
27. uses pronoun "you" for "I" (eg. Adult: "Do you want a biscuit?" Child: "You want a biscuit")
28. uses pronouns (eg. "them" and "it") without making it clear who's being referred to
29. has poor non-verbal communication (eg. doesn't use eyes, gesture or facial expression communicatively)
30. does not understand non-verbal communication
31. makes socially inappropriate remarks
32. interprets figurative language, irony, humour literally
33. does not understand indirect speech acts (eg. Adult: "Do you have the time?" Child: "Yes" rather than telling time

THANK-YOU VERY MUCH FOR TAKING THE TIME TO FILL IN THIS CHECKLIST.

DEBRA KERBEL
2. RESULTS OF THE CHECKLIST

The results of the Checklist are presented graphically in Figures 10 and 11 on the following pages. Shaded blocks indicate features circled for the relevant child. In the interests of clarity, the options "none of the above" and "absent" are not included for sections A, B, C and D, respectively. Each of the shaded blocks therefore represents some degree of impairment. In some instances, more than one feature was circled in sections A, B, C or D. In these cases, only the most severe manifestation marked as present is shaded on the chart.
FIGURE 10: CHECKLIST RESULTS - SECTIONS A-F
Kth neighbour discriminant analysis performed on the Checklist data was able to assign all of the LD and SP children to the correct groups. It was further able to discriminate between the "SPD" and "ASP" subgroups with 100% accuracy.

Interestingly, section G (language features), on its own, was found to be as powerful a discriminator as was the Checklist in its entirety. On the basis of this information alone, Kth neighbour discriminant analysis was able to assign all of the LD, "SPD" and "ASP" subjects to the correct groups. Furthermore, the LD group was found to differ significantly from each of the SP subgroups on the total number of language features scored positive for each child (Mann-Whitney Test: LD vs "SPD": U = 10, p = 0.0000; LD vs "ASP": U = 7.5, p = 0.0002). The difference between the "SPD" and "ASP" subgroups in the total number of language features scored positive was not significant. Table 72 below presents the proportion of each group and subgroup scoring positive for each of the language features.

Sections A to F collectively also discriminated groups to a considerable degree. SP6, SP14, SP15, SP16 and SP17 were, however, grouped with the LD subjects on the basis of these data. Between-groups differences were further explored by scoring one point for each section (A to F) in which the child had been noted to display some degree of impairment. A maximum of six points was thus possible for this index of "autistic continuum features". Notably, the LD group scored a mean of 0.27 (s.d. = 0.59) as compared with a mean of 2.18 (s.d. = 1.91) for the "SPD" subgroup, and a mean of 4.78 (s.d. = 1.3) for the "ASP" subgroup. Each of these pairwise differences was found to be highly significant on a series of Mann-Whitney Tests (LD vs "SPD": U = 44, p = 0.0007; LD vs "ASP": U = 0, p = 0.0000; "SPD" vs "ASP": U = 21, p = 0.0025). This result, not surprisingly, indicates that autistic features were more prevalent among the SP than LD subjects. Interestingly, it also suggests that children who appear to manifest fewer autistic features are more likely to be described as having semantic-pragmatic difficulties than Asperger
syndrome or high-functioning autism. The possibility cannot be ruled out, however, that similar behaviour may be interpreted differently depending on whether the child is considered to have semantic-pragmatic difficulties or Asperger syndrome / high functioning autism.
<table>
<thead>
<tr>
<th>FEATURE</th>
<th>LD GROUP</th>
<th>SP GROUP</th>
<th>&quot;SPD&quot; SUBGROUP</th>
<th>&quot;ASP&quot; SUBGROUP</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1</td>
<td>13.3%</td>
<td>30.7%</td>
<td>17.6%</td>
<td>55.6%</td>
</tr>
<tr>
<td>G2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3</td>
<td>6.7%</td>
<td>15.4%</td>
<td>17.6%</td>
<td>11.1%</td>
</tr>
<tr>
<td>G4</td>
<td>26.7%</td>
<td>42.3%</td>
<td>47.1%</td>
<td>33.3%</td>
</tr>
<tr>
<td>G5</td>
<td>6.7%</td>
<td>19.2%</td>
<td>17.6%</td>
<td>22.2%</td>
</tr>
<tr>
<td>G6</td>
<td>50%</td>
<td>76.9%</td>
<td>82.4%</td>
<td>66.7%</td>
</tr>
<tr>
<td>G7</td>
<td>6.7%</td>
<td>50%</td>
<td>41.2%</td>
<td>66.7%</td>
</tr>
<tr>
<td>G8</td>
<td>53.3%</td>
<td>50%</td>
<td>47.1%</td>
<td>55.6%</td>
</tr>
<tr>
<td>G9</td>
<td>6.7%</td>
<td>19.2%</td>
<td>17.6%</td>
<td>22.2%</td>
</tr>
<tr>
<td>G10</td>
<td>53.3%</td>
<td>50%</td>
<td>47.1%</td>
<td>55.6%</td>
</tr>
<tr>
<td>G11</td>
<td>13.3%</td>
<td>30.7%</td>
<td>17.6%</td>
<td>55.6%</td>
</tr>
<tr>
<td>G12</td>
<td>6.7%</td>
<td>19.2%</td>
<td>17.6%</td>
<td>22.2%</td>
</tr>
<tr>
<td>G13</td>
<td>50%</td>
<td>76.9%</td>
<td>82.4%</td>
<td>66.7%</td>
</tr>
<tr>
<td>G14</td>
<td>6.7%</td>
<td>19.2%</td>
<td>17.6%</td>
<td>22.2%</td>
</tr>
<tr>
<td>G15</td>
<td>6.7%</td>
<td>19.2%</td>
<td>17.6%</td>
<td>22.2%</td>
</tr>
<tr>
<td>G16</td>
<td>50%</td>
<td>76.9%</td>
<td>82.4%</td>
<td>66.7%</td>
</tr>
<tr>
<td>G17</td>
<td>6.7%</td>
<td>19.2%</td>
<td>17.6%</td>
<td>22.2%</td>
</tr>
<tr>
<td>G18</td>
<td>50%</td>
<td>76.9%</td>
<td>82.4%</td>
<td>66.7%</td>
</tr>
<tr>
<td>G19</td>
<td>6.7%</td>
<td>19.2%</td>
<td>17.6%</td>
<td>22.2%</td>
</tr>
<tr>
<td>G20</td>
<td>50%</td>
<td>76.9%</td>
<td>82.4%</td>
<td>66.7%</td>
</tr>
<tr>
<td>G21</td>
<td>6.7%</td>
<td>19.2%</td>
<td>17.6%</td>
<td>22.2%</td>
</tr>
<tr>
<td>G22</td>
<td>50%</td>
<td>76.9%</td>
<td>82.4%</td>
<td>66.7%</td>
</tr>
<tr>
<td>G23</td>
<td>6.7%</td>
<td>19.2%</td>
<td>17.6%</td>
<td>22.2%</td>
</tr>
<tr>
<td>G24</td>
<td>50%</td>
<td>76.9%</td>
<td>82.4%</td>
<td>66.7%</td>
</tr>
<tr>
<td>G25</td>
<td>6.7%</td>
<td>19.2%</td>
<td>17.6%</td>
<td>22.2%</td>
</tr>
<tr>
<td>G26</td>
<td>50%</td>
<td>76.9%</td>
<td>82.4%</td>
<td>66.7%</td>
</tr>
<tr>
<td>G27</td>
<td>6.7%</td>
<td>19.2%</td>
<td>17.6%</td>
<td>22.2%</td>
</tr>
<tr>
<td>G28</td>
<td>50%</td>
<td>76.9%</td>
<td>82.4%</td>
<td>66.7%</td>
</tr>
<tr>
<td>G29</td>
<td>6.7%</td>
<td>19.2%</td>
<td>17.6%</td>
<td>22.2%</td>
</tr>
<tr>
<td>G30</td>
<td>6.7%</td>
<td>19.2%</td>
<td>17.6%</td>
<td>22.2%</td>
</tr>
<tr>
<td>G31</td>
<td>6.7%</td>
<td>19.2%</td>
<td>17.6%</td>
<td>22.2%</td>
</tr>
<tr>
<td>G32</td>
<td>6.7%</td>
<td>19.2%</td>
<td>17.6%</td>
<td>22.2%</td>
</tr>
<tr>
<td>G33</td>
<td>6.7%</td>
<td>19.2%</td>
<td>17.6%</td>
<td>22.2%</td>
</tr>
</tbody>
</table>
APPENDIX D: MOST FREQUENTLY-OCCURRING IDIOMS IN TEACHER AND TELEVISION LANGUAGE SAMPLES

As noted in Chapter 4, a list of the most commonly-occurring idioms in the teacher and television language sample was compiled before the story was developed. The following idioms occurred in at least half of the morning-to-lunchtime and/or afternoon television sessions recorded, and in at least half of the mainstream and/or language unit teacher sessions.

1. all right
2. carry on
3. come back
4. come on
5. find out
6. first of all
7. get on with
8. have a go
9. in a second/minute/moment
10. just a minute
11. never mind
12. oh dear
13. wait a minute
APPENDIX E: THE STORY

1. MAIN STUDY VERSION

The story below was used in the main study reported in this document. Capital letters without underlining denote idioms with a clearly literal counterpart. Underlined capital letters denote idioms with a borderline literal/idiomatic counterpart.

One day, Sally and Rover were playing in the garden. Suddenly, Rover ran into the road. A green motorbike was driving on the road. The bike rider didn't see Rover and RAN OVER him. The bike rider TOOK OFF without stopping. Rover lay in the road. Sally's mum ran to Rover. Sally GAVE her mum A HAND to carry Rover to the white car. They put him inside. Sally's mum drove to the vet's office. She carried Rover into the room. She put Rover on the floor. There was a heap of bandages on the vet's table. The vet MOVED them UP so that there was enough space for Rover. Sally's mum put Rover on the table.

The vet examined Rover's head. Then he GOT ON TO Rover's legs. Rover's leg was broken.

Sally and her mum left the vet's office. They saw the bike rider walking in the road. Sally chased him. The bike rider started running. Suddenly the bike rider stopped ON THE SPOT. There was a policeman in the road. Sally TURNED the bike rider OVER to the policeman. The policeman took the bike rider to the police station. He put him in a cell and KEPT him UNDER LOCK AND KEY. The bike rider wanted his motorbike at the police station. The policeman went to PICK UP it up.

The next day, Sally's mum drove to the vet's office. She brought Rover home. Sally put Rover into his basket. Sally KEPT AN EYE ON him all day.
That night, the policeman DROPPED IN to visit Sally and Rover. He patted Rover. Rover was very happy. Rover WENT ROUND everyone and licked them all.

2. PILOT STUDY VERSION

The story below is the version used in the pilot study. Underlining denotes idioms removed from the story for the main study.

One day, Sally decided to brush Rover's hair. Sally took Rover and the brush into the garden. Rover did not like having his hair brushed. He wriggled away from Sally. He ran over the flowers and into the road. A green motorbike was driving on the road. The bike rider didn't see Rover and ran over him. The bike rider was scared and he took off without stopping. Rover lay in the road. Sally's mum ran to Rover and lifted him in her arms. Sally gave her mum a hand to carry Rover to their white car. They put him inside. Sally's mum drove to the vet's office and carried Rover into the room. There was a heap of bandages on the vet's table. They took up the whole table. The vet moved them up so that there was enough space for Rover. Sally's mum put Rover on the table.

The vet examined Rover's head. Then he got on to Rover's legs. Rover's leg was broken. The vet gave Rover a bowl of water. He drank the water, but he felt very ill and he threw it up on the table.

Sally and her mum left Rover at the vet's office. They got into their car to go home. While they were driving, they saw the green motorbike that hit the dog. They followed the bike as fast as they could. The bike rider was frightened. He knew he should have stopped when he hit the dog. Suddenly, his motorbike broke
down and would not go any further. The bike rider started running. Sally chased him across the street. Suddenly a policeman appeared. The bike rider stopped on the spot. Sally ran to the bike rider and waved to the policeman. The policeman listened to Sally's story and then Sally turned the bike rider over to the policeman. The policeman took the bike rider to the police station. He put him in a cell and kept him under lock and key.

The next day, Sally's mum drove to the vet's office to pick up Rover. She brought him home. Sally put Rover into his basket and kept an eye on him all afternoon. Soon Rover's leg healed and he ran around in the garden.

One day, the policeman dropped in to visit Sally and Rover. He patted Rover. Rover was very happy and went round everyone and licked them all.

3. AMENDMENTS MADE TO THE STORY PRIOR TO THE MAIN STUDY

Following a between-idioms analysis conducted on the pilot data, the four idioms listed below were removed from the story.

(a) "TOOK UP"

This expression was removed on the grounds that the prior arrangement of the bandages across the table gave rise to a high proportion of "appropriate no action" codes. In this context, children may have assumed that no action was required for the sentence, "They took up the whole table".
(b) "THREW UP"

It was recognised that in order for the methodology to be meaningful, the idioms included should be understood by the majority of mainstream children. Since the expression "threw up" was appropriately demonstrated by fewer than half of the mainstream subjects, this idiom was removed in the main study.

(c) "BROKE DOWN"

Although the Playmobil motorbike could be broken down into several components, it was likely that many children would fail to realise this. Others may have felt that it was not permitted to break a stranger's toys or might have lacked the necessary fine motor skills. This expression was therefore removed on the grounds that the nature of the prop introduced a bias against a literal demonstration.

(d) "RAN AROUND"

This expression was removed on the grounds that it was difficult to distinguish between its literal and idiomatic demonstration.
Before the story was played for the first time, the children were instructed simply to listen. As the story was played, the author noted the degree to which each child scanned the play set visually. "Full" scanning was recorded for children who appeared to be looking at the appropriate props or parts of the set for approximately 75% or more of the story. "Partial" scanning was recorded for children who looked at the appropriate props or parts of the set intermittently, but for more than 25% of the story. "None" was recorded for children who appeared to be scanning the set during approximately 25% or less of the story.

The number of trials attempted by each child before the play task was also recorded. Each child completed at least two trials. Additional trials were introduced if the child failed to act these out accurately. The task commenced when two consecutive trials had been accurately enacted.

<table>
<thead>
<tr>
<th>SUBJECT</th>
<th>SCANNING</th>
<th>NO. OF TRIALS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MI1</td>
<td>partial</td>
<td>3</td>
</tr>
<tr>
<td>MI2</td>
<td>full</td>
<td>2</td>
</tr>
<tr>
<td>MI3</td>
<td>full</td>
<td>2</td>
</tr>
<tr>
<td>MI4</td>
<td>full</td>
<td>3</td>
</tr>
<tr>
<td>MI5</td>
<td>full</td>
<td>2</td>
</tr>
<tr>
<td>MI6</td>
<td>partial</td>
<td>2</td>
</tr>
<tr>
<td>MI7</td>
<td>full</td>
<td>2</td>
</tr>
<tr>
<td>MI8</td>
<td>full</td>
<td>2</td>
</tr>
<tr>
<td>MI9</td>
<td>full</td>
<td>2</td>
</tr>
<tr>
<td>MI10</td>
<td>full</td>
<td>2</td>
</tr>
<tr>
<td>MI11</td>
<td>full</td>
<td>4</td>
</tr>
<tr>
<td>MI12</td>
<td>full</td>
<td>3</td>
</tr>
<tr>
<td>MI13</td>
<td>partial</td>
<td>2</td>
</tr>
<tr>
<td>MI14</td>
<td>full</td>
<td>2</td>
</tr>
<tr>
<td>MI15</td>
<td>full</td>
<td>2</td>
</tr>
</tbody>
</table>
### Table 74: MJ Group - Play Set Scanning and Number of Play Task Trials

<table>
<thead>
<tr>
<th>SUBJECT</th>
<th>SCANNING</th>
<th>NO. OF TRIALS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MJ1</td>
<td>full</td>
<td>2</td>
</tr>
<tr>
<td>MJ2</td>
<td>full</td>
<td>2</td>
</tr>
<tr>
<td>MJ3</td>
<td>full</td>
<td>2</td>
</tr>
<tr>
<td>MJ4</td>
<td>full</td>
<td>2</td>
</tr>
<tr>
<td>MJ5</td>
<td>full</td>
<td>2</td>
</tr>
<tr>
<td>MJ6</td>
<td>partial</td>
<td>2</td>
</tr>
<tr>
<td>MJ7</td>
<td>partial</td>
<td>2</td>
</tr>
<tr>
<td>MJ8</td>
<td>full</td>
<td>2</td>
</tr>
<tr>
<td>MJ9</td>
<td>full</td>
<td>2</td>
</tr>
<tr>
<td>MJ10</td>
<td>full</td>
<td>2</td>
</tr>
<tr>
<td>MJ11</td>
<td>none</td>
<td>2</td>
</tr>
<tr>
<td>MJ12</td>
<td>not recorded</td>
<td>2</td>
</tr>
<tr>
<td>MJ13</td>
<td>partial</td>
<td>2</td>
</tr>
<tr>
<td>MJ14</td>
<td>full</td>
<td>2</td>
</tr>
<tr>
<td>MJ15</td>
<td>full</td>
<td>2</td>
</tr>
</tbody>
</table>

### Table 75: LD Group - Play Set Scanning and Number of Play Task Trials

<table>
<thead>
<tr>
<th>SUBJECT</th>
<th>SCANNING</th>
<th>NO. OF TRIALS</th>
</tr>
</thead>
<tbody>
<tr>
<td>LD1</td>
<td>full</td>
<td>2</td>
</tr>
<tr>
<td>LD2</td>
<td>full</td>
<td>2</td>
</tr>
<tr>
<td>LD3</td>
<td>full</td>
<td>5</td>
</tr>
<tr>
<td>LD4</td>
<td>none</td>
<td>2</td>
</tr>
<tr>
<td>LD5</td>
<td>partial</td>
<td>7</td>
</tr>
<tr>
<td>LD6</td>
<td>partial</td>
<td>2</td>
</tr>
<tr>
<td>LD7</td>
<td>partial</td>
<td>2</td>
</tr>
<tr>
<td>LD8</td>
<td>full</td>
<td>2</td>
</tr>
<tr>
<td>LD9</td>
<td>full</td>
<td>2</td>
</tr>
<tr>
<td>LD10</td>
<td>full</td>
<td>2</td>
</tr>
<tr>
<td>LD11</td>
<td>full</td>
<td>3</td>
</tr>
<tr>
<td>LD12</td>
<td>full</td>
<td>2</td>
</tr>
<tr>
<td>LD13</td>
<td>partial</td>
<td>2</td>
</tr>
<tr>
<td>LD14</td>
<td>partial</td>
<td>3</td>
</tr>
<tr>
<td>LD15</td>
<td>full</td>
<td>5</td>
</tr>
</tbody>
</table>
### TABLE 76: SP GROUP - PLAY SET SCANNING AND NUMBER OF PLAY TASK TRIALS

<table>
<thead>
<tr>
<th>SUBJECT</th>
<th>SCANNING</th>
<th>NO. OF TRIALS</th>
</tr>
</thead>
<tbody>
<tr>
<td>SP1</td>
<td>full</td>
<td>3</td>
</tr>
<tr>
<td>SP2</td>
<td>full</td>
<td>3</td>
</tr>
<tr>
<td>SP3</td>
<td>partial</td>
<td>5</td>
</tr>
<tr>
<td>SP4</td>
<td>full</td>
<td>2</td>
</tr>
<tr>
<td>SP5</td>
<td>full</td>
<td>2</td>
</tr>
<tr>
<td>SP6</td>
<td>full</td>
<td>2</td>
</tr>
<tr>
<td>SP7</td>
<td>partial</td>
<td>3</td>
</tr>
<tr>
<td>SP8</td>
<td>full</td>
<td>4</td>
</tr>
<tr>
<td>SP9</td>
<td>partial</td>
<td>3</td>
</tr>
<tr>
<td>SP10</td>
<td>none</td>
<td>3</td>
</tr>
<tr>
<td>SP11</td>
<td>partial</td>
<td>4</td>
</tr>
<tr>
<td>SP12</td>
<td>partial</td>
<td>3</td>
</tr>
<tr>
<td>SP13</td>
<td>none</td>
<td>3</td>
</tr>
<tr>
<td>SP14</td>
<td>full</td>
<td>4</td>
</tr>
<tr>
<td>SP15</td>
<td>full</td>
<td>2</td>
</tr>
<tr>
<td>SP16</td>
<td>full</td>
<td>2</td>
</tr>
<tr>
<td>SP17</td>
<td>partial</td>
<td>2</td>
</tr>
<tr>
<td>SP18</td>
<td>partial</td>
<td>2</td>
</tr>
<tr>
<td>SP19</td>
<td>full</td>
<td>2</td>
</tr>
<tr>
<td>SP20</td>
<td>partial</td>
<td>4</td>
</tr>
<tr>
<td>SP21</td>
<td>full</td>
<td>2</td>
</tr>
<tr>
<td>SP22</td>
<td>none</td>
<td>5</td>
</tr>
<tr>
<td>SP23</td>
<td>none</td>
<td>3</td>
</tr>
<tr>
<td>SP24</td>
<td>partial</td>
<td>3</td>
</tr>
<tr>
<td>SP25</td>
<td>full</td>
<td>2</td>
</tr>
<tr>
<td>SP26</td>
<td>partial</td>
<td>4</td>
</tr>
</tbody>
</table>

As discussed in Chapter 6, the fact that most of the children appeared to be scanning the play set, at least intermittently, during the first reading of the story suggests that they did make the connection between the characters and objects of the story and their concrete referents. That most children required no more than three trials further supports the view that these children did grasp the requirements of the play task.
In the interests of consistency, the prompts used during the play task and definition task were standardized.

1. PLAY TASK

If the child failed to act out the first or second sentence of the story and/or repeated the sentence after it was played, the author prompted with the words, "You do it."

If the child failed to demonstrate a sentence (other than the first or second sentence), looked at the author apparently seeking help, asked "What does that mean?" or otherwise requested clarification, the author responded with the words, "Let's listen again and then you do it." The relevant sentence was then replayed.

If the child asked the author how to demonstrate a sentence or suggested an action with questioning intonation, the author responded with the sentence, "Do whatever the story says."

If a prop was damaged (to a minor degree) during the play task, the author responded with the sentence, "Don't worry about that."

2. DEFINITION TASK

During the children's definitions, the author provided back channel support in the form of "uhu", produced with neutral intonation. This was also used in response to incorrect or incomplete definitions and when the child paused during a definition.

If the child responded to the author's question with prolonged
silence, a tangential response, a repeat of the whole idiom or of the verb followed by a different preposition, the author prompted with the question, "So what does (ran over) mean?"

If the child responded to the author's question with a non-verbal demonstration, the author prompted with the words, "Can you say it?"

If the child's response was inaudible or unintelligible, the child was asked to "Say that again."
APPENDIX H: EXAMPLE OF DATA TRANSCRIPT AND CODING

The following examples all relate to the idiom "pick up". This occurred in the context, "The bike rider wanted his motorbike at the police station. The policeman went to pick it up."

1. PLAY TASK

In the transcript of play task actions, the symbol # denotes the position of the relevant props immediately prior to the idiom being heard on tape. The symbol * denotes the child's action and comments from the time the idiom is heard. The abbreviation, "v/o", is used for "vet's office". "P/s" stands for "police station". "HA" ("hands away") refers to the child drawing his/her hands away from the props.

MJ1 #  bike parked between v/o and p/s; policeman in p/s; HA
  * stands toppled Sally, walks policeman to bike, holds policeman's hand on handlebar, policeman walks alongside bike with hand on handlebar into p/s, parks bike on edge of p/s, stands policeman in p/s, HA

SP15 #  bike lies toppled between v/o and p/s; IDIOMATIC
        policeman stands in p/s; moving bike rider in cell
  * knocks p/s as lifts policeman, "oops", seats policeman on bike, drives policeman into p/s, removes policeman from bike, parks bike in p/s, stands policeman in p/s, HA

LD5 #  policeman stands in p/s; bike parked at the edge of p/s; touching front wheel of the bike
  * HA, moves the bike with its wheels on the ground and parks the bike in the p/s, HA

386
SP11 # standing Sally; policeman stands outside p/s; bike lies at edge of v/o
* drives bike into the p/s, walks policeman to the bike, begins to seat policeman on the bike, stands policeman at the edge of the p/s, HA

MI15 # bike parked between v/o and p/s; holds policeman in the air
* takes policeman to bike, holds bike up against the policeman's chest, lowers and parks the bike in the same spot, holds the policeman standing next to the bike

SP21 # bike lies at the edge of the p/s; policeman lying on p/s floor
* pushes the policeman towards the bike causing the cell door to open, pulls the bike and policeman in lying position out of the p/s and closes the cell door, holds the policeman upright with the bike held vertically against him for four seconds outside the p/s

MI2 # bike parked between v/o and p/s; policeman toppled in p/s; holding bike rider standing in cell
* brings policeman to bike, lifts bike into policeman's arms, policeman walks into p/s with bike pushed up against him, parks bike in p/s, stands policeman in p/s, HA

SP7 # policeman stands in p/s; bike lies topples outside p/s; HA
* takes policeman to bike, lifts policeman and bike in one hand to p/s, lowers bike on its side on p/s floor, stands policeman in p/s, HA
lifting bike from near the house (off camera); policeman lies toppled in the p/s lowers bike near p/s, policeman walks to bike, holds the policeman by the head and the bike with wheels off the ground and walks them to the road outside the house, lowers bike on its side, holds the policeman above the road, "mm hallo hallo hallo hallo hallo (motorbike ...)", holds the policeman standing as adjusts policeman's posture, holds bike stationary alongside the policeman in the middle of the set

2. DEFINITION TASK

In the definition task transcripts, "C" denotes "child" and "A" denotes "author". A slash (/) denotes a short pause of between 1.5 and two seconds in length. "Silence" refers to a pause of between 2.5 and seven seconds. "Long silence" refers to a pause longer than 7.5 seconds. Bold type denotes stressed words.

SP25 fetch something

SP15 go and get it

SP10 C: um take pick it up and take it to someone
A: uhu so what does pick up mean?
C: pick up means pick it up
A: uhu
C: take it to someone

LD3 um to get it

SP3 C: means you hold it high
A: uhu
C: and if you um don't hold it high it means you low down
A: uhu
C: but if you do it you go and then if you be good as gold you out of jail
LD9  C: um you're just picked something up from the floor
A: uhu
C: and you holding it

SP18  C: pick something pick the motorbike up with his hands
A: uhu so what does pick up mean?
C: pick it up

SP26  C: like if you're picking up something and if it's dirty you'll drop it
A: uhu
C: so don't need to pick it up just check if it needs be- pick it up or something
A: uhu / so what does pick up mean?
C: pick up means like you're picking up a real thing
A: uhu
C: if you're sure that you want to pick up one thing pick it up if you don't then don't

SP22  C: (long silence)
A: the policeman went to pick it up /
what do you think pick up means?
C: to put in prison

MI7  C: it was like on the floor and then he picked it up
A: uhu so what does pick up mean?
C: he got it up
A: uhu
C: straight

LD4  C: it's when people pick a sandwich up or pick um / the motorbike up and
A: so what does pick up mean?
C: um (long silence) ah (silence) don't know

LD5  C: pu- put it in the police station
A: uhu
C: and / and shut the door
A: uhu
C: (silence) and tell the policeman he can't drive it any more

389
SP7  C: picking up
A: uhu so what does pick up mean?
C: picking things up

LD13  C: because it might get nick
A: uhu
C: because he he might want because he
    want to pick up because it doesn't
    want to get lost
A: uhu so what does pick up mean?
C: pick up his bike

SP23  umm / give it to him?
The following guidelines were used in the coding and second-coding of the play task and definition task data. Items numbered (a), (b), (c) etc. describe acceptable alternative demonstrations or definitions. Items numbered (al), (a2) etc. describe the series of actions required to be demonstrated for the given code.

The criteria listed under "idiomatic" and "literal" must be met in order for these codes to be scored. The examples included for other codes are not exhaustive, but represent a guide to distinguishing these from other codes. In other words, an item should not be coded "idiomatic" or "literal" unless it fulfils all of the necessary criteria. An item may still qualify for one of the other codes, even if it does not match the examples given, as long as the general sense of the code is adhered to.

If the child anticipates the expression (i.e. acts it out before it is played), a subscript "A" should be inserted after the relevant code (eg. "idiomatic_A"). (Specific guidelines for the coding of anticipated actions are included for "on the spot", but the demonstration of any idiom should be marked "A" if it occurs prior to the playing of that idiom.) If the child acts out the expression only after it has been replayed, a subscript "R" should be inserted after the relevant code eg. "idiomatic_R".

Unless stated to the contrary, a definition phrased in a tense other than that used in these guidelines still qualifies for the given code. For example, for the idiom, "pick up", a code of "idiomatic" should be scored for "fetch", "fetched" or "will fetch".

391
1. RAN OVER - PLAY TASK

IDIOMATIC
(a) Rover is knocked to lying position with the bike's (front or back) wheel while the bike rider is on the bike and THE WHEELS ARE IN CONTACT WITH THE GROUND
(b) Child lowers Rover to lying position with his or her own hand while the bike wheel/s touch/es Rover and THE WHEELS ARE IN CONTACT WITH THE GROUND
Note: The child may lift the bike immediately before or after hitting Rover provided this is only to adjust the balancing lever.

INFERRED IDIOMATIC
(a) A character other than Rover is knocked to lying position with the bike and/or a character other than the bike rider rides the bike
(REASONING: idiomatic action - incorrect character)
(b) Rover is nudged with the wheel/s of the bike but does not fall down
(REASONING: ODCIE definition: "knock (sb) down and possibly pass over his body - therefore this action cannot be coded "idiomatic")
(c) Child lowers Rover to lying position when the bike wheels are not in contact with Rover and subsequently knocks Rover with the bike wheel/s while they are IN CONTACT WITH THE GROUND

LITERAL
(a) Bike rider (off the bike) is moved over Rover's body

INFERRED LITERAL
(a) Bike rider (on his bike) is moved over Rover without the wheel/s making any contact with him
(REASONING: Later in the story, many children demonstrate "The bike rider started running by moving the bike rider on his bike")
(b) A character other than the bike rider is moved over Rover's body
   (REASONING: literal action - incorrect character)
(c) Bike rider (on or off his bike) is moved over a character/prop other than Rover
   (REASONING: literal action - incorrect character)

AMBIGUOUS - IDIOMATIC/LITERAL
(a) Bike wheels nudge or knock Rover down and the bike is then placed/driven astride Rover or driven over him
   (IDIOMATIC: ODCIE definition: "knock (sb) down and possibly pass over his body"
   INFERRED LITERAL: (a) bike rider is moving over Rover therefore the child may be interpreting "over" as a modifier of the verb "run" - especially given that (b) later in the story, "the bike rider started running" is often demonstrated by moving the bike rider while on his bike
   COMMENT: It is an inherent problem of this idiom that the acting out of the second part of the dictionary definition of its idiomatic sense may also be considered as a literal demonstration.)
(b) Knocks Rover to lying position from above

2. RAN OVER - DEFINITION TASK

IDIOMATIC
(a) expresses the concept of being driven over
(b) expresses the concept of being gone over, knocked over, pushed over or hit by a vehicle
(c) "knocked down", "knocked over"
   (REASONING: These idioms describe the idiomatic sense and do not have literal meaning which relate to the literal sense of "ran over").
INFERRED IDIOMATIC
(a) describes an injury (eg. death or pain)
(b) mentions an accident
(c) repeats the idiom in conjunction with a vehicle
(d) "crash!", "bang!" or other onomatopoeia associated with road traffic accidents

LITERAL
(a) expresses the concept of walking or running across Rover or any non-specific item

AMBIGUOUS - IDIOMATIC/LITERAL
(a) "go over", "went over"
(REASONING: This could mean going over with a vehicle or walking over)
(b) "something/it goes over you/one"
(REASONING: The pronoun "it" or "something" implies that an inanimate object ran over rather than a person BUT pronoun confusion is associated with SPD and coding cannot therefore be justified on the grounds of pronoun choice.)

3. TOOK OFF - PLAY TASK

IDIOMATIC
(a) drives the bike away from Rover
(b) continues driving the bike away from Rover
(c) drives the bike away from Rover, lifts the bike to adjust the balancing lever and immediately drives the bike on or parks it

INFERRED IDIOMATIC
(a) drives the bike away, lifts the bike as it reaches the bend in the road in order to change the hand manipulating it or because s/he can't reach with fully stretched arm
(b) drives or pushes the bike a very short distance (\(<= 1\)"

394
(c) (the bike is in the air when the expression is played), Child immediately lowers it and drives it away

LITERAL
(a) removes the bike rider from the bike
(b) lifts the bike rider on the bike in the air without adjusting the balancing lever
(c) lifts one wheel of the bike

INFERRED LITERAL
(a) removes the bike rider or policeman's helmet
(b) lifts any character other than the bike rider off the set
(REASONING: Literal action - incorrect character)

4. TOOK OFF — DEFINITION TASK

IDIOMATIC
(a) expresses the concept of driving away, driving on, going away or continuing to "go on"

INFERRED IDIOMATIC
(a) "goes" or "went" without adding "away"
(b) "drive/ride off"
(REASONING: The child may be simply repeating "off" from the expression making this definition less robust than "drive/rode away/on" which is very obviously idiomatic. The child has, however, grasped the notion of it being a vehicle moving rather than a person being lifted away from the road/bike. The item therefore qualifies for INFERRED IDIOMATIC.)

LITERAL
(a) "go/get/move off the road/bike/table"
AMBIGUOUS - IDIOMATIC/LITERAL

(a) "go off"

OTHER FIGURATIVE SENSE

(a) expresses the concept of a plane taking off

5. GAVE A HAND - PLAY TASK

IDIOMATIC

(a1) brings Sally, Mum and Rover together
AND
(a2) walks/lifts Sally and Mum with Rover between them to the car

INFERRRED IDIOMATIC

(a1) brings Sally to Rover
AND
(a2) Sally carries Rover to the car (with/without Mum being taken to the car separately)
(REASONING: Sally is helping Mum by carrying Rover for her)

LITERAL

(a) Child tries to remove Sally's hand
(b) Child brings Sally's hand in contact with her mother without Rover being involved

FUZZY

(a) Mum carries Rover to the car alone
(b) Child walks/lifts Rover to the car
6. GAVE A HAND - DEFINITION TASK

IDIOMATIC
(a) expresses the concept of helping

LITERAL
(a) expresses the concept of removing a hand

INFERRRED LITERAL
(a) mentions or demonstrates a real hand

7. MOVED UP - PLAY TASK

IDIOMATIC
(a) the bandages are pushed across or along the table with the child's own hand or with the vet's hand, arm, trunk or feet
Note: The action is still coded as idiomatic if some of the bandages fall off the table and the child leaves them where they fall.
(REASONING: Pushing the bandages along the table without dropping any off is being considered an age-related dexterity skill)

LITERAL
(a) lifts some/all of the bandages up above the table and holds this position or replaces them as before
(b) (some/all of the bandages are already on the floor), Child lifts them onto the table
(c) (some/all of the bandages are already on the floor), Child lifts them up in the air, holds this position or replaces them as before
INFERRED LITERAL

(a) lifts a prop/character other than the bandages up in the air, then holds this position or replaces as before
(literal action - incorrect prop)
(REASONING: Lifting Rover from the floor onto the table is not sufficient for this code. It could merely be an anticipation of the next action. If the child holds Rover up in the air and then replaces him as before, then it does qualify for this code.)

AMBIGUOUS - IDIOMATIC/LITERAL

(a) child pushes the bandages across the table, some or all fall off, the child lifts them back onto the table or onto the pile
(b) child lifts some of the bandages onto the pile further along the table

FUZZY

(a) child pushes all of the bandages off the vet's table onto the floor
(REASONING: Can't be sure what the child is actually doing. It isn't literal but also isn't clearly idiomatic. Resembles "moved off" more than "moved up". See (2) comments.)
(b) child lifts some/all of the bandages just off the vet's table and puts them on the floor
(REASONING: This is really more of a sideways and downwards movement than a clearly upward movement and does not therefore qualify for a code of LITERAL. Secondly, unlike (1), in this case the child may remove all or some of the bandages. When the child pushes the bandages and some fall off, this may be due to immature fine motor skills. It is therefore coded as idiomatic. When some are lifted off, however, this cannot be idiomatic.
COMMENT: MI8 and others say "move off" before doing it. This may result from a misunderstanding / "mishearing" /
don't recognise the expression "move up" and so change it to one they do understand. Whatever the reason, the fact that they express "move off" supports the decision to code this item as fuzzy rather than as a demonstration of the literal sense of "move up")

8. MOVED UP - DEFINITION TASK

IDIOMATIC
(a) expresses the concept of the bandages being:
shifted } away / over / out of the way / to the side /
pushed } to one side
OR
moved away / out of the way / to the side / to one side

Note: If the child says "shift(ed)", this need not necessarily be followed by one of the above adverbs.
If the child uses the words "push" or "move", these must be followed by one of the adverbs/adverbial phrases listed next to it.
If "move" is followed by "over", it should not be coded IDIOMATIC.
(REASONING: The word "push" implies movement in contact with the surface i.e. not above the surface. Hence "push over" is acceptable as IDIOMATIC. "Move over", however, may be LITERAL or IDIOMATIC and is therefore coded AMBIGUOUS-ID/LIT.)

INFERRED IDIOMATIC
(a) "push"
(b) "push up"
(REASONING: "push" implies moving while retaining contact with the surface. This is therefore not LITERAL. The repetition of the particle "up", however, reduces this item from IDIOMATIC to INFERRED IDIOMATIC.)
LITERAL
(a) expresses the concept of picking up / lifting / moving in an upward direction

AMBIGUOUS - IDIOMATIC/LITERAL
(a) "put the bandages on the other ones"
(REASONING: This could mean lifting the bandages onto the pile (LITERAL) or moving them out of the way (IDIOMATIC)
(b) "move the bandages or things over"
(REASONING: This could mean "lift the bandages over" which would involve an upward movement (LITERAL) or this could mean shift them across (IDIOMATIC).)

REPEAT-NARRATIVE
(a) "moving them so there's more room/space"
(REASONING: similarly to "stopped on the spot", the child may be responding just to the verb i.e. just to "moved" and not to the whole expression "moved up". The rest is just a repeat/paraphrase of the rest of the sentence.)
(b) "moved them"

FUZZY
(a) "move off the table"
(REASONING: Although this is not LITERAL, the child is not treating the expression as a unit (therefore not IDIOMATIC). The child breaks up the expression and changes the preposition.)

9. GOT ON TO - PLAY TASK

IDIOMATIC
One or more of the following actions representing "the vet examined Rover's head" is/are then applied to Rover's leg. Different actions may be used for the head and legs:-
(a) vet is held or stands leaning forward with his face close
to Rover's head/leg
(b) vet's hand or arm is placed on Rover's head/leg
(c) stethoscope is placed on Rover's head/leg
(d) bandage is placed on or wound around Rover's head/leg

Note: The vet may be held with his feet on or above the table as long as he is clearly not standing on any part of Rover's body AND as long as one of the above actions is being executed at the time.

**INFERRED IDIOMATIC**
(a) Applies one of the actions listed under IDIOMATIC to the head and then legs of a character other than Rover (REASONING: idiomatic action - incorrect character)
(b) no action / incomplete action for head examination followed by one of the above actions applied to Rover's leg/s
(c) moves the vet from the head-side of the table to the leg-side of the table OR rotates Rover so that his legs face the vet

**LITERAL**
(a) the vet is stood on Rover's leg/s

**INFERRED LITERAL**
(a) stands a character other than the vet on Rover's leg/s (literal action - incorrect character)

10. **GOT ON TO - DEFINITION TASK**

**IDIOMATIC**
(a) expresses the concept of progressing from one activity or part of an activity to another OR specifically from Rover's head to Rover's leg(s)
INFERRED IDIOMATIC
(a) "looked at", "examined" or "did" (Rover's legs / them)
(REASONING: This is not robust enough for IDIOMATIC because
it does not express the progression clearly enough. It
does, however, have the sense of doing something to the
legs by way of examination and is clearly not a definition
of the literal sense. Although vague, "did" is acceptable
in view of the restricted vocabulary associated with young
children. A child who has "examined" as part of his/her
passive vocabulary may not yet have it as part of his/her
active vocabulary.)
(b) a sense of progression or sequence is weakly expressed (eg.
solely by "(did) something else") OR expresses progression
within an ungrammatical utterance

LITERAL
(a) expresses the concept of moving upwards, climbing up onto
or being on top of something
(b) expresses the concept of standing on Rover's leg(s)

INFERRED LITERAL
(a) expresses the concept of putting or lifting non-specific
objects or specific objects - other than the vet - onto
other objects
(b) "above" / "on top of" (without "climbed" / "went" or "moved"
i.e. the active sense is omitted)

AMBIGUOUS - IDIOMATIC/LITERAL
(a) "moved up to (Rover's legs)"
(REASONING: Depending on how the three words are broken
down, this expression could be interpreted in three
different ways:-
(i) (move) + (up) + (to) where each word is interpreted
in its literal sense as in "The book is on the floor.
Please move it up to the shelf."
(ii) (move up) + (to) where "move up" means "shift" as in
"Move up to the end of the bench please."
(iii) (move) + (up to) where "up to" means "as far as" as in "I left from London and made it up to Bristol before running out of petrol."
In the context of this task, (i) would be coded LITERAL, (ii) would be coded INFERRED IDIOMATIC and (iii) would be coded NARRATIVE. Since it is unclear which sense the child intends, the item is coded AMBIGUOUS.

(b) uses "moved on" alone
(REASONING: This expression could mean "progressed" (IDIOMATIC) or "moved up onto" (LITERAL).)

(c) "went/moved/got onto/on to (Rover's legs)"
(REASONING: This could translate as "progressed to" (IDIOMATIC) or as "got up onto" (LITERAL).)

NARRATIVE
(a) "to check Rover's legs" / "to see if Rover's legs were all right"
(REASONING: The use of the infinitive indicates that the child is proposing an explanation for the action of "got on to" rather than defining the action.)

11. ON THE SPOT - PLAY TASK

IDIOMATIC
(a) immediately stops the bike rider (on foot or on the bike) where he is and then does not move him
Note: If the bike rider topples after being stopped immediately and where he is, the child may stand him as long as he is stood in the same place as before.
If the bike rider is on the bike when stopped, the child may then lift the bike to pull down the balancing lever provided it is then parked on exactly the same place as before.
IDIOMATIC
(a) (holds the bike rider stationary on foot or on bike before the expression - i.e. hasn't released grip on bike rider), releases bike rider and takes hands away or manipulates other props
(b) (bike rider has already stopped after running without his path being blocked), Child does not move bike rider further

INFERRED IDIOMATIC
(a) immediately stops the bike rider where he is and then rotates him or shifts him fractionally
(b) (the bike rider is held off the ground before the idiom), Child immediately lowers the bike rider and stands him where he is

INFERRED IDIOMATIC
(a) bike rider has had his path blocked or has toppled while running and therefore has come to a stop before the expression is played
   Note: Child may be shifting the obstruction or standing the bike rider
(b) (bike rider is stationary before the expression is played without his path being blocked), Child rotates or shifts the bike rider a very small distance

LITERAL
(a) moves the bike rider (on foot or on the bike) to the green dot

INFERRED LITERAL
(a) moves the bike rider onto the green dot and then rotates or shifts him fractionally

OTHER FIGURATIVE SENSE
(a) moves the bike rider to the place where the accident happened
   Note: This is specified in the transcript where it happens.
AMBIGUOUS - IDIOMATIC/LITERAL
(a) the bike rider is on/just next to the green dot when the expression is played and is immediately stopped there (REASONING: Prepositions are not consistently understood in children with language disorders/SPD. It is conceivable that children may interpret "on" as "next to")

APPROPRIATE NO ACTION
(a) stops the bike rider (on foot or on the bike) AND takes own hands away BEFORE the expression AND does not touch the props AFTER the expression

12. ON THE SPOT - DEFINITION TASK

IDIOMATIC
(a) expresses the concept of stopping or doing something immediately and/or where one is at that moment

INFERRED IDIOMATIC
(a) expresses the concept of standing still

LITERAL
(a) expresses the concept of stopping / moving to / being on the (green) dot

OTHER FIGURATIVE SENSE
(a) "in like one place" / "stopped on a certain place" / "run on the spot" / "where it happened" or any other utterance which indicates that the child is interpreting the word "spot" as "a particular place, definite locality" (REASONING: The child has broken the expression down into its component parts therefore it is not a definition of the idiomatic sense of the whole expression. Having broken the expression down, the child has then defined the figurative sense of the single word "spot" - "particular place, definite locality" (Concise Oxford Dictionary) - as opposed to the literal sense of the word - "small part of
the surface of a thing distinguished by colour, texture etc usually round or less elongated than a streak or stripe" (Concise Oxford Dictionary).

REPEAT
(a) "stopped" without adding where
(REASONING: This does not demonstrate any interpretation of the expression "on the spot".)

13. TURNED OVER - PLAY TASK

IDIOMATIC
(al) moves Sally to the bike rider
AND
(a2) moves the bike rider and Sally closer to the policeman (either together or one after the other) WITHOUT ROTATING THEM
OR
(b1) moves Sally to the policeman
AND
(b2) moves Sally and the policeman to the bike rider (either together or one after the other) WITHOUT ROTATING THEM
OR
(c) (Sally is already placed close to the bike rider), Child moves Sally and the bike rider closer to the policeman (either together or one after the other) WITHOUT ROTATING THEM

INFERRRED IDIOMATIC
(a) acts out any of the options listed under IDIOMATIC using the incorrect character(s)
(REASONING: idiomatic action - incorrect character)
(b) (the bike rider is already standing close to the policeman), brings Sally to them
(c) moves the bike rider to Sally and then moves them both to
the policeman (together or one after the other) WITHOUT
ROTATING THEM
(REASONING: The child has the sense of taking someone to the
authorities but is not following the logic that the offender
wouldn't fetch the other party, or is confusing the
characters, or fails to understand the grammar)
(d) (Sally is standing close to the bike rider and the
policeman), Child grips Sally before/after/while moving the
bike rider closer to the policeman
(e) (Sally is standing very close to the bike rider and the
policeman), Child moves bike rider towards the policeman
WITHOUT ROTATING HIM and then HA or holds in this position

LITERAL
(a) rolls the bike rider or turns him upside down
(ODCIE definition: "(cause somebody/something to) face in
another direction by rolling)"
Note: Since the next word in the sentence is "to", the child may
roll or rotate the bike rider towards the policeman and
still be coded as literal or inferred literal rather than
ambiguous-both.

INFERRED LITERAL
(a) rotates the bike rider more than is necessary to make him
face the policeman
(b) rolls, turns upside down or rotates any character or prop
other than the bike rider
(c) lays down the bike rider or any other character on his/her
back or front

AMBIGUOUS - IDIOMATIC/LITERAL
(a) rotates the bike rider to face the policeman
(could be idiomatic/inferred literal 1)
FUZZY
(a) moves the bike rider to the policeman but does not touch Sally
(REASONING: Child may simply be ignoring the expression altogether and moving on to the next bit of the story i.e. the policeman taking the bike rider to the p/s. This overrides the possibility that Sally is telling the policeman that this is the offender in instances when she is standing close to the bike rider and the policeman.)

14. TURNED OVER - DEFINITION TASK

IDIOMATIC
(a) expresses the concept of giving / handing / taking / bringing someone / the bike rider to the policeman / authority figure
(REASONING: Is it an integral part of the idiomatic sense that the receiver should be an authority figure.)

INFERRED IDIOMATIC
(a) expresses the concept of giving / handing / taking to somebody without making it clear that it is to a policeman or authority figure
(REASONING: see IDIOMATIC above)

LITERAL
(a) expresses the concept of rolling somebody / something / the bike rider OR flipping him over

INFERRED LITERAL
(a) repeats "turned" in conjunction with a phrase indicating a change in direction
(b) expresses the concept of rotating rather than turning over
(c) introduces the analogy of an inanimate object which is turned over
OTHER FIGURATIVE SENSE
(a) expresses the concept of a vehicle accident

NARRATIVE
(a) expresses the concept of telling the policeman what has happened
(REASONING: To score IDIOMATIC or INFERRED IDIOMATIC, the child must show that s/he understands the concept of giving or taking the offender to the authorities. Just telling the policeman what happened does not reach the threshold.)
(b) expresses the concept of arrest

15. UNDER LOCK AND KEY - PLAY TASK

IDIOMATIC
(a1) Child puts the bike rider into the cell or leaves him there if he is already in
AND
(a2) Child closes the lock around the door and wall bars
Note: The child does not have to remove the key but if s/he does remove it, then s/he can place it anywhere outside the cell as long as it is clearly not on top of or beneath anything else.
Note: The transcript specifies when the lock is closed. If the transcription reads "lock looped around bar" and does not state that it is subsequently closed, then the item should not be coded IDIOMATIC.

INFERRED IDIOMATIC
(a) the child loops the lock around one or more of the cell bars but does not close it
(b) the child closes the lock around the incorrect bar/s of the cell (i.e. not around the door and wall bars at the opening edge) and leaves it there

409
LITERAL
(a) places the lock and key on top of the bike rider

INFERRED LITERAL
(a) places the lock and key beneath the bike rider
(b) places the lock and key on top of/beneath another character
   (REASONING: literal action - incorrect character)

INAPPROPRIATE NO ACTION
(a) the child takes the bike rider into the cell but withdraws
    his/her hands and does not touch the lock and key
    (REASONING: The sentence reads "The policeman put the bike
    rider in the cell and kept him under lock and key." The
    child is therefore demonstrating the first clause of the
    sentence but does not demonstrate the second clause
    containing the idiom.)

FUZZY
(a) the child does something with the lock and key but does not
    leave it looped/locked over the bar/s

16. UNDER LOCK AND KEY - DEFINITION TASK

IDIOMATIC
(a) expresses the concept of locking or being locked up / in

INFERRED IDIOMATIC
(a) expresses the concept of being unable to get out
Note: the child may express this as "can't" or "won't"
(REASONING: Children in this age group may not grasp the subtle
   distinction between "can't" (unable to) and "won't" (will not try
to).)

410
LITERAL
(a) expresses the concept of the lock and key being on top of someone / something

INFERRED LITERAL
(a) expresses the concept of someone / something being on top of the lock and key

17. PICK UP - PLAY TASK

IDIOMATIC
(al) brings the policeman to the bike OR moves them towards each other

AND

(a2) takes the policeman and the bike to the police station OR to just outside the police station with the bike wheels in contact with the ground.

Note: The child may lift the bike either to adjust the balancing lever or to remove the policeman from the bike. (These are acceptable because it is difficult to do this with the wheels in contact with the ground.)

INFERRED IDIOMATIC
(a) brings the bike to the policeman and then moves them to the police station with the wheels in contact with the ground

(b) drives the bike to the police station with its wheels in contact with the ground

(REASONING: Using a prop as the agent of a play action is an age-related skill.)

(c) brings the policeman to the bike but then drives the bike and walks the policeman separately to the police station

LITERAL
(a) the child holds or places the bike across/between the policeman's arms OR holds the bike vertically against the
policeman OR lifts the bike using the policeman's hand OR lifts the bike with the hand holding the policeman without then driving it to the p/s

INFERRED LITERAL
(a) lifts a prop other than the bike up in the air or onto something else
   (literal action - incorrect prop)

AMBIGUOUS - IDIOMATIC/LITERAL
(a) lifts the bike to the police station
   (This could represent literally picking up the bike or the idiomatic sense of taking it to the police station inferred literal or inferred idiomatic)
(b) lifts the bike and the policeman to the police station
(c) lifts and lowers the policeman and/or the bike before/after driving the bike into the police station
(d) lifts the bike into/onto the policeman's arms before/after driving it into the police station

18. PICK UP - DEFINITION TASK

IDIOMATIC
(a) expresses both the concept of going to the bike and that of bringing it back in one utterance eg. "fetch" / "collect" / "went to get"

INFERRED IDIOMATIC
(a) expresses the concept of going to the bike and bringing it back within an ungrammatical utterance
(b) expresses the concept of bringing the bike to the police station but not if going out to get it
LITERAL
(a) expresses the concept of lifting / carrying / holding something / the bike

INFERRED LITERAL
(a) expresses the concept of getting the bike using one's hands or arms
(b) introduces the analogy of something else being lifted up
(c) repeats "pick up" in conjunction with the concept of dropping or lowering

OTHER FIGURATIVE SENSE
(a) expresses the concept of standing the bike/something upright after it has been in lying position
(b) expresses the concept of arrest

19. KEPT AN EYE ON - PLAY TASK

IDIOMATIC
(a) moves Sally closer to and/or to face Rover
(b) moves Sally towards Rover, then away, then towards him etc.
(c) moves Sally around the edge of the basket always facing it and always in an upright position

Note: Sally does not balance in sitting position. Provided the child has attempted to seat Sally facing the basket, it is acceptable if s/he then props Sally up against the wall - even if Sally is consequently looking past the basket.

INFERRED IDIOMATIC
(a) stands a character other than Sally close to and facing the basket
   (REASONING: idiomatic action - incorrect character)

LITERAL
(a) places/holds Sally with her face up against Rover
20. KEPT AN EYE ON - DEFINITION TASK

IDIOMATIC
(a) expresses the concept of looking at / looking after / watching

INFERRED IDIOMATIC
(a) expresses the concept of looking at / looking after within an ungrammatical utterance
(b) expresses the concept of keeping Rover safe or indoors

LITERAL
(a) expresses the concept of putting / placing / holding an eye on / against Rover

21. DROPPED IN - PLAY TASK

IDIOMATIC
(a) policeman walks into the house
(b) policeman is lifted into the house through the side/front entrance
(c) policeman is gently lowered into the house through the roof and is stood/seated on the floor
(Provided the child is clearly not dropping the policeman from above, this is acceptable on the grounds that the children often lower other characters into the house through the roof as an alternative to using the other entrances.)

**INFERRED IDIOMATIC**
(a) policeman is brought to the side/front entrance of the house but is not brought inside

**LITERAL**
(a) policeman is dropped from above into the house

**INFERRED LITERAL**
(a) policeman is held and then dropped anywhere on the set other than into the house

22. DROPPED IN - DEFINITION TASK

**IDIOMATIC**
(a) expresses the concept of coming-going to the house/a place
Note: The child may use idioms like "pop over" and "go round" which do not have a literal sense corresponding to the literal sense of "dropped in".
(b) "come/go to visit/see (Rover and Sally)"
(REASONING: Although "visit" occurs in the quoted sentence, the child is given the benefit of the doubt in this one case because "visit" constitutes a perfect definition of the idiomatic sense of "dropped in".

**INFERRED IDIOMATIC**
(a) "come", "come in" or "went in"
(REASONING: These are too weak for IDIOMATIC but are clearly not LITERAL.)
Note: "went" on its own should NOT be coded INFERRED IDIOMATIC.
(b) Child describes a scenario of visiting eg. approaching a house, ringing the bell and walking in

**LITERAL**

(a) expresses the concept of dropping / falling

**NARRATIVE**

(a) "to see how Rover was", "to check if Rover was all right"
(Reasoning: These are explanations for the visit rather than definitions of "dropped in". These differ from "went to see" which may mean "visit".)

**FUZZY**

(a) child just says "went"
(Reasoning: This lacks the relation to "visit" that "come" or "went in" have. It therefore fails to attain the threshold for INFERRED IDIOMATIC)

**23. WENT ROUND - PLAY TASK**

**IDIOMATIC**

(a) Rover approaches all the people in the house
Note: Rover may or may not make snout contact.
   Rover may or may not also approach other characters on the set.

**INFERRED IDIOMATIC**

(a) brings everyone in the house to Rover
(b) Rover approaches at least 2 but not all of the people in the house
   (Reasoning: Story reads "went round everyone".)
(c) takes a character other than Rover to everyone in the house
   (idiomatic action - incorrect character)

**LITERAL**

(a) Rover runs/walks in a circle without approaching anyone
INFERRED LITERAL
(a) A character other than Rover runs/walks in a circle
   (literal action - incorrect character)

24. WENT ROUND - DEFINITION TASK

IDIOMATIC
(a) expresses the concept of going "from person to person" or
going to each one "in turn"

INFERRED IDIOMATIC
(a) expresses the concept of going from person to person within
   an ungrammatical utterance
(b) describes going to Mum and then to Sally and then to the
   policeman etc.
(c) "one at a time"
   (REASONING: This does not necessarily mean going to each
   one immediately after another.
(d) "go to each/every single person"
   (REASONING: The child adds the sense of going to each
   individual as opposed to the items listed for NARRATIVE.)

LITERAL
(a) expresses the concept of Rover/someone moving in circular
   motion or around the edge of something

INFERRED LITERAL
(a) "turned around"

AMBIGUOUS - IDIOMATIC/LITERAL
(a) "went round in a circle"
   (REASONING: This could mean turn 360° (LITERAL) or it could
   mean go from person to person while everyone is sitting in
   a circle (IDIOMATIC).
(b) "walked round"
(REASONING: This could mean walk around the edge of something (LITERAL) or it could mean walk from person to person (IDIOMATIC).

NARRATIVE

(a) "went to/licked everyone/everybody"

(REASONING: The child is not adding anything to the quoted sentence.)
APPENDIX J: SENTENCES REPEATED DURING THE PLAY TASK

Sentences were repeated during the play task when it appeared to the author that the child had become distracted (indicated with a bracketed "A" - denoting "author initiated" - in the table below) or in response to a child's request for clarification eg. "What did it say?" or "Lift what up?" (indicated with a bracketed "C" - denoting "child initiated" - in the table below). Table 77 indicates which idiom sentences were repeated for which children during the play task.

**Table 77: Children for whom sentences were repeated during the play task**

<table>
<thead>
<tr>
<th>IDIOM</th>
<th>MI</th>
<th>MJ</th>
<th>LD</th>
<th>SP</th>
</tr>
</thead>
<tbody>
<tr>
<td>RAN OVER</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOOK OFF</td>
<td></td>
<td></td>
<td></td>
<td>SP22 (A)</td>
</tr>
<tr>
<td>GAVE A HAND</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MOVED UP</td>
<td>MI8 (C)</td>
<td>MJ1 (A)</td>
<td>MJ7 (A)</td>
<td>SP11 (C) SP12 (C)</td>
</tr>
<tr>
<td>GOT ON TO</td>
<td>MI12 (A)</td>
<td></td>
<td>LD3 (C)</td>
<td>SP22 (A)</td>
</tr>
<tr>
<td>ON THE SPOT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TURNED OVER</td>
<td>MI1 (C)</td>
<td></td>
<td></td>
<td>SP19 (C) SP24 (C)</td>
</tr>
<tr>
<td>UNDER LOCK &amp; KEY</td>
<td></td>
<td></td>
<td></td>
<td>SP3 (C) SP20 (C) SP23 (C)</td>
</tr>
<tr>
<td>PICK UP</td>
<td></td>
<td></td>
<td></td>
<td>SP23 (A)</td>
</tr>
<tr>
<td>KEPT AN EYE ON</td>
<td></td>
<td></td>
<td></td>
<td>SP3 (C) SP22 (A) SP23 (A)</td>
</tr>
<tr>
<td>DROPPED IN</td>
<td></td>
<td>LD3 (C)</td>
<td>LD9 (A)</td>
<td>SP8 (C) SP22 (A) SP23 (A)</td>
</tr>
<tr>
<td>WENT ROUND</td>
<td></td>
<td></td>
<td></td>
<td>SP19 (C) SP20 (A) SP22 (A) SP23 (C)</td>
</tr>
</tbody>
</table>
APPENDIX K: IDEAS FOR THERAPEUTIC INTERVENTION FOCUSING ON IDIOM COMPREHENSION

Specific strategies for tackling idiom comprehension difficulty in children with semantic-pragmatic difficulties are included in Culloden et al. (1986), Hyde-Wright & Cray (1991) and in the ICAN pack, "Semantic Pragmatic Disorder: Suggestions for Teaching and Therapy Activities". Ideas for teaching idioms to children with language disorders in general (Abkarian, Jones & West, 1990; Nippold, 1991; Wiig & Secord, 1989), to children with learning difficulties (Ezell & Goldstein, 1992) and to normally-developing children (D'Angelo Bromley, 1984) may also be adapted. The second-language acquisition literature also offers some valuable ideas (Irujo, 1986).

The play task might also be used as a context in which to teach specific idioms as well as the process of idiom interpretation to children with semantic-pragmatic difficulties. The pragmatic aspects of interpretation could be explained and/or revised with reference to idioms which the child has accurately interpreted. (It would be useful to video the child's actions for this purpose.) The child might be encouraged to explain what else the expression could mean and why the (target) idiomatic sense is the most appropriate in the context. This process could then be applied to idioms which the child has interpreted inappropriately. The child could thus be taught not only the semantic meaning of new idioms, but also how to interpret those idioms in context. Clearly, teaching of idioms in the context of the play task would need to be reinforced and generalized through teaching and repetition in other contexts.

The story and play set could also be adapted to include idioms appropriate to particular children's experience and needs.1

---

1 Beyond idioms, the procedure might be further useful for assessment and intervention with regard to phonological and syntactic difficulties, for example. For these purposes, stories might be constructed to include minimal pairs or a variety of syntactic structures, respectively.
Indeed, whatever method of teaching idioms is used, attention needs to be paid to the nature of the idioms selected for teaching. In general, the teaching of idioms commonly used in the classroom, in the playground and in the home are most likely to promote social interaction and learning. The teaching of these idioms would also benefit from many opportunities for their rehearsal. It may be preferable to begin by teaching concrete rather than abstract idioms, the latter tending to be more difficult for young children to understand (Levorato, 1993). In the case of children who have very limited interests, it may be useful to begin with idioms related to those interests. For example, a child who has an obsessional interest in time and time-tables might be taught such expressions as "in time", "in a minute", "time and again", "clock on/off" and "after hours". Working from the child's obsession may help not only to motivate the child but also to encourage flexibility of thought.