ASSESSMENT OF COGNITIVE DEVELOPMENT IN FOUR TO EIGHT YEAR OLD CHILDREN BY MEANS OF DRAWING TASKS

BY

CHRISOULA KARAKITSOU

DEPARTMENT OF PSYCHOLOGY

UNIVERSITY OF STIRLING

THESIS SUBMITTED AS REQUIREMENT FOR THE DEGREE OF PHD

DECEMBER, 1999

To my Mum and Dad,
for giving me space to draw and trying to understand my pictures

ABSTRACT

The present thesis explores the link between children's drawings and cognitive development. The aim of this study is to investigate the intellectual abilities of the child draughtsman with good depiction skills and to evaluate the merit of the drawing technique in the assessment of conceptual maturity. The standardised Goodenough-Harris Drawing Test (GHDT) of intellectual maturity was administered to 115 children between 4 to 8 years of age against criterion ability measures (Wechsler scales). Its psychometric properties are examined in respect to its norms and scales, its reliability and validity at different age levels and ranges of intelligence. Early theories in the area of pictorial representation were directed towards identifying features characteristic of different developmental periods (Kerschensteiner, 1905; Luquet, 1927/1977). At the same time Piaget and Inhelder (1948/1967) incorporated these stage theories into their model of spatial intelligence. Yet, the recent experimental study of children's drawings has disclosed a number of variables which interfere during the course of production, challenging the view that drawings can be seen as the royal route to access children's concepts. Stage theories are re-evaluated by means of fourteen experimental drawing tasks with various degree of difficulty. The tasks – administered to the same children tested with the standardised instruments –are spatial in nature and have been sampled from two widely researched areas related to the pictorial representation of partial occlusion and of spatial axes (horizontal/vertical). The acquisition of the pertinent spatial concepts by means of drawings is examined, considering competence-deficiency and competence-utilisation accounts of children's performance at different ages. Finally, overall performance on spatial tasks is compared with performance on conventional (Wechsler scales) and non-verbal (GHDT) measures of intellectual functioning, considering the optimum method to assess children's abilities by means of drawings. In general, drawing performance is reasonably sensitive to children's level of intelligence, yet the significance of drawing varies at different ages and ranges of IQ. Finally, the establishment of steadfast developmental trajectories falls short in the field of pictorial representation. The variable performance, particularly from the children at intermediate ages, suggests that the stages of intellectual or visual realism should be seen as relative and not as absolute.

ACKNOWLEDGEMENTS

Drawing is a creative but also a problem-solving activity. Studying and writing about children's drawings was equally so. At this point I would like to thank the people who provided me with the much-needed mental and emotional tools through my studies. First and foremost I owe my thanks to my supervisor, Dr Robin N Campbell, for his initial suggestion to undertake this study and his constant encouragement to complete it. Working with him has been an inspiring and self-growing experience.

I am deeply thankful to the headteachers, Mr Bull and Mrs Marshall, who kindly allowed me access to their schools and created the ideal conditions for the collection of the data; to all the primary and nursery teachers for their much appreciated co-operation despite my never-ending testing sessions; to the parents and guardians who granted their children permission to participate in the study, but over and above special thanks go to the children who have been generous in their sharing of their creative work with me, made this thesis a reality and the time I spent with them the most enjoyable part of my studies.

Coming up to less zestful stages, I am grateful to Dr Peter Griffiths for all his help and guidance with the scoring of the psychometric tests and especially to Dr Ranald MacDonald and Dr Katie Howie for enlightening me with their expertise and settling my statistical afflictions. I am utterly thankful to Bruce Sutherland and Peter Hucker who on two occasions resurrected my hard disc from its dogged silence, restored my faith to miracles (not to computers) and gave me back a timid smile.

Of the many people who have crossed my path, I feel some deserve a special mention. At times that I was short of words to understand and communicate my bewilderment, being with my office mates was comforting and reassuring. A big thank you to Dave for the music (it wasn't too dark for me here, after all), to Verena for staying a friend (parcels were lovely) and to Trish for being so special (I'm missing you already). Studying here, for what seemed an infinite period of time, I've been lucky to become close with a few people. Septembers were bringing but also were taking away from me my friends. Ντινα and Παυλινα spiced up my life, Μαρια took care of me at difficult times and Πεπη made my last house a home.

A special thanks to Φαιδωνας for his words of wisdom, and Ελλη – to tell and make stories with you kept me going. Σπυρος, by sharing a big part of his life with me, made me richer. A big hug to Βασιλης who make the impossible possible for me and was full of love and support. I would have liked to thank Σενη but she won't let me – so how about you have been my window with a sanguine view. Finally the two people who have contributed the most to what I feel is good about this work and what feels good in me are Γιωργος and my sister Βαλια. Thank you for all the love and all the waiting. Both.

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

APS	Axial point score.
ASS	Axial structural score.
CEFT	Children's Embedded Figure Test.
CPS	Composite point score.
DAM	Draw-A-Man.
EFT	Embedded Figure Test.
FD	Field Dependent/Field Dependence.
FI	Field Independent/Field Independece.
FO	Fine occlusion.
FSIQ	Full scale IQ.
GH	Goodenough-Harris.
GHDT	Goodenough Harris Drawing Test.
GHSS(s)	Goodenough Harris standard score(s).
HF	Human Figure.
HFD(s)	Human Figure Drawing(s).
HLE	Hidden Line Elimination.
MO	Medium occlusion.
MS	Man scale.
MST	Mountainside task.
OPS	Occlusion point score.
OSS	Occlusion structural score.
PEFT	Preschool Embedded Figure Test.
PIQ	Performance IQ.
PL(Ts)	Plumb line (tasks).
RFT	Rod and Frame Test.
SS	Self scale.
VIQ	Verbal IQ.
WISC III^{UK}	Wechsler Intelligence Scale for Children.
WL(Ts)	Water level (tasks).
WPPSI	Wechsler Preschool and Primary Scale of Intelligence.
WS	Woman scale.

CONTENTS

CHAPTER IV

CHAPTER V

CHAPTER VI

CHAPTER VII

CHAPTER VIII

CHAPTER IX

CHAPTER X

CHAPTER XI

CHAPTER XII

CHAPTER XIII

Once, when I was six years old I saw a magnificent picture in a book, called True Stories from Nature, about a primeval forest. It was a picture of a boa constrictor in the act of swallowing an animal ...I pondered deeply, then, over the adventures of the jungle. And after some work with a coloured pencil I succeeded in making my first drawing. My Drawing Number One. It looked like this:

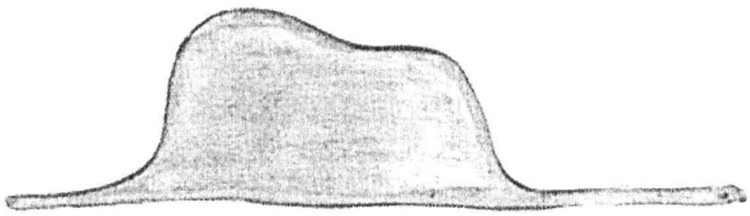


Figure 1.1 Drawing from The Little Prince by Antoine de Saint Exupéry. Copyright 1943 and renewed 1971 by Harcourt Brace & Company.

I showed my masterpiece to the grown-ups and asked them whether the drawing frightened them. But they answered: "Frighten? Why should any one be frightened by a hat?"

My drawing was not a picture of a hat. It was a picture of a boa constrictor digesting an elephant. But since the grown-ups were not able to understand it, I made another drawing: I drew the inside of the boa constrictor so that the grown-ups could see it clearly. They always need to have things explained. My drawing Number Two looked like this:

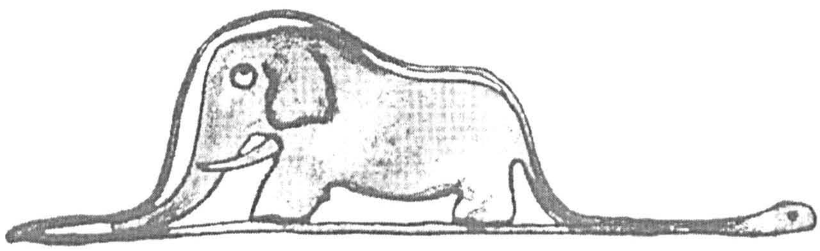


Figure 1.2 Drawing from The Little Prince by Antoine de Saint Exupéry. Copyright 1943 and renewed 1971 by Harcourt Brace & Company.

The grown-ups' response, this time, was to advise me to lay aside my drawings of boa constrictors, whether from the inside or the outside, and devote myself instead to geography, history, arithmetic and grammar. That is why, at the age of six, I gave up what might have been a magnificent career as a painter.

From *The Little Prince* by Antoine de Saint Exupéry (1943, p. 3-4).

CHAPTER I

INTRODUCTION

The representation of events, objects and relations by means of a system of signifiers (semiotic function) is widely acclaimed as instrumental in development and, according to Piagetian theory (Piaget, 1962; Piaget & Inhelder, 1969), a principal milestone of the pre-operational period. The major signifying systems to emerge during this time are language, mental imagery, symbolic play and drawing. The functional role of drawing in the overall development of children has been raised within many theoretical frameworks, and naturally is far from being resolved. Amidst the diversity of suggestions, most theorists seem to agree that the drawing medium is a symbolic medium and, as the manner of symbol-use changes in the course of development so does the role of drawing activity throughout its 'golden period' (2 years to 10 years of age).

Its appearance is marked by the so called scribbling; a relatively solitary activity, whose significance has been explained by some solely on the basis of its motoric aspects (Bender, 1938) and by others on the visual pleasure resulting from trace-production[1] (Gardner, 1980; Gibson, 1969; Kellogg, 1970), whereas a few claim that scribbles might have an elusive representational content (Freeman, 1980, p.4-6; Matthews, 1984). Although there might be no consensus about the role of scribbles, once a child accidentally discovers that the traces, which his strokes leave on paper, resemble real objects, drawing begins to acquire its symbolic function.

[1] Children, given a pencil-like tool that failed to leave a mark, ceased scribbling.

A number of analogies have been drawn between children's developing mastery over the graphic medium and language acquisition in its written and verbal form (Goodnow, 1977; Karmiloff-Smith, 1990; Pemberton, 1990; Willats, 1977 & 1992). For instance, the repetitive line configurations of scribbling have been related to the monotonous sounds of babbling, and the plurifunctional and general-purpose properties of early graphic forms have been compared to one-word sentences (see Eng, 1931, p. 181-182). Besides, some authors, either implicitly or explicitly, argue that drawing can even facilitate the acquisition of early writing skills (De Goes & Martlew, 1983; Ferreiro 1985; Goodnow 1977; Kellogg 1970; Martlew, 1988), and that there is a functional interrelatedness between the linguistic and the pictorial semiotic systems (see Stetsenko, 1995).

The assumption that learning to use a system of pictorial symbols promotes subsequent acquisition of other representational systems can be elucidated by looking at the mental operations that drawing entails. In the course of experimentation with lines and shapes, the child gradually discovers the dual nature of these graphic elements; that they can exist on their own material substance and they can stand for things as well. This discovery leads to further development as the child explores the potential of the medium, by selecting and combining elements in a manner that will better serve his/her communicative intentions. This process not only bears strong similarities with the process of mastering language (its semantic and pragmatic aspects), but also shows how the graphic medium equips the child with an alternative mode of expressing and interpreting the world.

Yet, graphic forms are not arbitrary symbols that stand for objects or ideas in the manner that words do. Not only do they have a referent in the real world but also its visual attributes codetermine how credible its graphic equivalents are. The fact that "pictures are made by people with people in mind" – as Costall nicely puts it (1985, p. 28) –, bestows a veritable representational status on the process of picture making, and ultimately spurs drawing development towards pictorial realism. To clarify matters, if we retrace the opening quotation from the story of *The Little Prince*, we understand that children not only use drawing as a self-expressive outlet – wherein ideas and feelings become tangible to them – but intend these to be readable by others. That's why the drawing of a boa loses its appeal to the six-year-old child once he discovers to his great disappointment that it fails to convey the element of fear to the grown-ups (see Figure 1.1). The realisation that one's standards of what constitutes an adequate representation of a subject matter are not shared by the beholder (peer or adult) instigates greater concern towards visual resemblance between the signifier and the signified. And this ultimately leads to the expansion of children's graphic vocabulary and to the acquisition of drawing conventions by trial and error.

Figural differentiation as well as discovery and use of conventions (intrinsic to the drawing medium as well as to the social milieu the drawing emanates from) is not an easy enterprise. As the present thesis aims to elucidate, the production of a realistic picture requires a lot of physical and mental effort. This explains why children – often older than the six-year-old child in the story of de Saint Exupéry – become dissatisfied with their pictures and cease drawing as they cannot overcome the taxing problems imposed by the medium. Notwithstanding, the acquisition of graphic competence enables the child to perform a range of operations within this particular system of symbols, which possibly promotes the mastery of other more abstract social-semiotic modes of communication (language). Having briefly outlined the communicative, expressive and cognitive aspects of drawing, one might be more prepared to appreciate why pictorial representation has been considered a product of thinking, or as Freeman expressed it 'a manifestation of someone's mind on paper' (1995). Along this line, it seems now reasonable that as the child expresses, interprets and reworks ideas through this medium, he/she comes to understand better him/herself and the world. At this point it is relevant to present in short the main approaches to children's drawings, before delineating the theoretical and empirical framework of the present work.

1. APPROACHES TO CHILDREN'S DRAWINGS

The study of children's drawings as a topic of psychological enquiry has a long and well-documented history, which extends just over 100 years. Its initial appearance was made through theories of intellectual development. The notion that "to make is to know" dictates the initial theoretical framework in which drawings were seen as direct products of children's inner world of ideas (concepts). Research interests were directed towards identifying uniform features, characteristic of different developmental periods. This kind of universal orientation to developmental phenomena looked for general principles, which cut across different contexts. Ultimately, the conception of a relatively invariant succession of stages or phases in drawing development was the central theme of those times (Burt, 1921; Cooke, 1886; Goodenough, 1926; Luquet, 1927/1977). The realisation that drawing reveals developmental aspects of mental maturity culminated in the use of the graphic medium for the assessment of conceptual development (Goodenough, 1926; Harris, 1963). Now, not only a number of conventional intelligence tests and developmental scales contain a subtest where the child has to draw something (i.e. WPPSI: Geometric Design; Stanford-Binet/BAS: Copying; MSCA: Draw-A-Design, Draw-A-Child)[2] but there are a number of standardised non-verbal tests which assess general/specific abilities or intellectual disorders, solely based on the drawing technique (Goodenough-Harris Drawing Test: Harris, 1963; Developmental Test Visual-Motor Integration: Beery, 1967; Memory for Designs Test: Graham & Kendall, 1960; Bender-Gestalt Test: Bender, 1938; Benton Revised Visual Retention Test: Benton, 1974).

[2] WPPSI = Wechsler Preschool and Primary Scale of Intelligence, Wechsler, 1967; Stanford-Binet, Terman & Merrill, 1960; MSCA = McCarthy Scales of Children's Abilities, McCarthy 1970; BAS = British Ability Scales, Elliot, 1983.

Embracing the use of drawings as a diagnostic tool, but arguing from a different conceptual framework, another school of thought viewed drawings as a direct projection of children's inner world of feelings and focused on the expressive aspects of the graphic forms. The pictorial expression of affect linked drawings with the assessment of personality and psychological adjustment, broadening the analysis of drawings beyond a strictly cognitive assessment of mental ability. After being on the sidelines for some time, children's drawings regained the interest of psychologists by mid-70s. Hitherto, the treatment of the subject from a variety of theoretical perspectives caused a refinement in the understanding of child art and revealed the conceptual and methodological inadequacies of the earlier approaches.

Drawing tests of intellectual development or personality were based on the aesthetic norm of naturalism or realism. The visually correct perspective projection was taken as the standard, deviations from which were interpreted as signs of conceptual or emotional immaturity. However, a number of authorities, by anchoring drawings in the field of visual perception, rejected visual realism as isomorphic to perception and maintained that the origin of graphic forms is not conceptual but perceptual, hence governed by the laws of perceptual organisation (see Arnheim, 1974; Gibson, 1979; Hochberg, 1978). Although the unappreciated significance of natural perception and picture perception in the study of drawings demonstrated how limited the earlier approaches were, still many felt that drawing development continued to be explained outside the medium within which it takes place.

Ultimately, recent theoretical advances have laid stress on the nature of depiction, viewing drawings not as print-outs of affective states or mental and visual concepts, but as constructions determined by the way children solve organisational and procedural problems (see Goodnow, 1977; Freeman, 1980; Van Sommers, 1984). Within this framework, performance factors were placed in the heart of this problem-solving activity and drawing research on a more explicit experimental footing. To understand the implications of these advancements, one has to recall Willats' comment on this 15 years ago:

> *To this extent the study of visual representation is beginning to come to age, just as the study of verbal language began to come of age with Chomsky's distinction between competence and performance (1985, p.84).*

The major contribution of this theorising was that it demonstrated that pictorial representation is far more complex that early theories made us believe, as a number of interactive processes cluster (sensorimotor coordination, cognitive/perceptual factors, memory capacity). The realisation of its multifaceted nature ultimately restored children's art to a central position in child cognitive psychology, and led many theorists to adopt an information-processing approach in an attempt to account for the plethora of factors involved in picture production (Morra, 1995; Van Sommers 1984, p.228 & 1989; Willats, 1995, see also Burton, 1995)

The first impression one gets after this brief introduction to the topic is that the theories of depiction are as many as the ways we can see and conceptualise pictures. Inevitably, the scope of any research project is liable to offer just one possible view of the topic. The present thesis provides a cognitive account of pictorial representation and explores the link between children's drawings and cognitive development. Two interrelated themes form the basis from which various other issues are raised. First, what are the mental characteristics of the child draughtsman with good depiction skills and second what conclusions can one draw about children's conceptual abilities by studying their drawings. In answering these questions, earlier theories as well as current approaches are juxtaposed, whereas a reconciliation is attempted at the end of the thesis.

2. THE PLAN OF THE THESIS

To become acquainted with the range of issues and topics, which will be studied here, a brief account of the structure of this thesis will follow, pinpointing the rationale guiding the present set of research interests. The sequence of the chapters corresponds to a scale which parallels the course of change in theoretical perspectives and methodological techniques adopted through the years. Since early studies on drawings began with an emphasis on their use in determining intellectual level, this naturally marks the starting point of the present research. Besides its obvious relevance to the main theme, investigating the utility of the drawing technique in intellectual assessment is of paramount importance for the following reasons. First, the topic is rather controversial. On the one hand some question its rationale and argue against its use, while others underscore its practical usefulness in educational services (see Witt, 1993). Second, since the appearance of the first drawing test in 1926, a number of researchers have followed this tradition up to the present times (Nagliery, 1988). Simply, this exemplifies the significance of systematically assessing its merit.

Considering that the most extensively used drawing instrument in intellectual assessment is the Goodenough-Harris Drawing Test (GHDT) and that, since its construction in 1963, it has never been revised, it seemed reasonable to use this as the empirical foundation for tackling the subject. Chapter II traces the origin of the test back to the first descriptive studies of the late 19th century, provides a detailed account of its features and offers a comprehensive review of the pertinent literature, integrating current research findings. Having identified problematic or least studied areas, specific as well as general hypotheses are tested in Chapter III and results are discussed in detail. The issues addressed are related to its psychometric properties in respect to its norms and scales, the reliability of the scoring method, its internal consistency as well as its validity as a measure of conceptual maturity.

Most of these issues are examined considering various subject variables such as age level, gender and ranges of intelligence, obtained from conventional measures.

A cursory glance into the drawing literature shows that research interests can be roughly divided into two topic areas; one which deals with Human Figure Drawings (HFDs) and the other with spatial problems of a specific or general nature. Early studies, by being mainly cross-sectional, revealed developmental changes in both areas and provided the empirical ground for the categorisation of drawings into series of distinctive stages. It was against the background of developmental stages that the GHDT was developed using the HFDs as the stimulus theme. And at the same time, Piaget and Inhelder (1948/1967) proposed a model of spatial intelligence incorporating stage theories of drawing development. This is the next area covered through Chapters IV to XI, which deal with the development of representational space. And although it is the author's firm belief that not all drawing is an exercise in geometry, a fair amount of diversity in the number of drawing themes studied is needed to obtain a balanced and comprehensive view of children's drawing competence.

As the method of studying children's drawings moved from the naturalistic approach to the experimental one, certain peculiarities symptomatic of earlier stages of development were put under scrutiny. Researchers, by setting carefully designed drawing tasks, revealed that conceptual knowledge is not always readily mapped onto paper. Since then, full attention has been given in setting tasks, which in fairness to children and to the nature of the medium reveal their true abilities. This is the principal objective running through Chapter IV to XI. The chapters have been equally divided into two domains with progressive emphasis on geometry. Both areas have engendered as much research as HFDs.

Chapters IV to VII address the question of how children solve spatial problems of relative position and Chapters VIII to XI deal with problems of relative orientation. The former problems are often translated into scenes where one object is partially or totally occluded by another, close to it. Pertinent to this issue is the difficulty that the six year old faces in the story of *The Little Prince*; making the elephant visible through the body of the boa, although this is fully masked by it (see Figure 1.2). Besides the problems of depth or object solidity, another domain causing representational difficulties is that of spatial orientation. Everyone is familiar with the charming albeit awkward drawings of houses with chimneys falling off roofs. Similar phenomena, first studied by Piaget with his water-level and plumb-line tasks, fall under the scope of the remaining chapters, dealing with the representation of the invariant horizontal and vertical axes.

For many years these "mistakes" have been interpreted in the line of conceptual immaturity, of limitations of spatial reasoning in the child. Piagetian theory of spatial intelligence predicts that spatial relations become graphically explicit once the child is able to conceptualise space in a way that the Euclidean and projective properties of objects are preserved. And this skill is acquired broadly speaking at the age of 9, a time where children's drawings become visually realistic. In addition, earlier research have shown that the aforementioned spatial problems elicit errors, characteristic of different developmental stages (Barnes, 1897, p.283-294; Cox, 1978; Freeman *et al.*, 1977; Piaget & Inhelder, 1948/1967; see also Leeds *et al.*, 1983). These errors have been often described as biases, a term which describes their obstinate nature (Freeman 1980; Pemberton, 1990). However, recent empirical data have shown that children are not firmly anchored on a certain stage. Not only there is a considerable asymmetry between picture perception and production (Fayol *et al.*, 1995), but within-subject variability across different modes of drawing task (mnemonic, copying, picture completion tasks). Most important, neither are their errors always systematic nor the biases always conceptually-driven (Arrowsmith *et al.*, 1994; Bremner, 1985; Cox, 1981; Davis, 1983; Freeman, 1983; Perner *et al.*, 1984). A general impression formed reading current research findings is that children younger than 8 years of age have been underestimated and that they are a great deal more "view-specific" than psychologists have given them credit for.

Congenial is the view of a number of researchers who feel that the set of criteria employed in previous studies to judge children's drawings has been often too narrow, if not unfair. Current advances in the understanding of picture production have shown that taking the standard of naïve photographic realism as the absolute criterion for successful representation is fallible. One line of reasoning is that art neither imitates nature nor does any adult draw what he sees in the sense that he creates a photographic record of a scene (Costall, 1995; Golomb, 1992 p. 124). Another point often raised is that for any representational problem there are a number of strategies, which bear varying degrees of credibility towards the visually accurate solution. The common practice of classifying drawings as correct or wrong certainly conceals children's growing abilities towards view-specificity. As most authorities seem to agree, development consists of the progressive discovery of more effective strategies. The study of these strategies can shed light on the nature of the problems that children face. Some of them show that the child totally lacks the requisite spatial concept whereas other reveal difficulty in implementing it in the course of production. Yet, as the editors of *Visual Order* affirm

> *The conclusion is that in the past (not to speak of the present) people have unwittingly operated a sleight-of-hand: by prejudicing the criteria for successful depiction they have, sensu strictu, prejudiced their chances of explaining development (Freeman & Cox, 1985, p. 5)*

Another criticism of early methodological procedures is their preoccupation with the surface structure of the drawings and their neglect of the process of production (Freeman, 1972, 1980; Goodnow, 1977). Drawing activity is a sequential process, which involves a number of successive steps and subsequent decisions at various points along the process. The well-known saying "Well-begun is half done" rightfully applies to children's drawings, as early steps have been shown to affect later ones either by increasing or limiting children's options. Further, some sequential patterns are guided by formal constraints whereas other by representational forces (Van Sommers, 1984, 95-114, 1991; Cox, 1992, p.113-116; Ingram, & Butterworth, 1989).

Due to the advances made in the experimental study of children's drawings, researchers are now discussing performance factors and describing how they interfere with the representation of view-specific information. However, although they are sceptical about the explanatory power of traditional stage theories, a number of issues are still unsettled. Performance factors have been used as an umbrella term to describe constraints of diverse origin. Similarly, these operate differently at different age levels. The drawing tasks presented in this thesis were designed with these considerations in mind. The key question addressed by the present research is what type of errors are age-dependent and which are task-specific. In other words, at what age are "competence-deficit" and "competence-utilisation" interpretations more credible. And this brings us back to the central theme of the thesis; finding the optimum way to assess children's representational abilities reliably and to make accurate assumptions about their conceptual maturity on the basis of their drawings.

Literature reviews, pertaining to the two spatial areas of drawing competence are to be found in Chapter IV and Chapter VIII respectively. These chapters expand on issues briefly introduced here, incorporating earlier and currently important theoretical propositions and research findings. On this basis, empirical questions are formulated and tested by means of drawing tasks where various factors are manipulated. The studies on occlusion are presented in Chapters V & VI, whereas those on spatial axes are presented in Chapters IX & X. Each drawing task is described separately followed by results and discussion sections. Although children's performance between tasks is compared as the studies progress, an overall review and general conclusions are presented in the final chapters dealing with these spatial problems (in Chapter VII for occlusion and in Chapter XI for spatial axes). These chapters answer the following questions: (a) What types of error appear to be age-dependent and which are task-specific? (b) when is the requisite spatial concept acquired and how?

Finally, an attempt is made in Chapter XII to provide an inclusive picture of drawing development bringing together the three rather independent areas of research studied (representation of human figure-GHDT: Chapters II & III, representation of spatial position-Occlusion: Chapters IV-VII, and representation of spatial orientation-Coordinate axes: Chapters VIII-XI). One objective is to assess their relative efficacy in revealing developmental changes and conceptual abilities. The second objective is to explore which method based on the drawing technique better reveals information of children's intellectual level of functioning. As has been previously mentioned, developmental changes were observed both in HFDs and spatial tasks. However, the basic difference between these two topic-specific lines of research was that the changes in HFDs have been related to both chronological and mental age, whereas the stages in the differentiation of pictorial space have been exclusively related to children's age. This issue is examined empirically in Chapter XII, by comparing overall performance on the drawing tasks and in each spatial domain separately with GHDT and conventional measures of intelligence. Finally, Chapter XIII concludes this thesis with summaries of the main findings and suggestions about future directions in the study of children's drawings. Before proceeding to the main body of this thesis, a general description of the methodology employed follows.

3. GENERAL METHODS

3.1 Participants

Children were recruited from two local nurseries and primary schools. No specific socioeconomic status measures were obtained from the subjects; however, an effort was made to sample schools that would represent a fairly broad spectrum of socioeconomic classes residing in the area. For those school classes that were targeted as samples, a letter was sent to the parents/guardians informing them about the general nature of the project and asking their permission for children's participation. A total of 115 right-handed children provided the data for the present research. The overall sample consisted of 42 boys and 73 girls, ranging from 3:7 to 8:8 years of age, and was grouped into four age levels as follows: 34 children with mean age of 4 years and 2 months (range 3:7-4:8 years), 26 children with mean age of 5 years and 6 months (range 5:3-6:2 years), 28 children with mean age of 7 years (range 6:8-7:5 years), and 27 children with mean age of 8 years and 3 months (range 7:8-8:8 years). Due to sporadic occurrences, where children either did not produce any data or their drawing couldn't be scored, there are occasional fluctuations in the sample sizes across studies. Subsequently, specific details are given in the relevant sections.

3.2 Instruments

Three standardised test instruments were administered to all children: (a) the complete Goodenough-Harris Drawing Test (b) a short form of the Wechsler IQ scales and (c) Embedded Figure Tests as a measure for field dependence/independence. As there was no single conventional measure of intellectual ability (b) and cognitive style (c) which would be age-appropriate for all the children, two equivalent forms of each test battery were used. Thus, the Wechsler Preschool and Primary Scale of Intelligence (WPPSI: Wechsler, 1967) and the Preschool Embedded Figure Test (PEFT: Coates, 1972) were administered to nursery children (youngest age group), whereas primary school children (remaining age groups) were tested with the Wechsler Intelligence Scale for Children (WISC-III^{UK}: Wechsler, 1992) and the Children's Embedded Figure Test (CEFT: Witkin *et al.*, 1971). A full description of the particular features of these instruments is provided in Chapter III, Methods section. All psychometric instruments were administered and scored by the author, after receiving formal training from a Charted Clinical Child Psychologist.

In all, 14 non-standard drawing tasks were administered. Apart for some drawings, which were made from memory, for copying and picture completion tasks, children were presented with A6 stimulus cards, each portraying a line-drawing picture of a simplified array. Some cards were used exclusively to assess one of the two spatial domains of representational ability. Others were multipurpose as they contained features for the assessment of the pictorial representation of partial occlusion and of the spatial axes. Children were provided with a standard test booklet and black pencils with an eraser. As each scene had to be drawn on a separate page, the paper of the children's booklet was sufficiently thick so as to prevent previous drawings and subsequent pre-drawn forms to be visible during drawing. Schematic illustration and detailed description of the task materials are given as each drawing task is presented separately in Chapters V-VI and Chapters IX-X.

3.3 Procedure

Children were tested individually by the author in a quiet room close to their classroom. Prior to the initial testing session, children's classes were visited and time was spent especially with the nursery children to familiarise them with the author. The testing with the psychometric instruments took place in the morning. Except for a few cases with WPPSI, each measure was administered in a single session, following the standard procedures described in the pertinent manual. The order of administration was fixed with Embedded

Figure Test completed first, followed by GHDT and Wechsler scales. Each child was tested with all the instruments in less than one-month time.

On completion of data collection from standard measures of ability, the drawing tasks were administered next. Each drawing was done without time constraint, and all the tasks were completed within three or four sessions, depending on the children's age and level of concentration. Most testing sessions lasted between 15 to 20 minutes, whereas a few were terminated earlier in the presence of signs of tiredness. Within each spatial area of drawing competence, the order of task presentation was fixed and comparable to the order followed in Chapters V-VI and IX-X. Yet, as it has been already mentioned, some stimulus cards were multipurpose, containing features relating to occlusion and spatial axes[3]. Considering that each child had to draw all the features of a scene before proceeding to the next one, in due course, tasks alternated across the two spatial domains. This was done to reduce practice effects between successive items assessing similar spatial skills and to sustain children's motivational level.

All children were taken from their class and escorted to the testing room. Children were seated on age-appropriate chairs, while the experimenter was always seated to the left of the child. After some preliminary conversation and once it was ascertained that the child was at ease, he/she was presented with one task at a time. For those tasks where children were drawing from a model, the stimulus cards were placed on the table directly above children's paper and remained in view throughout the drawing. In general, children were asked to look carefully at each card and copy the drawing (or complete their own drawing) so that "it looks exactly like this" (point to the card). There was no explicit reference made to the particular spatial relationship between the items of the array. If subsequently children asked questions

[3] For example, one card showed a hill with a sun partially visible behind one of its slope (occlusion task) and an upright tree resting on its other slope (vertical axis task)

about what should be drawn, the original instructions were repeated. On completion of each array, the card was removed and children were instructed (or helped) to turn to the next page of their booklet. All drawings were done freehand, whereas the order of drawings' construction and other relevant information were recorded.

3.4 Scoring

For the scoring of the standard measures, the specific instructions in the test manuals were followed. Occasionally the author consulted a Charted Clinical Psychologist to resolve uncertainties over the scoring of particular protocols. Specific details for the scoring of the Embedded Figure Test (PEFT & CEFT) and the GHDT will be found in Chapter III at Methods section. Samples of HFDs are also provided in Appendix A.

The drawing tasks demanded the selection of scoring criteria and the device of a scoring scheme. In the main, previous studies provided the empirical foundation for setting the drawing criteria, which in all cases were multiple, accounting for both successful drawing responses as well as alternative strategies. For matters of convenience and to elucidate the link between scoring procedures and theoretical considerations, the scoring method will be presented in Chapter V for occlusion tasks, in Chapter IX for horizontal tasks and in Chapter X for vertical tasks. Scoring procedures peculiar to individual tasks are described at relevant sections in relation to these tasks and sample drawings are presented in Appendix B.

CHAPTER II

MEASURING INTELLECTUAL MATURITY WITH HUMAN FIGURE DRAWINGS

1. A HISTORICAL OVERVIEW

Early interest in children's art dates from the late nineteenth century. Approximately between 1885 and 1920s, educators and psychologists from several different countries started to collect children's spontaneous drawings in a attempt to describe and classify them. The paper by Ebenezer Cooke 'Our art teaching and child nature' (1886) is one of the first publications solely devoted to children's artistic development, whereas Corrado Ricci's monograph *'L'Art dei Bambini'* (1887) possibly constitutes the earliest collection of children's drawings on the record. A few years later, the construction of objective criteria for the taxonomy of drawings originated, with the work of Schuyten (1904) and Lobsien (1905) being the first systematic approaches in the study of human figure drawings (HFDs). Coincidentally, the collection of drawings under rigorously controlled procedures was undertaken by Lamprecht (1906). Along with the conspicuous observation that the sophistication of children's artwork increased with age, Claparède (1907) also proposed the investigation of the possible relationship between drawing performance and intellectual ability as shown by school work. Ivanoff (1909) confirmed this conjecture after scoring the drawings of Swiss children on a 6-point scale and finding that they correlated positively with teachers' ratings.

These early studies were instrumental in subsequently formulating stage theories to explain the path of pictorial representation. Three researchers possibly made the most significant contribution in outlining developmental sequences; Kerschensteiner (1905), Rouma (1913) and Luquet (1913, 1927/1977). Kerschensteiner, studying thousands of drawings by schoolchildren in Munich, proposed a rough age-related succession from schematic and visually realistic forms to drawings capturing the three-dimensionality of space. His ideas were latter embodied and developed by Luquet (1913, 1927/1977) into a five-stage unifying theory of drawing development which eminently influenced the subsequent work of Piaget. At the same time, Rouma (1913) conducted a very extensive and well-designed study as he collected drawings made by normal and learning disabled schoolchildren over a period of 10 months. His book, which is amply illustrated by example drawings, furnished a picture of developmental aspects particularly in terms of HFDs. This study contributed new insights to the understanding of the topic; not only did he describe stages largely born out by subsequent investigations but he also discerned that the changes that the human form undergoes are more closely related to children's increasing mental age rather than to their chronological age.

The studies conducted and the theoretical formulations developed until then set up the working framework for the study of children's drawings and marked the beginning of a new era. The most influential achievements of these early investigations can be encapsulated as follows. The initial exploratory studies disclosed the developmental character of drawings and showed that the graphic forms undergo both quantitative and qualitative changes. They also laid the basis in identifying a number of typical features in the drawings of children of different ages. Developmental trends were next delineated by stage theories which, by putting some order on the plethora of obtained information, offered a lucid description of the course of development and an inclusive account of the nature of this progression. The gist of these ensuing theories was the recognition that the art, or act, of drawing is related to conceptual maturity. When Luquet proposed that an internal model (*modéle interne*) guides pictorial representation and suggested age-related drawing stages, the prospect of using the graphic medium to assess intellectual development became a theoretically defensible and promising enterprise.

A more analytical temper seems to characterise subsequent work. Researchers adopted rigorous sampling techniques and standardised methodological procedures in order to devise an objective scoring system and to develop valid age norms. In 1921 Cyril Burt published the book *Mental and scholastic tests* where he offers a crude age scale for HFDs. In the same manner that later Harris developed the Quality scale for HFDs, Burt selected a median sample drawing for each year and described its typical characteristics. Although he stated that the drawings didn't correlate highly with intellectual ability, he acknowledged their usefulness with children whose verbal or writing skills are limited.

The culmination of this energetic research into the psychology of children's drawings was reached in 1926 with the construction of the Draw-A-Man test (DAM) by Florence Goodenough. The test was an important psychometric achievement of its day and vindicated the supposition that an intellectual component underlay drawing development. Thus, Goodenough pinpointed the theoretical link between the study of children's drawing with the psychometric study of intelligence which dominated the literature up to 1940. The selection of the drawing topic for this purpose was further investigated by subsequent research. Although developmental trends were evident with various drawing subjects (animals: Du Bois, 1939; Desai, 1958; ships: Bender & Wolfson, 1943; houses: Kerr 1937; Markham, 1954), the human figure appeared to be the stimulus of choice for the following reasons: it seemed to be the most frequently drawn topic (McCathy, 1924), often in the absence of any other contextual forms (Lark-Horovitz *et al.*, 1939), and its predominant features yielded the least variance as a result of being exposed to different environments.

Goodenough recognised that children's HFDs not only reveal conceptual but also emotional aspects and discussed its application in the study of children's personality traits in addition to intelligence[1]. This idea began to take hold around 1940 with the advent of projective techniques and was actualised by Buck (1948, House-Tree-Person test) and Machover (1949), whose Draw-A-Person test was firmly rooted in the psychoanalytic thought. The problematic aspects and assumptions of Machover's work were underscored later by Roback (1968) and Swensen (1968). After an extensive review of empirical findings, they argued against the specific interpretations of individual signs (features) and expressed their reservations about its use in the diagnosis of pathology. Later, Koppitz (1968) attempted to link the two opposing approaches –using drawings either exclusively as a developmental test or a projective method – and devised a test on HFDs assessing both intellectual maturity and emotional adjustment.

An affective component is possibly present in children's drawings and a number of drawing tests continue to be developed for the assessment of children's emotional adjustment (Kinetic Family Drawings: Burns & Kaufman 1970; Draw-A-Person: Naglieri *et al.*, 1991). However, the scope of this thesis is delineated along the relationship of children's drawings to the development of cognitive abilities. As such, the rest of the chapter will review the literature that falls readily into the use of the DAM technique for the assessment of intellectual abilities.

2. THE DAM TEST AND ITS REVISION

When Goodenough devised the DAM test, the drawing was scored on a 51 point scale for the number of details, correct proportions between body parts, and motor co-ordination as

[1] Many researchers have used HFDs to investigate the body-image hypothesis and sex-role identification.

demonstrated by the fluency of lines and integration of parts[2]. By adding all the items present in the drawing, an overall score was obtained known as the Goodenough Intelligence Quotient. Early research on the properties of the scale suggested that performance on the test was related to gender (Goodenough, 1926; Dennis, 1942; Ford *et al.*, 1974) social class (Adler, 1970; Georgas & Georgas, 1972) and ethnic background (Lindler, 1962; Manuel & Hughes, 1932) and that learning disabled surpassed normal children on the scoring of details whereas the reverse was true regarding proportion and organisation (Israelite, 1936; McElwee, 1932; Spoerl, 1940). Finally, the original allowance of Goodenough that the test levels off approximately at the age of 12 was borne out (see Harris, 1963, p. 20).

From its development until the late 50's, the Goodenough DAM enjoyed remarkable popularity world-wide and it was ranked among the top 15 instruments mostly used in the clinical setting (Sundberg, 1961). Scott (1981) noted that this was an impressive finding considering that the test was essentially used with children. Its success seems to be attributed to its following attractive features: Drawing, by being a natural activity for young children, rendered the assessment of intellectual abilities by this means unobtrusive and non-threatening. Performance in the test remained largely unaffected by various factors related to its administration and scoring. For example, the test could be administered individually or in groups equally reliably (Harris, 1963), and the degree of expertise of the examiner was not related to the scoring of the drawing (Evans *et al.*, 1975; Harris, 1963, p. 21; Stricker, 1967)[3]. Its brevity and non-verbal nature also rendered it ideal for the assessment of children with limited attention span and/or language skills (young children, bilingual, linguistically or

[2] The items were selected after rigorous formal analysis and their validity was determined using both chronological age and school grade.

[3] Group versus individual administration was investigated by Harris (1963) with the GHDT. Using a counterbalanced design with 8th graders, he found no significant result. However there was some indication that some activities preceding the time of testing might exert some influence on the drawing performance (i.e. physical exercise, art instruction, Harris, 1963, p. 23, 30-31, 93) whereas inconclusive results were obtained in terms of whether the sex of the administration affects the sex of the figure drawn (see, Farylo & Paludi, 1985)

auditory impaired, children unwilling to interact verbally). In addition, the universal appeal of the drawing topic and the non-academic content of the test allowed it to be considered as less culturally loaded compared with conventional tests of intelligence (e.g. Wechsler Scales).

The favourable response of the scientific world for this new method of assessment prompted Harris to undertake a major revision of the DAM test to enhance its psychometric properties in response to former recommendations (e.g. Buros, 1953). In 1963 Harris introduced the revised system as the Goodenough-Harris Drawing Test (GHDT). The domains that were improved and modified were the following:

1. Harris validated the scoring items using more rigorous procedures[4] and revised the scoring system. He also added more items (73 compared to 51) for the drawing of the Man which, by sampling different aspects of cognitive ability, aimed to increase the reliability and validity of the scale (see Appendix A, Table GH-M). Some of these new items were found only in the drawings of adolescents in an attempt to raise the ceiling of the DAM test beyond its 12-year-old limit.

2. Two more drawings, the woman and self, were introduced as possible alternatives to the man drawing. The observation that young girls tended to draw a same-sex drawing and were losing points on items related to masculine details, led to the construction of the Woman scale which consisted of 71 items (see Appendix A, Table GH-W). The drawing of the self was included as a more appropriate stimulus for the assessment of personality and affective variables. However, no scale was devised specifically for the self-drawing which was supposed to be scored according to the Man or the Woman scale. Harris also stated that this increase in the number of drawings demanded was within the capacities even of pre-

[4] Three requirements needed to be satisfied for the inclusion of each item in the scale: (a) a regular and (b) rapid increase in the percentage of children passing each item at successive ages and (c) discriminative power of GH items for children of the same age range but different school grade.

schoolers, although some reluctance from older children was noted with respect to the self drawing.

3. The Man scale (MS) and Woman scale (WS) were standardised on more representative samples ranging from 5 to 15 years of age, drawn from four areas of the United States and controlled for the occupational distribution of parents. Ratio IQ scores in the original DAM test were replaced with deviation IQs for all the scales. In the pertinent section of his book (1963, p. 88) Harris argues that this substitution was called for as the more recent method of measurement used the standard score formula (e.g. Wechsler). Likewise, he provided tables for the conversion of raw scores to standard scores which translated the mean of the distribution of raw scores to 100 and the standard deviation to 15, at each age level.

4. To further enhance its utility as an economically administered test, Harris developed Quality scales for the man and woman drawings. Thus, a global assessment of children's conceptual maturity was feasible by comparing their drawings to a set of 12 model drawings of varying degree of sophistication.

5. The notable gender differences obtained from Harris' standardisation sample (1963, p. 126-130) corroborated similar findings reported by Goodenough (1926, p. 13, 56ff) and called for the development of separate norms for the sexes. To rectify the overall superiority of girls compared to boys – more noticeable in female drawings – Harris credited boys and girls of the same age with equivalent standard scores when girls obtained a slightly better raw score.

Finally, Harris introduced a significant qualification in the rationale of the test. He regarded the GHDT as a measure of intellectual, or preferably, conceptual maturity and not of intelligence as previously sustained[5]. In the introductory chapter of his book, he states that this modification follows the changes in the conceptualisation of the notion of intelligence of

[5] This qualification also implies that the test presumably measures a child's actual rather then potential level of intellectual functioning.

his times. The long-held view of "general intelligence" as a unitary construct was abandoned in favour of the idea that it embodies a large number of mental abilities (operations) which gradually become so complex that psychologist referred to them as cognitive processes. The products of these processes are the concepts. Following this, he maintained that intellectual maturity involves the process whereby ideas gradually lose their concrete properties and become more abstract. Harris considered that the mental operations involved in the process of concept formation are the ability to perceive and discriminate similarities and differences, to generalise and abstract them, and classify objects accurately. Based on the premise that pictorial representation requires conceptual understanding, Harris concluded that one can use the HFD as an index of the children's complexity of the pertinent concept and a rough guide to their developing concepts in general.

Since 1963, the literature on the revised GHDT has been voluminous. The major research interests have been on special populations (culturally disadvantaged, learning disabled, mentally retarded, emotionally disturbed) and on the predictive validity of the test as a measure of intellectual functioning and academic achievement. The remaining section of this chapter will review empirical findings from subsequent studies concerned with the psychometric properties of the test. The principal focus will be on normal population. Cultural influences, perceptual-motor skills and personality variables will not be discussed.

2.1 Normative data

2.1.1 Test inflated norms

Towards the end of 1960s a national survey on children's health status was undertake in which Harris was involved. The GHDT was administered to thousands of non-institutionalised American children between the ages of 6 to 15, rigorously sampled according to many demographic characteristics and socioeconomic variables. During 1970s, a number of papers, published from this survey, revealed that from the age of 7 children obtained mean standard scores approximately 10 points below the test mean at each age level (Harris & Pinder, 1974; Harris & Roberts, 1972; Harris *et al.*, 1970). Subsequent studies seemed to confirm this initial finding. Scott (1981), reviewing published papers for 1963 to 1977, maintained that the upward bias in Harris standardisation norms was upheld by a large number of studies which showed that GH norms were generally inflated and this pattern was more evident for children younger than 7 and older than 12 years of age. In addition, although the majority of studies were conducted with MS, there was a clear indication that this bias was more pronounced for the norms of the WS relatively to those of the MS.

However, results from more recent large-scale studies are inconclusive. Sutter & Bishop (1986) and Oakland & Dowling (1983) tested children of 7 and 6 to 10 years olds respectively, using rigorous sampling procedures and large sample sizes. The first authors report a mean score for the total sample whereas the second provide mean scores at each age level. In any case, there was no evidence in their results in support of the claim that GH scores underestimate children's abilities. Comparable results were obtained by Carvajal *et al.* (1987) whose 8 year old sample obtained a mean score only 3 points lower than the test mean (100). On the contrary the 10 point upward bias in GH norms was found in the study of Mehryar *et al.*, (1987) who used a large subject pool (N = 1.195) stratified according to various criteria. Although this pattern was present at all age levels from 6 to 11 years of age their results might not be that conclusive since the study was carried with Iranian children. Similarly, the study of Fabry & Bertinetti (1990), whose results from their 8 year olds sample are in line with those of Mehryar *et al.*, suffers from methodological shortcomings as their sample had a wide age range (N = 31; 6-10 yrs.) and was possibly not as representative of normal population[6].

[6] The authors reported that all children had been referred for various problems. However they claimed that they were of average intelligence.

Scott (1981) offered two explanations for the presumed upward bias in Harris' standardisation norms; both were related to his sampling procedures. First she stated that Harris' sample was drawn from 10 different geographical regions but it was not controlled for racial distribution and consisted only of children attending regular classes (Harris, 1963, p. 100). Other studies – and in particular the federal Health examination survey – sampled children from wider geographical locations (25 states) and didn't include school attendance as a sampling requirement. Thus, the exclusion of unschooled children and possibly the underrepresentation of some ethnic minorities from Harris' standardisation sample may have accounted for the upward bias in his norms.

In fact results from subsequent studies have not totally precluded the effect of race on performance at the test. Although some have found no significant differences between different racial-ethnic groups (black/white: Adams & Lieb, 1973) others report contradictory results (black/white: Harris & Roberts 1972; O'Keefe *et al.*, 1971; Mexican: Laosa *et al.*, 1974; American Indian: Cundick 1970). However, in the more recent study, Oakland & Dowling (1983) attributed the absence of a significant difference between their ethnic groups to their better sampling method. They suggested that at low SES levels ethnic minorities perform worse than the control group –something supported also from the results of Harris & Roberts (1972). However when SES was kept constant across the overall racial distribution these differences were eradicated.

2.1.2 Test ceiling and floor

Harris (1963) acknowledged that his attempt to extend the DAM test up to 15 years of age had been unsuccessful. Subsequent studies confirmed the conjecture that performance on the GHDT reaches a plateau approximately at 12 years of age (Harris & Pinder, 1974; Harris & Roberts, 1972). The recent literature on the GHDT however suggests that the test has been mainly used to assess presumably normal children up to the age of 10, whereas studies involving older children were predominately concerned with special populations. Although Goodenough and Harris reason that the test taps children's ability to form abstract concepts, it is more likely that for older normal children these concepts are readily expressed with instruments having a verbal component. The recognition that children seem to abandon this medium for self-expression in favour of language (see Arnheim, 1969; Gardner, 1980) by middle childhood – which coincides with the Piagetian stage of formal operations – corroborates the previous claim. Yet the limited number of scorable features in the scale or the shift of interest towards other drawing themes can somewhat contribute to the abrupt levelling off of the test.

Although Harris' standardisation sample ranged from 5 to 15 years of age, standard scores for children of 3 and 4 years of age are also provided. However, Harris suggested at a footnote (1963, p. 294) that these figures should be used as a "tentative guide" of pre-schoolers' abilities since they have been calculated from less representative samples[7]. The scarcity of normative data from pre-school children in subsequent studies does not allow the investigation of his claim[8]. Studies with young children had either used other scoring systems (Goodenough DAM: Ford *et al.*, 1974; Koppitz DAP: Groves & Fried, 1991) or were rather poor in their sampling procedures[9] (Short-DeGraff *et al.*, 1989). Review of the literature suggests that the main interest is either in the development of an abbreviated scoring method from the GH scales to account for the sparse drawings of pre-schoolers (Ayres & Reid, 1966; Goldman & Velasco, 1980; Short-DeGraff & Holan, 1992; Simner, 1985) or in the evaluation of the test's predictive validity with respect to school readiness (Simner, 1985; Willis, 1987). Thus, no firm conclusion can be drawn in terms of the test's performance with young children.

[7] Harris hinted that these values might be somewhat inflated.
[8] Groves & Fried (1991) provide normative data for pre-schoolers but they used a different scoring scale (Koppitz, DAP) and their results are reported per item only, and not on total drawing score.

2.1.3 Gender differences

When Goodenough (1926) originally devised the DAM test she observed that there was a moderate but consistent tendency of girls to outperform boys at all age levels, and that this difference was pronounced with particular items. However, she minimised these differences, claiming that they were mainly qualitative rather than quantitative. Harris obtained the same pattern with the revised GHDT. However, he maintained that these differences were substantial in all respects, particularly with the female figure drawing. On the basis of this, he developed separate norms for the sexes for both scales, where he corrected the obtained differences on mean raw scores. Also, by conducting an item analysis to further explore the gender differences, he confirmed most of Goodenough's original observations. He found that girls were significantly better rendering the head of the figures (hair, eyes, mouth), scored higher on all the items on motor co-ordination and did better on arm proportion. On the contrary, boys were better on feet proportion, on single facial features (i.e. nose), and their human form had plasticity by depicting some action. Further, for the drawings of Man, girls were not superior to boys with items relating to clothing, except that they were able to avoid transparencies at an earlier age than boys. Conversely, for the drawings of Woman, girls demonstrated general proficiency in all aspects of outfit (see Harris, 1963, p. 126-130).

Harris stated that Goodenough minimised the obtained gender differences in her sample since the prominent views of her time were in support of differences favouring boys (Burt 1921[10]; Kerschensteiner, 1905). Possibly certain modifications in the curriculum at school during the 60s might had changed the direction of the gender difference. Numerous other explanations had been also offered in terms of psychological, biological, cultural and personality factors. Thus, girls' general superiority in drawings was attributed to their faster maturation rate, their

[9] Short De-Graff *et al.*, (1989) recruited their 15 subjects from a screening clinic, and obtained a wide range of standard scores for the Man scale of the GHDT (i.e. 62 to 130).
[10] Burt found the eminent superiority of boys hardly surprising since curriculum and extra-curriculum activities for boys were favourable to their drawing development.

"docility and diligence", their earlier identification with their sex role and social interest and awareness.

Subsequent research on the nature of gender differences in HFDs generally supported the previous claims. Factor analysis on the revised GHDT (Sinha, 1970), as well as analysis per item (Brown, 1990; Groves & Fried 1991; Koppitz, 1968; Snyder & Gaston, 1970), confirmed the pattern of qualitative differences between the sexes. Although there were fluctuations at different age levels, in general girls were better elaborating the head whereas boys' graphic forms showed more flexibility and agility. Only Ford *et al.*, (1974), scoring the drawings of 3 to 5 year olds with the Goodenough technique, found that girls were significantly better in all aspects of the drawing (detailing and body-concept development). On the contrary recent evidence from experimental studies on HFDs suggests that boys were not ahead with respect to formal aspects, and that contouring, profiling and depiction of motion were found at comparable frequency in the HFDs of both sexes (Cox, 1993).

However, subsequent findings for the overall performance of the sexes on the GH scales are not encouraging. Scott (1981), reviewing results from 9 studies between 1967 to 1977, reported that there was no significant gender difference in mean raw scores in 24 out of the 33 comparisons across ages of 3 to 18. Only in 7 out of the 9 cases where significant differences were obtained, were these in favour of girls. Although the vast majority of research had focused on the MS, overall findings suggest that, when gender differences are present, they are more likely to be found with the female drawing (WS). Gender differences were absent in three studies testing children between 5 to 10 years of age with the MS (Coyle *et al.*, 1977; Oakland & Dowling 1983; Phillips *et al.*, 1973) and only Mehryar's *et al.*, (1987) study with Iranian children of 6 to 11 years obtained differences in favour of girls. On the contrary O'Keefe *et al.*, (1971), found girls' superiority only when performance on the

MS and WS was averaged and not within each scale separately, and similarly Strommen, & Smith (1987) didn't find gender differences comparing scores from same-sex drawings (MS/boys vs. WS/girls). Finally, Ward & Eliot (1987) reported that girls' superiority was present only for the WS and only Sinha (1970) found difference in the same direction for both scales (MS, WS).

Although these results are not firmly conclusive they are rather worrying. They don't fully support the suppression of girls' raw scores and raise the possibility that Harris' norms might result in a serious underestimation of girls' true abilities. It is very likely that gender differences, present at the time of test's construction, are absent nowadays or limited to certain ages or scales.

2.2 Reliability

2.2.1 Intra- and inter-rater reliability

The reliability of the Goodenough technique had been investigated by various methods which in general have yielded satisfactory results. The consistency of the rater's scoring method was examined by obtaining correlation estimates between the re-scoring of the same drawing by the same judge (intra-rater reliability) and by independent judges (inter-rater reliability). Intra-rater correlation values seem to fall in the low 90s (median value = .93, see Scott, 1981). Dunn (1967a), testing children of 6 to 11 years old with the MS, obtained an intra-rater correlation estimate of .93 from self-taught naïve raters with a one-week interval. This is comparable with Struempfer's (1971) value of .91 from 11 year olds with a 15-month interval between the self-scorings of the same drawing by experienced raters. However his value for the WS was somewhat lower (.83) [11].

[11] Struempfer's study was carried with Harris' Quality scales.

The degree of agreement between the scoring standards of different raters was investigated by Harris in a single extensive study with 8 and 10 year olds (Harris, 1963, p. 90). The values were in the range of 90 for both the MS and WS. Subsequent studies yielded overall values ranging from low 80s to low 90s. Three studies with 11 year olds obtained correlations for the MS around .90 between two raters (Strümpfer & Mienie 1968; Struempfer, 1971) and from .83 to .91 between three raters (Philips *et al.*, 1973). Similarly, the studies of Evans *et al.*, (1975) and Short-DeGraff *et al.*, (1989) with 5 year olds yielded a median value of .92 and .95 respectively. Although overall correlations from a wide age range were comparable (.88, in Abell *et al.*, 1996 with 5 to 15 year olds, and in Brown 1977 and Dunn 1967a with 3 & 6 to 11 year olds respectfully), Sinha (1970) demonstrated that these values can fluctuate substantially according to age level and drawing. Scorer reliability ranged from .65 to .88 across the ages of 7 to 11 and overall coefficients for the MS and WS was as discrepant as .81 and .98 respectively.

Few studies went beyond correlational evidence to investigate the nature and the origin of the obtained differences between pair of scores for the MS. Yule *et al.*, (1967) discovered that besides the high and significant inter-rater coefficients, significant differences existed between the mean scores assigned to the total group by each of their three raters. These results can be partly explained by Harris' previous claim that a rater eventually develops his own scoring standards which reduce the random error in his score and possibly render the inter-scorer variance systematic rather than random. Harris though didn't seem to consider that the idiosyncrasies of raters' scoring methods might originate from the insufficient guidance in Manual for the scoring of particular items. Philips *et al.*, (1973), performing an item analysis, identified 12 items whose ambiguous scoring criteria rendered the pattern of disagreement systematic rather than random. On the other hand, Evans *et al.*, (1975) found that at low levels of overall score differences (<3 points)[12] their three scorers were equally variable with no indication that they were scoring consistently leniently or harshly.

The authors attributed their results to the fact that they obtained scorer reliabilities from a younger age group (5 year olds) compared to previous studies (10 & 11). Their argument was that the immature drawing of young children can often cause uncertainty and discourage the use of consistent standards, more likely to develop with the advanced drawings of older children. These qualifications indicate that a considerable fluctuation in the intra- and inter-scorer reliability coefficients will be obtained as a function of age (i.e. Sinha 1970) and that certain ambiguities in the scoring method might further hinder the comparability of individual scores. On the basis of these results some authors are more cautious than others with the use of one score from a single judge and suggest that using a mean of judges' scores might be a more accurate assessment of children's abilities.

2.2.2 Test-retest reliability

Very limited information is provided by Harris with respect to the temporal stability of children's drawing performance. A single study yielded no significant differences in the scores obtained from the drawings of kindergarten children made on 10 consecutive days (Harris, 1963, p. 91). However, it is unclear which scoring technique was used (DAM, GHDT). Later studies investigating temporal stability after an interval of a couple of weeks (Brown, 1977; Evans *et al.*, 1975; McGilligan *et al.,* 1971) to three/four months (Mehryar *et al.*, 1987; Struempfer, 1971; Strümpfer & Mienie, 1968) yielded test-retest coefficients in the low 70s. However two studies obtained rather inconsistent results. Brown (1977), testing children of 3 to 11 years old, found no change in the items drawn between a two-weeks interval in 96% to 100% of the cases; a rather impressive figure. On the contrary, Sinha (1970), without specifying the interval between the two administrations, reported stability

[12] This accounted for the 85% of their cases of disagreement.

coefficients ranging from .57 to .74 across the ages of 7 to 11, with overall reliabilities considerably low (MS = .65, WS = .58).

Some anecdotal reports suggest that children's drawing performance can demonstrate great variability even within a short period of time, particularly at certain ages (Griffiths, 1945, p. 218; Kellogg, 1970, p. 191). Following this, some have proposed the development of norms at quarterly or biyearly intervals to enhance the scales' precision in the assessment of children whose skills develop rapidly (Sattler, 1982). However, there is always the possibility that this variability reflects variation in the testing procedures (instructions, materials) rather than in children's true abilities.

2.2.3 Internal consistency reliability

Measures of consistency regarding content sampling have been primarily investigated by the split-half method and by obtaining alternative-form reliability estimates from GH scales. Although the internal consistency of the revised scales was not examined by Harris, in two occasions he referred to split-half reliabilities of .89 and .77 for the Goodenough DAM test (Harris, 1963, p. 21 &106). Results from subsequent studies fall into the desired 80s to 90s range. Strümpfer & Mienie (1968) obtained a split-half reliability of .88 for the MS and .92 for the WS using the Spearman-Brown formula, whereas with the same method Mehryar *et al.*, (1987) reported values at low 80s for the MS. Two studies investigated item-homogeneity with the KR-20 formula for both scales. Although the values of Sinha (1970) ranged from .84-.90, those of Strommen & Smith (1987) varied from .63 to .92 for different subgroups. Possibly this variation can be attributed to the relative smaller sample size of the age groups in the latter study and might not be as troublesome since the overall interitem coefficients remained at the high 80s for both scales.

The relationship between children's performance on the MS, WS and SS has not been investigated at length partly due to dearth of research comparing the three scales. Harris reported correlation coefficients between the MS and WS across the ages of 6 to 14. The values, ranging from 71-81, didn't vary substantially with age or gender and reached the mid 70s for the total group and each sex separately (1963, p. 106). Harris noted that these values are comparable to the split-half reliabilities and thus recommended that the two scales can be regarded as alternative forms. Later studies reported similar correlational estimates between the scales. Two studies with 11 year olds yielded correlations of .78 (Struempfer 1971; Strümpfer & Mienie 1968), Dunn (1967b) with 9 year olds obtained a value of .87 using the standard scores, whereas McGilligan *et al.*, (1971), testing 6-8 year olds, reported a correlation of .75, identical with that of Harris.

However some anomalies exist in the results from three studies investigating the relationship between all the GH scales. The study of McGilligan *et al.*, (1971) is the only which reported inter-scale correlations by age and gender. With respect to MS and WS, boys obtained values 10 points lower compared to girls, overall and at each age level. Further, when the adult drawing (Man/Woman) was compared to a same sex drawing (Self), the gender differences were even more pronounced and the direction was reversed at each level. In two more recent studies using standard scores, correlation coefficients fell below the mid 70s between all the scales (Bardos *et al.*, 1989), and only between the MS and WS in the study of Short-DeGraff *et al.*, (1989). Further, in the latter study the magnitude of correlations varied substantially between any two scales, with the correlation between the drawings of woman with self almost 20 points higher than the one between man and self. The authors call upon the order of scales' presentation and the possible interplay of certain psychological factors in explaining the closer relationship between the WS and SS. Yet, no firm conclusion can be drawn since the authors applied a different scoring system for the drawings of self, used the raw scores from these drawings but the standard scores from the adult drawings, and studied a single age group. However, these results suggest that the relationship between the scales might change with age and gender, and no firm conclusion can be drawn as to the comparability of the self scale with the other two, due to insufficient evidence.

2.3 Validity

Harris did not investigate the validity of the new GHDT during his revision. Instead, he reviewed findings from correlational studies investigating the relationship of Goodenough DAM test with criterion measures of intelligence (Harris, 1963, p. 35-36, 95-99). The majority of the studies used the Stanford-Binet (9 studies) and the Wechsler scales (5 studies) whereas the remaining studies (7 studies) investigated validity issues with miscellaneous measures (i.e. Porteus Maze Test, Primary Mental Abilities Test, the Raven's Progressive Matrices). Half of the studies were conducted with special populations and the other half with normal children across a wide age range (kindergarten to 15 years). Validity coefficients albeit positive and significant were notably variable, ranging from mid 10s to 90s, depending on the age range and the characteristics of the sample, as well as the criterion measure and the scores used.

The fluctuations in the magnitude of correlations were not systematic. Correlations of the same range were found with the major individual tests of intelligence. The DAM test yielded coefficients with Stanford-Binet between .26-.92 with a median value of .60, and with WISC between .13-.70 with a median value of .50. Although it might seem that the drawing test was somewhat more closely related to Stanford-Binet, some of these correlations were calculated using mental age (MA) scores rather than IQ scores, naturally inflating the strength of the relationship. Further, the correlation coefficients at the upper and lower limit of the aforementioned ranges were obtained by special groups, and a comparable magnitude of coefficients was also obtained with normal samples. Thus, the pattern of relationship between the Goodenough DAM test with conventional measures of intelligence was as variable with normal children as it was with more diverse populations.

Harris cited results from a few studies which reported correlations between the DAM test and subtests/scales of criterion measures of known factorial composition. He claimed that the overall results corroborated Goodenough's initial assertion that the drawing test is equally related to performance as well as to verbal skills. However, he speculated that the contribution of various mental abilities to drawing performance might change with age. Comparing results from two extensive studies[13] which administered the same criterion measure separately to kindergarten and ten year old children, he noted that, although the drawing test score correlated fairly closely with the criterion test score in both samples (.41, .46), coefficients with its particular subtests varied at the two age levels.

The findings suggested that at pre-school ages the Goodenough score was more strongly associated with numerical ability and motor coordination skills, but at the age of 10 it was more closely linked with reasoning, perceptual speed and accuracy, whereas spatial aptitude seemed to play a role in both age groups. Thus, he came to the following conclusions:

a) the DAM test measures comparatively more successfully intellectual abilities than other variables (personality factors, artistic skills)
b) the significance of the drawing test as a measure of intellectual maturity is likely to vary with age. Mental components such as verbal abstraction and perceptual speed might be more readily tapped at older ages and quantitative and spatial concepts at younger ages. However, on the basis of sparse evidence no firm conclusion was drawn.

[13] Harris in an unpublished study tested 164 kindergarten children and Ansbacher (1952, cited in Harris, 1963) 100 ten year olds with the Primary Mental Abilities test.

c) Yet, in view of the variable relationship between the DAM test with criterion measures he conceded that "its value as an index of intelligence is perhaps not quite so firm" (Harris, 1963, p. 36).

In reviewing subsequent research on the utility of the GHDT in intellectual assessment, the best word to describe the findings would be inconsistent. Despite extensive research for more than 30 years, evidence to date indicates that the relationship between performance on the GHDT and intelligence remains equivocal. Scott (1981) summarised results from 40 studies up to 1977. Correlations of the revised GHDT with measures of intelligence stayed in the range of .24 to .83, with a mean value of .49. A review of studies after 1977 as well as those not included in Scott's report is summarised in Table 2.1. Inspection of the validity coefficients indicates that the values are in keeping with previous findings; correlations between the GHDT and comprehensive measures of intelligence are still mediocre.

Although most validation studies were conducted with the Man scale and little information is provided regarding the Woman or the Self scale, the mean correlation for the Woman scale was determined to be approximately .59 (Scott, 1981). However, there is some inconsistency between studies; some yielded comparable validity coefficients for the MS and WS (Dunn, 1967b; Sinha, 1970; Struempfer, 1971) whereas in others there was a difference in favour of the MS (Croake *et al.*, 1973; Strümpfer & Mienie, 1968). There is also a dearth of studies investigating the strength of relationship between the Goodenough DAM test and GHDT with criterion measure of intelligence. Limited research suggests that the Goodenough scores may be more accurate estimates of IQ scores than the GH scores (Phillips *et al.*, 1973; Vane 1967; see also Scott, 1981). Yet, due to insufficient evidence no firm conclusion can be drawn with respect to which of the two drawing systems or GH scales is more closely related with criterion measures of intelligence.

In view of the mediocre concurrent validity coefficients, a number of authors didn't preclude the possibility that these values might have attained significance due to the large size of their samples. They also recognised that although statistical significance was established in the majority of the cases, practical significance is poor as the validity coefficients appear to account for a small percent of common variance between the measures (Aikman *et al.*, 1992; Oakland & Dowling, 1983; Strümpfer & Mienie, 1968; see also Abell, *et al.*, 1998 & 1996). However, some others maintain that correlations of this magnitude are acceptable, considering that the drawing test is a brief screening measure using a nonverbal medium whereas more extensive batteries of intelligence include a more verbal crystallised component.

Some researchers went beyond the common practice of reporting a single validity estimate of GH scores with an overall IQ score from a criterion measure. Instead, they explored the convergent validity of the drawing test by establishing its relationship with subtests/scales of these measures. The expectation that the drawing score should correlate more highly with the performance rather than with the verbal aspect of IQ received either minimal or full support (Abell *et al.*, 1998 & 1996; Fabry & Bertinetti, 1990; Oakland & Dowling, 1983; Short-DeGraff *et al.*, 1989; Short-DeGraff & Holan, 1992; White, 1979) or it was well refuted – in keeping with Harris' premise (Dunn, 1967b; Laosa *et al.*, 1973; Struempfer, 1971; Strümpfer & Mienie1968; Wisniewski & Naglieri, 1989). A few studies not only reported correlations between GH scores with PIQ and VIQ (or other similar measures) but also provided coefficients across the whole range of subtests (Dunn, 1967b, c; Sutter & Bishop, 1986;White, 1979). The pattern was rather variable in most of the studies.

Specifically, although Dunn (1967c) obtained comparable correlations between GH scores with PIQ and VIQ, he observed that those with the verbal subtests of the WISC were consistently lower than those with the performance subtests. He also claimed that the overall results from children between 6 to 10 years of age suggested that performance on the GHDT is more strongly related to performance on the Block Design subtest. A similar conclusion was also drawn by Laosa *et al.*, (1973) and White (1979). On the contrary, while Sutter & Bishop (1986), testing a larger sample of the same age range, corroborated this different magnitude of correlations across the subtests of the Wechsler scales, the highest *rs* were obtained with the Object Assembly and Coding subtests. In sharp contrast are the results from two earlier studies; Dunn (1967b) obtained higher correlations with the verbal subtests for the WS although the reveres trend was seen for the MS, whereas Strümpfer & Mienie (1968) with 11 year olds and a different criterion measure found the strongest relationship with both verbal and non-verbal subtests.

Inconclusive are also the results from instruments using exclusively a non-verbal medium. As expected the validity coefficients particular with other drawing tests were quite substantial (*House-Tree-Person Test*: Abell *et al.*, 1998; *McCarthy Draw-A-Boy/Girl Scales*: Fine & Tracy, 1968; *Rutger's Drawing test*: Croake *et al.*, 1973; *Naglieri Draw-A-Person*: Bardos *et al.*, 1989; Naglieri, 1988). However, results from the studies, summarised in Table 2.2, suggest that the GHDT is not more allied to instruments which load heavily on performance factors (i.e. Bender Gestalt Test, Embedded Figure Test, Test of Visual-Motor Skills) than it is with more comprehensive criterion measures. In general, visuo-motor and graphic abilities appear to account as much as the ability to develop abstract-verbal concepts in performance on the GHDT. Again, inconsistent results were obtained as to their effect at different ages. Some researchers suggested that the relationship between visual-motor abilities and drawing performance tends to increase with age (Oakland & Dowling, 1983; Phillips *et al.*, 1973) whereas others found substantial correlations with perceptual-motor abilities at young ages (Short-DeGraff & Holan, 1992).

Taken together, although these results are a far cry from being conclusive, they suggest that a number of factors can change the nature of relationship between the GHDT and measures of intellectual abilities. As Harris speculated, the test seems to tap different abilities at different ages. Sample sizes, age range, sampling procedures and type of criterion measure against which the GHDT is evaluated can all affect the magnitude of the relationship.

A number of studies went beyond correlational evidence to investigate how closely the actual drawing and IQ scores approximate each other. The findings showed that the GH scores did not closely approximate IQ scores (Gayton *et al.*, 1971; Reisman & Yamokoski, 1973), and in spite of significant correlations they underestimated children and adolescents either in every aspect of intellectual functioning (WISC-VIQ, PIQ and FSIQ: Abell *et al.*, 1998 & 1996; Stanford Binet: Carvajal *et al.*, 1987; Johnson & Johnson, 1971) or in certain areas (PIQ: Wisniewski & Naglieri, 1989[14]). A number of researchers went further, hypothesising that the GHDT might have differential validity with individuals of various ages and intellectual abilities. The fact that in a number of studies diverse samples on a number of attributes (age, SES, intelligence) had been treated as homogeneous might have obscured the effectiveness of the GHDT in the assessment of intellectual abilities of individual cases.

The few studies which had stratified the subjects into appropriate IQ levels showed that the discrepancy between drawing scores and IQ scores is more apparent with children of average and above-average intelligence (see Scott, 1981). Empirical findings suggest that the GHDT is a more accurate estimate of IQ scores at lower levels of intelligence (Ables, 1971; Abell *et al.*, 1994; Aikman *et al.*, 1992; Wells & Pedrini, 1971) either on the basis of the magnitude of correlations or in terms of the classification accuracy rates. However, some contradictory findings have been reported with normal and disadvantaged children (Abell *et al.*, 1996;

[14] A different drawing technique was used in this study namely the Naglieri DAP test.

Johnson & Johnson, 1971). In addition, results from normative data and rate of drawing development suggest that intellectual abilities seem to be more readily tapped at younger ages via the human drawing technique (see Mortensen, 1991). A recent study by Abell *et al.*, (1996) investigated the differential validity of the GHDT at two age levels. Validity coefficients with all the IQ scores from WISC-R were notably higher for children younger than 10 years of age. However the authors claimed that when the *r*'s were transformed to Fisher's *z*, no significant differences were found in the magnitude of correlations between the younger and the older children.

2.3.1 Academic Achievement

Correlations between academic achievement and GH scores have yielded lower coefficients than those with IQ scores. Scott (1981), examining empirical findings up to 1977, concludes that global ratings of academic achievement correlate with drawing performance at around .30. Results from recent studies, summarised in Table 2.3, corroborate her appraisal. Also, some studies yielded correlations of the same magnitude across a range of skills (Oakland & Dowling, 1983; White, 1979), some found a closer relationship between GH scores with arithmetic and computational skills than with reading-related skills (Sinha, 1970; Sutter & Bishop, 1986) whereas other obtained the reverse pattern (Struempfer, 1971; Strümpfer & Mienie, 1968; Pihl & Nimrod, 1976). Yet, the latter studies were conducted with older children (>10 years), whose verbal abilities might have contributed more to their drawing performance.

2.3.2 Special Populations

A final note deserves to be made in terms of the GHDT's performance with children having learning difficulties, mental retardation or behavioural problems. Comparisons across studies are difficult to make since these terms are ill-defined, the root of disabilities are often variable within and between studies, and generally such studies lack normal controls. Some have reported encouraging results, finding that GH scores were higher than IQ scores and correlated substantially with all aspects of WISC-R for children with cerebral palsy (Kifune, 1984) and learning disabilities (Tramill *et al.*, 1980). On the contrary, others obtained discrepant scores and poor relationship between performance on the drawing test and conventional tests of intelligence with learning disabled (Pikulski, 1972) and behaviourally disturbed (Lehman & Levy, 1971). Some authors endorsed the view that factors often associated with learning disabilities and behavioural disturbances – such as perceptual-motor deficits and body-image problems – might account for the discrepancies often observed between GH scores and IQ scores. However, this hypothesis was not upheld by empirical findings (see Scott, 1981).

In conclusion, the validity coefficients for the GHDT are mediocre. Mean correlation estimates with comprehensive intelligence tests are approximately .50 whereas those with batteries of academic achievement are even lower, around .30. Some researchers in defence of the drawing technique argue that screening tests, such as the GHDT, measure a different aspect of intelligence and obviously do not provide the same quality and quantity of information as more comprehensive instruments. Extensive batteries such as Wechsler scales operationalize intelligence differently and contain tests of achievement whereas the HF drawing test has not an explicit academic content. Besides, Harris warrants against the use of the drawing score as equivalent to IQ. Others maintain that its concurrent validity is weak and that screening portions or short forms of major intelligence tests are clearly superior alternatives for screening purposes (Kamphaus & Pleiss, 1991). On the positive side, the HFD technique has certain appeal on the basis of the popularity of the drawing theme among children and its non-verbal and less culturally loaded content. In addition the GHDT appears to be a reliable instrument with rater reliabilities and coefficients of internal consistency at the rate of .80 to .90, whereas test-retest reliability around 70s.

Obviously, a factor that can account for the obtained discrepancies and low validity coefficients between GH scores and IQ scores is the difference in the conceptualisation and measurement of intelligence. However, these discrepancies are likely to have been accentuated by various other factors namely the presumed inflated norms of Harris' standardisation sample and the outdated scoring system. The review of the literature on the GHDT showed that a number of issues are still unsettled. There is a dearth of findings evaluating the three GH scales in terms of aspects of reliability, accuracy of Harris's norms and gender differences. Further, factors which seem to affect the performance of the GHDT – age, gender, level of IQ – have not been adequately investigated. The next chapter deals with these aspects in an attempt to evaluate the psychometric properties of the DHDT in a more systematic way considering its continuous popularity.

Table 2.1 Validity coefficients for the GHDT from tests of intelligence.

	Population					Correlation coefficients (*r*)		
Reference	N	Age	Status	Instrument		Man Scale	Woman Scale	Self Scale
Abell *et al.* 1998	200	14-15	racially diverse	WISC-R	FSIQ	.29		
					VIQ	.19		
					PIQ	.35		
Abell *et al.* 1996	125	5-15	clinical	WISC-R	FSIQ	.51		
					VIQ	.40		
					PIQ	.57		
	74	5-12	clinical	Stanford-Binet		.37		
Aikman *et al.* 1992	216	6-18	clinical	WISC-R & WAIS-R	FSIQ	.48	.49	
Carvajal *et al.* 1987	23	8.2-9.10	normal	Stanford-Binet IV		$.40^{ns}$		
				PPVT-R		.25		
				CMMS		.22		
Fabry & Bertinetti 1990	31	6-10	diverse	WISC-R	FSIQ	.62		
					VIQ	.45		
					PIQ	.69		
Gayton *et al.* 1974	50	6-12	clinical	WISC	FSIQ	.64		
Oakland & Dowling 1983	188	8-10	diverse	WISC-R	FSIQ	.55		
					VIQ	.47		
					PIQ	.55		
Short-DeGraff & Holan 1992	32	nursery	normal	WPPSI	Information	.20		
Short-DeGraff *et al.* 1989	15	nursery	normal (?)	WPPSI	Information		$.36^{ns}$	$.27^{ns}$
Sutter & Bishop 1986	360	7-10	normal	WISC-R	FSIQ	.38		
					VIQ	.27		
					PIQ	.41		
Tramill *et al.* 1980	100	6.8-16.1	learning disabled	WISC	FSIQ	.55 boys .63 girls		
Ward & Eliot 1987	204	7-12	culturally disadvantaged	CFIT		.30		
White 1979	30	7-11	diverse	WISC-R	FSIQ	.55		
					VIQ	.50		
					PIQ	.57		

WISC-R = Wechsler Intelligence Scale for Children-Revised; WAIS-R = Wechsler Adult Intelligence Scale; WPPSI = Wechsler Preschool and Primary Scale of Intelligence; PPVT-R = Peabody Picture Vocabulary Test-Revised; CMMS = Columbia Mental Maturity Scale; CFIT = Culture Fair Intelligence Test.

Table 2.2 Validity coefficients with various non-verbal tests.

Reference	Population N	Age	Status	Instrument	Correlation coefficients (*r*)
Aikman *et al.* 1992	216	6-18	clinical	Bender Gestalt	-.30
Oakland & Dowling 1983	157	8-10	diverse	Bender Gestalt	-.34
Philips *et al.* 1973	337	11	normal	Bender Gestalt	-.49
		5	normal	Ability to copy forms	.33
Saracho 1985	300	3-5	normal	PEFT	.99
				ABC	.90 (mean)
Saracho 1986	480	6-8	normal	EFT	.05 -.22
				ABC	.87 -.97
Ward & Eliot 1987	204	7-12	culturally disadvantaged	CEFT	.34
Short-DeGraff & Holan 1992	32	nursery	normal	VMI	.55
				TVMS	.49

(C)EFT = (Children) Embedded Figure Test; VMI = Developmental test of Visual -Motor Integration; TVMS = Test of Visual-Motor Skills; PEFT = Preschool Embedded Figure Test; ABC = Articulation of Body Concept.

Table 2.3 Validity coefficients with measures of academic achievement.

Reference	Population N	Age	Status	Instrument	Correlation coefficients (*r*)
Aikman *et al.* 1992	216	6-18	clinical	MAT/PIAT	.36 (Man)
					.39 (Woman)
Mehryar *et al.* 1987	510	6-13	normal	Teacher's rating	.24 -.38 (Man)
	569	6-13		Examination Marks	.30 -.45
Oakland & Dowling 1983	188	8-10	normal	CAT	.34 -.46
Sutter & Bishop 1986	360	7-10	normal	PIAT	.29
White 1979	30	7-11.6		PIAT	.47

MAT = Metropolitan Achievement Test; PIAT = Peabody Individual Achievement Test; CAT = California Achievement Test.

CHAPTER III

INVESTIGATING THE UTILITY OF THE GHDT IN INTELLECTUAL ASSESSMENT

1. AIMS

The present chapter is concerned with the various psychometric properties of the GHDT. Besides the interest in norms, issues related to the reliability of the scoring method, the internal consistency of the test and the relationship between the GHDT and criterion measures of intellectual ability are thoroughly examined. Although these aspects have been previously investigated, the following weaknesses were identified in reviewing the literature. The majority of the studies were conducted with the Man scale, whereas there is a dearth of results from all three scales (Man, Woman, Self). Variables such as age level, and gender were not always considered when reporting estimates of reliability and validity. The validity of the test was mainly examined with correlational methods. The common practice was to obtain an estimate for the entire distribution of ages, ranges of intelligence and criterion test variables[1]. As the literature suggests, this can obscure the changing relationship between the test and criterion measures of ability at different levels of these variables. In addition very few studies went beyond correlational evidence and investigated the comparability of the mean scores of both measures. Finally, although the GHDT is still very popular among psychologists, research relating to the performance of the test is rather sporadic the last 10 years. In this respect, the following analysis, considering the aforementioned shortcomings, attempts to present a more comprehensive evaluation of the properties of the GHDT.

[1] With few exceptions, most studies report correlations between the GHDT and the Full Scale IQ without investigating its relationship with the different scales and subtests.

2. METHOD

2.1 Instruments

All children were administered the full Goodenough-Harris Drawing Test (GHDT) with Man scale (MS) first, followed by Woman scale (WS) and Self scale (SS). For the assessment of intellectual abilities, Wechsler scales were used since they are the most frequently included instruments in psycho-educational appraisal and validation studies of GHDT. As there was no single Wechsler scale which would be age-appropriate for all the children, the Wechsler Preschool and Primary Scale of Intelligence (WPPSI), designed for ages of 4 to 6½, was administered to the first two age groups (4 & 5½ year olds) and the Wechsler Intelligence Scale for Children (WISC-III^{UK}) was used for the two older age groups (7 & 8 year olds). An abbreviated version of WPPSI and WISC-III^{UK} was chosen in consideration of the large number of instruments and experimental tasks that needed to be administered to each child within a short period of time (approximately one month), and the limited attention span of the younger children[2].

The short form of the Wechsler scales consisted of six subtests; three from the Verbal and three from the Performance Scale[3]. The selection of the Wechsler subtests was based on the magnitude of correlations of each with the Full Scale and with the subscale to which they belong (Verbal and Performance)[4]. Further, recommendations from previous research

---

[2] Some studies even used one subtest (Information) from the Wechsler Scales as a rapid screening method and assessed the concurrent validity of GHDT on the basis of this (Short-DeGraff *et al.*, 1989; Short-DeGraff & Holan, 1992).

[3] The WISC-III^{UK} consists of 12 subtests, of which 2 are used as supplementary or alternative. From the 10 standard subtests, 5 are classified into the Verbal Scale and 5 into the Performance Scale. Similarly, the WPPSI comprises 10 subtests for assessing IQ which are equally divided into Verbal and Performance Scales.

[4] For the WISC-III^{UK} the three verbal subtests used here correlated more highly with the Verbal and the Full Scale IQs and with the sum of scaled scores for the Verbal Comprehension factor at each age level. Intercorrelations of similar magnitude were obtained with the three performance subtests administered here and the Performance, Full Scale IQs and the Perceptual Organisation factor (see Wechsler, 1992, p. 79 and Tables C.7-8). Likewise, for the WPPSI, the intercorrelations of the subtest scores with their corresponding Scale and Full Scale scores were satisfactory although the magnitude

assessing abbreviated scales of the WISC and WPPSI were consulted in choosing the subtests (Kaufman, 1972, Sattler, 1982; Silverstein, 1968a, 1968b, 1970, 1971). The two Verbal subtests were Information and Vocabulary and the third was Similarities from the WISC-III^{UK} and Comprehension from the WPPSI. The two Performance subtests common to both Wechsler Scales were Picture Completion and Block Design whereas the third was Picture Arrangement from the WISC-III^{UK} and Mazes from the WPPSI.

Finally, field independence (FI)/field dependence (FD) was assessed for primary school children with the Children's Embedded Figure Test (CEFT: Witkin *et al.*, 1971) and for nursery children (4 year olds) with the Preschool Embedded Figure Test (PEFT: Coates, 1972). This measure was included due to its perceptual/spatial component which renders it closer to the Performance than the Verbal Scale of conventional tests of intelligence (Anastasi, 1990). Thus, although it is not a measure of intellectual ability, the convergent validity of the GHDT can be further explored. It also has an attractive aspect as its perceptual element is regarded as part of a broader personality construct described as global *vs* articulated cognitive style. Consequently, it offers the chance to explore the relation of the GHDT with a construct bridging cognition and personality development. The relationship between figure/ground perception and spatial orientation, measured by the EFT, with the GHDT has been previous explored with younger (Saracho, 1985) and socioeconomically deprived children (Ward & Eliot, 1987) but the results were incompatible.

2.2 Administration

Each measure was administered following the standard procedures described in the pertinent manual. The order of administration was fixed. The EFT was administered first followed by the GHDT and the Wechsler scales. Each measure was administered on a single session

---

of correlations for the verbal subtests were higher compared to the performance subtests (see Wechsler, except for a few cases where the WPPSI was administered in two sessions, no more than four days apart. Each child was tested with all the instruments within less than a month.

2.3 Scoring Method

2.3.1 GHDT

All the drawings from the four age groups (115×3) were scored by the author according to Harris's instructions. The whole set of drawings is available at http://pics.psych.stir.ac.uk whereas examples are included in Appendix A-sample drawings. Before scoring, each drawing from the total sample (115×3; Man, Woman, and Self drawings) was coded and put in a random order so that the scorer was blind on child's age, sex and drawing performance on preceding Scales. This procedure was adopted following Phillips's *et al* (1973) observation that often their judges were inadvertently affected by knowing the age of the child or by judging the drawings of a single age group. To obtain intrascorer reliability estimates, the total sample of drawings was scored again after a five month interval following the same procedure.

2.3.2 Wechsler Scales

After obtaining a raw score for each subtest by summing the item scores, this was converted to a scaled score using the appropriate age-specific Tables in the Wechsler Manuals. To calculate Verbal IQ (VIQ), Performance IQ (PIQ) and Full Scale IQ (FSIQ) for each child, the sums of the three Verbal and the three Performance subtest scaled scores were prorated separately for each Scale[5], and the resulting Verbal and Performance scores were summed to obtain Full Scale scores. Again, the three scores (Verbal, Performance and Full Scale) for each child were converted to their IQ equivalents using the relevant Tables in the Manuals.

1967, Tables 11 & 14, p. 26 & 29)

[5] For this procedure, the specific instructions given in the Manuals were followed (Wechsler, 1992, p. 53-54; Wechsler, 1967, p. 42-43).

2.3.3 Embedded Figure Test (EFT)

The two versions of the Embedded Figure Test (EFT) had a series of pictures of complex figures (CEFT: 28; PEFT: 24) in which a simple form had been embedded, and children had to identify it. Performance was measured on the total number of figures where the embedded form had been correctly located and traced within a predetermined time limit. Thus, high scores indicated FI and low scores indicated FD[6].

3. RESULTS[7]

3.1 Reliability

3.1.1 Intra-rater reliability

The objectivity and the precision of the GHDT was assessed in this study by having each drawing scored twice by the author on two different occasions separated approximately by five months. Reliability coefficients for the overall sample and by scale, age and gender are reported in the Appendix A, Table R1, using raw scores. The correlations between self-scorings were .99 for all the scales and remained in the high 90s with age and gender considered. A negligible increase in these estimates occurred between the age of 4 to 5½ for the scoring of the self drawings. The present estimates are somewhat higher compared with earlier studies which report intra-rater reliability estimates with a median value of .93 (see Chapter II/ Section 2.2.1). Possibly this might be attributed to the fact that having to score a large volume of drawings on the first instance promoted the development of consistent marking standards, limiting intra-scorer variance. This might also apply to the scoring of mature drawings which present fewer occasions of uncertainty over the interpretation of human features.

[6] A number of studies with young adolescents and adults used the number of seconds taken to solution as an alternative measure of performance on EFT (Liben, 1978, 1991b). However, studies with younger samples (Signorella & Jamison 1978; Foorman 1981) assess FD/FI on the total number of pictures where the embedded form was correctly traced; consequently this was considered as a more appropriate measurement of performance for the age range tested.

Besides the fact that intra-scorer reliability coefficients were uniformly high and significant, an item by item comparison between the two self-scoring occasions was performed. This procedure revealed a number of cases where inconsistencies occurred due to insufficient guidance in the Manual. Similarly with previous findings (Phillips *et al.*, 1973; Sinha, 1970), the types of disagreement tended to be rather systematic reflecting certain ambiguities in the definition of some scoring items and criteria, and sometimes inadequate scoring examples. These problematic aspects along with some scoring suggestions are illustrated by example drawings from the sample in Appendix A-Scoring suggestions.

3.1.2 Interitem reliability

The internal consistency of the GH scales was investigated using the split-half method corrected by the Spearman-Brown formula. The reliability coefficients were .89 for the Man scale (MS), .92 for the Woman scale (WS), and .89 for the Self scale administered to boys (SS_M), and .88 for the Self scale administered to girls (SS_W)[8]. These estimates are directly comparable with previous findings reporting coefficients of internal consistency in the high 80s – for MS (.89) Harris, 1963, p.21 and Strommen & Smith, 1987; (.88) Mehryar *et al.*, 1987 and Strümpher & Mienie, 1968; (.87) Serwer *et al.*, 1972; (.84-.90) Sinha, 1970; for WS (.87) Strommen & Smith, 1987; (.92) Strümpher & Mienie, 1968.

3.1.3 Alternative-form reliability

Two early studies assessed only the comparability of MS and WS (Harris, 1963; Strümpfer & Mienie, 1968), whereas the few studies reporting correlation estimates between the three scales either tested older children (Bardos *et al.*, 1989; Dunn, 1967b; Harris & Pinder, 1974; McGilligan *et al.*, 1971) or a single age group (Burns & Velicer, 1977; Short-DeGraff *et al.*,

---

[7] Apart from results in Section 3.1.1, subsequent analyses of drawings had been performed on the drawing scores from the first self-scoring occasion.

1989). Table 3.1 reports the correlation coefficients between the three scales by age and gender[9]. The overall coefficient estimate for the MS/WS was .79, close to Harris's finding of .75 for children of 6-14 years of age (Harris, 1963, p. 106). Comparing the correlations of this study with those of Harris's by gender, a somewhat lower correlation was obtained for boys and higher for girls (boys: .71 *vs* .76; girls: .81 *vs* .77)[10] but these differences are not substantial. Comparable to the correlations between the MS and the WS were also the estimates for the Self and the same sex drawing in the overall sample.

Table 3.1 Correlation coefficients for the GH Scales by age level and gender.

Age		N	GH-Scales		
			Man-Woman	Man-Self	Woman-Self
4 year olds	Boys	14	.66[a]	.77	.70[b]
	Girls	20	.90	.88	.81
	Combined	34	.82	.84	.78
5 ½ year olds	Boys	10	.76[a]	.67[a]	.32[c]
	Girls	16	.92	.88	.85
	Combined	26	.89	.85	.77
7 year olds	Boys	8	.88[b]	.87[b]	.73[a]
	Girls	20	.89	.80	.84
	Combined	28	.89	.81	.81
8 year olds	Boys	10	.75[a]	.88	.83[b]
	Girls	17	.73	.92	.76
	Combined	27	.72	.88	.76
Total	Boys	42	.71	.79	.64
	Girls	73	.81	.82	.81
	Combined	115	.79	.82	.78

Note 1: The correlation coefficients are significant at .001 except in the following cases
a = the correlation coefficient is significant at .05
b = the correlation coefficient is significant at .01
c = NS
Note 2: All correlations are bivariate except for those for the cells for totals. These are partial correlations, controlling for age.
Note 3: Underlined coefficients between the MS and the WS with SS represent the relation of two same-sex HFDs differentiated in terms of age (man-boy vs woman-girl).

---

[8] The coefficients remained high when the analysis was repeated by gender. Boys yielded a reliability coefficient of .91 for both scales and girls .88 for MS and .93 for WS.
[9] Similar analysis was performed on standard scores and the results are presented in the Appendix A, Table R2.
[10] The correlation coefficients of the present study are underlined.

Further, no discernible trend among the correlation values was observed when these estimates were examined by age or by gender. For the MS/WS, the lowest coefficient was obtained by the 8 year olds, as in the case of WS/SS_W where 8 year old girls attained somewhat lower correlations compared to their younger counterparts. However, for the MS/SS_M the reverse pattern was seen among boys as 4 and 5½ year olds yielded lower correlations than their older counterparts. However, no statistical age differences were found in the magnitude of correlation coefficients for any two scales.

When the correlation values were examined by gender, girls obtained higher correlations for the MS/WS at younger ages compared to boys (4 & 5½). Also, at the first two age levels girls yielded higher correlation estimates for the WS/SS_W than did boys for the MS/SS_M but the reverse was true at the two older levels. Evidently, the unequal numbers of boys and girls in each age group affected the magnitude of the gender-related correlations. Subsequently, the statistical analysis showed that only at the age of 4 did the MS/WS correlate significantly higher for girls than it did for boys ($z = 1.75$, $p<.05$ for one tailed comparison).

Harris, commenting upon the modest inter-correlation of .75 between the MS/WS, suggested that the scales might tap different abilities due to the fact that they are partly made of a number of different items (see Appendix A, Tables GH-M and GH-W). He added that they might address a mixture of several abilities both common and different in both scales, and that factor analytic studies might be informative, identifying clusters of items loading on various dimensions. In fact Strommen (1987) undertook such a task and distinguished two dimensions for both scales. The first consisted of 14 core items, which simply assess the presence and absence of certain basic features, whereas the second contained 23 elaborate items. The total of these 37 items were taken from the set of 44 items common to both scales.

In the present study an attempt was made to evaluate Harris's proposition using Strommer's dimensions. Thus, inter-scale correlation coefficients were calculated isolating only the core items (14), the elaborate items (23) and also the items common to both scales (44). Full report of these values is given in the Appendix A, Table R3, and the main findings of this analysis will be summarised here. For the overall sample, the coefficients for MS/WS were .82 for the core items, .70 for the elaborate items and .78 for the common items. The striking correspondence of these figures with the one obtained from the total items in the scales (.79) clearly suggests that the differences between the items in the scales do not suppress the magnitude of correlation coefficients as Harris has suggested.

Again, for all sets of items (core, elaborate, common), the lowest correlation estimates for the MS/WS were obtained at the age of 8, but no significant age differences were found. Regarding gender, the correlations for boys were lower than those of girls for the core and common items of MS/WS (core: .68 *vs* .89; common: .69 *vs* .81), and this difference was statistically significant for the core items ($z = 2.97$, $p<.01$). Investigating gender differences within age levels, the same pattern of results was obtained at the younger ages (4 & 5½) for the core and common items, and at the ages of 4 and 8 for the elaborate items. Systematic comparisons revealed that only at the age of 4 did girls yield significantly higher correlation coefficients than boys for the MS/WS on core and common features (.65 *vs* .91, $z = 1.95$, and .70 *vs* .92, $z = 1.86$ respectively, both $p<.05$).

Looking at the relationship between the MS and the WS with the SS within gender, the range of coefficients was between mid 70s to .80 for all the sets indicating that the Self drawing can be used as an alternative form to the MS and WS. However, the magnitude of the correlations fluctuated within gender according to age and set of items. For boys, the same-sex drawings did not correlate on the core features at the ages of 5½ and 8 (.27, .40) and on the elaborate items at the age of 4 (.34), whereas at the age of 8 the correlation, albeit significant, was relatively lower than those of the intermediate age groups (.65 *vs* .83 for 5½ and .87 for 7 year old boys). On the contrary, for girls the correlations were significant but relatively lower at the two older age groups when the comparison was between the core features, and at the two younger groups when the comparison was between the elaborate features. Considering the relationship between the same sex drawings on common items the range of correlations was satisfactory for girls and boys at each age level (boys: .70-.92; girls: .67-.82).

3.1.4 Summary and discussion

The obtained correlation estimates for the MS and the WS in the overall sample is of similar magnitude to those reported elsewhere. Comparable are also the estimates for the Self and the same sex drawing with those of previous studies (McGilligan *et al.*, 1971; Harris & Pinder, 1974). McGilligan *et al.*, testing children of 6-8 years of age, reported median correlations between Self/Man for boys and Self/Woman for girls of .81 and . 83 respectively which are equivalent to .79 and .81 found here. As might be expected, self-drawing scores should correlate approximately equally with Man scores for boys and Woman scores for girls. The present results confirm this hypothesis in agreement with previous studies and refute Short DeGraff's *et al.*, (1989) claim who attributed their obtained lower correlation for pre-school boys compared to girls to the test order[11]. In the present sample, only at the age of 5½ was such a difference discernible in favour of girls whereas the reverse pattern was obtained at the age of 8.

The comparability of the obtained correlation coefficients between any two pairs of scores corroborates previous assertions that the scales tap similar abilities and can be used as

[11] The authors maintained that the administration of the MS first, followed by the WS and the SS could result in higher correlation between the last two compared to the first two as a result of practice effect. This might apply more to the case of very young children.

alternative measures. Further support is also provided from the results on the reduced sets of items between the scales. A note deserves to be made here with respect to the SS. The results show that the SS is reasonably comparable to the MS and WS although it was originally developed as a potential device to assess self-concept and personality. Possibly the fact that neither normative data nor a peculiar scoring system had been developed for this scale might have contributed to its parallelity with the other two scales.

Taking the overall results on the reliability of the GHDT the following conclusions can be drawn. Intra-rater estimates are remarkably high as a 99% agreement was achieved between the self-scoring of the same drawing on two different occasions. This value is higher than the inter-rater reliabilities which generally fall in the low 90s, supporting the view that eventually the scorer develops his own consistent standards which subsequently reduce rater variance (Harris 1963, p.90). Yet, on the basis of the systematic inconsistencies observed here and reported previously (Phillips *et al.*, 1973; Yule *et al.*, 1967), the revision of the scoring criteria is recommended particularly for certain drawing features. Appendix A-scoring suggestions deals with these issues and may be helpful to the researchers willing to take up such an endeavour systematically, as well as to practitioners in clarifying their uncertainties.

The correlation estimates for the between-scales consistency is approximately .10 points lower than the split-half reliabilities assessing the within-scale consistency (inter-scale: .79 MS/WS, .82 MS/SS, .78 WS/SS; split half: .89 Man, .92 Woman, .89-.88 Self). In view of these results it is reasonable to maintain that the scales tap overlapping skills; yet these are not truly parallel. It is therefore advisable to administer the full test which will provide a more global assessment of children's abilities. However when a quick screening procedure is required the administration of a single scale is justifiable.

Having completed the analysis on the reliability of the GHDT, the next section deals with issues related to the precision of Harris's normative data, comparing the performance of the present sample with that of the American standardisation group. Variables such as age, gender and scale are included in this analysis. Scale differences are explored over and above the correlational evidence from the previous analysis, since the scales can be related, yet their mean scores might be discrepant. Further, on the basis of empirical results (see Chapter II/Section 2.1.3), gender differences seem not to be as well-established as Harris had thought when he developed separate norms for the sexes, correcting for these differences. In the subsequent analysis this issue is also addressed along with its implication for the precision of assessment of children's abilities.

3.2 Sample performance

Before presenting the results from the statistical analyses, the obtained mean raw scores will be compared with those obtained from Harris's normative data for the standardisation of MS and WS (1963, p. 102-103, Tables 11 & 13). Comparisons cannot be made for the 4 year old sample as Harris's standardisation group ranged from 5 to 15 years of age. In addition, there are certain differences in the sampling procedures between the studies which need to be outlined first. Harris centred his age groups at the mid-year of the age interval, with an equal number of boys and girls. In the present study, girls outnumbered boys at each age level and only the 5½ year old sample seems to have the right age range for direct comparison with Harris's 5 year old standardisation group. Since each of the older samples here (7 & 8) fall centrally between two consecutive age groups in Harris's study (i.e. 7 year olds here: 6-7 in Harris's study), their mean scores were compared against an average of means of the two corresponding standardisation subgroups.

Table 3.2 presents the means and standard deviations of GH raw scores by age and gender for each scale separately. In general, the mean raw scores from this sample were somewhat higher than the American norms. For the MS, they were 1-3 points above Harris's values at the age of 5½, 1-2 points at the age of 7 and 4-5 points at the age of 8. For the WS they corresponded fairly closely to his results only at the age of 5½, where there is only 1 point of difference, with boys scoring lower and girls higher resulting in virtually identical means for the total sample. At the age of 7 the difference is +3 points and at the age of 8 the values are again higher by 1-3 points.

3.3 ANOVAs

3.3.1 Raw scores

The initial analysis consisted of a 4 (age) × 2 (gender) unweighted ANOVA on raw scores, separately for each scale. As expected, the main effect of age was significant for all scales (MS: $F_{3,107} = 100.22$; WS: $F_{3,107} = 72.82$; SS: $F_{3,107} = 92.74$, all $p<.001$) with substantial gain in raw score points from the age of 4 to the age of 8. Within each scale, girls as a group scored approximately 2 points higher than boys, however the main effect of gender was not significant for the MS ($F_{1,107} = 2.90$, $p = .09$) but it was significant for the WS ($F_{1,107} = 4.69$, $p<.05$) and the SS ($F_{1,107} = 6.10$, $p<.05$). The overall sex difference obtained with the SS suggests that girls obtained higher scores when drawing a woman (SS_W) compared to boys drawing a man (SS_M). This result virtually reflects the overall gender difference found for the WS but not for the MS. Age didn't interact with gender, although for all the scales gender differences were more pronounced at intermediate ages[12]. Comparing boys and girls at each age level within scale, the observed differences in favour of girls were significant at the age of 5½ for the WS and the SS (WS: $t_{24} = 2.40$; SS: $t_{24} = 2.17$, both $p<.05$) and at the age of 7 for the SS ($t_{26} = 2.10, p<.05$).

Table 3.2 Means and standard deviations of raw scores on GH Scales by age and gender.

			Man Scale		Woman Scale		Self Scale	
Age level	Sex	N	Mean	SD	Mean	SD	Mean	SD
4	Boys	14	6.50	3.25	7.29	2.55	7.14	2.48
	Girls	20	8.30	4.24	7.70	3.67	7.60	4.27
	Combined	34	7.40	3.75	7.49	3.11	7.37	3.38
5 ½	Boys	10	16.80	4.02	15.10	4.61	15.00	4.35
	Girls	16	20.06	4.73	20.56	6.17	19.63	5.76
	Combined	26	18.43	4.38	17.83	5.39	17.31	5.06
7	Boys	8	21.50	5.32	24.25	5.73	20.88	6.29
	Girls	20	23.80	5.64	27.45	8.36	26.60	6.60
	Combined	28	22.65	5.48	25.85	7.05	23.74	6.46
8	Boys	10	29.80	4.29	28.80	6.21	29.40	5.54
	Girls	17	29.12	6.76	30.18	8.00	28.71	5.17
	Combined	27	29.46	5.25	29.49	7.12	29.05	5.36
Total	Boys	42	18.65	4.22	18.86	4.78	18.10	4.67
	Girls	73	20.32	5.34	21.47	6.55	20.63	5.45
	Combined	115	19.49	4.78	20.17	5.67	19.37	5.06

Note *The combined means reported are unweighted.*

Newman-Keuls post-hoc analysis for difference (<.05) was carried out to examine how well the scales discriminate children at different age levels. The analysis revealed that for all the scales each age group yielded significantly different scores from the others. Only for the WS were 7 and 8 year olds not significantly different from each other. As expected, the magnitude of increase in score points declined between successive age groups. For all the scales there was a substantial gain in test scores, approximately of 10 points, from the age of 4 to 5½. However, a gradual decrease in the rate of gain was obtained from 8 to 3 points for the WS and from 7 to 4 points for the SS comparing 5½ with 7 year olds, and 7 with 8 year olds respectively. On the contrary, for the MS the amount of increment in score points was smaller from the age of 5½ to 7 than it was from the age of 7 to the age of 8.

[12] Girls outperformed boys at intermediate ages by 3 to 5 raw score points, and only at the age of 8 did

Scale differences

An inspection of Table 3.2 suggests that for the total group no discernible differences were obtained in the mean raw scores for each scale. However, Harris, having observed that the WS yielded considerably higher raw scores than the MS, developed different raw-to-standard score conversion tables for man and woman drawings assigning similar standard scores when a slightly better raw score was obtained for the WS. To investigate whether scale differences were actually present, mean raw scores were analysed by a 4 (age) × 3 (scale) ANOVA with a repeated measure on the last factor, separately for each sex. This procedure was adopted to make the comparisons with SS more meaningful since a self drawing is scored with a different scale (Man/Woman) depending on children's gender (boy/girl). Only the results from the within-subjects analyses are of interest here.

The overall scale effect was not significant for boys ($F_{2,76} = 1.16, p>.05$) but it was significant for girls ($F_{2,138} = 3.84, p<.05$). Each analysis also yielded a significant interaction between age × scale (boys: $F_{2,76} = 2.35, p<.05$; girls: $F_{2,138} = 3.11, p<.01$), suggesting that besides the absence of substantial scale differences for boys, there was some variability in the pattern of scales' superiority across the age levels for both sexes. To examine where the scale differences lie, pairwise comparisons were performed within gender, with an experimentwise error rate of $p = .01$. Specifically, boys obtained better scores drawing a woman at the ages of 4 and 7, while they were better drawing a man at the ages of 5½ and 8. Nevertheless, there was merely one raw score point difference in the majority of these comparisons, and only at the age of 7 was the superiority of the WS to the MS more pronounced, albeit not significant (21.50 *vs* 24.25; $t_7 = 2.87, p = .024$). Also, the self drawings by boys obtained the lowest mean score at intermediate ages whereas they were their second best at the age of 4 and 8.

---

boys perform slightly better than girls at the MS and the SS.

Girls on the other hand performed better at the WS compared to the MS ($t_{72} = 2.51, p = .014$) and this pattern was consistent at all ages except at the age of 4 where this pattern was the reverse. They also obtained the lowest score on the SS with the exception of the 7 year old group. These scale differences were significant only at the age of 7, where the mean scores for both the adult and the self drawings were superior to that of the man drawing (WS-MS: $t_{19} = 3.89, p<.001$; SS_W-MS: $t_{19} = 3.15, p<.01$). Thus, the general results from the repeated measure analysis suggest that scale differences are present only for girls, where the superiority of WS is more noticeable at the age of 7.

3.3.2 Standard scores

Having found gender differences on the WS and the self drawings (SS_M-SS_W), as well as scale differences with girls, it was thought that repeating the analyses on standard scores would be instructive. In fact, it was expected that if Harris's different test score conversions for boys and girls within scale are correcting for gender differences, in the analysis of standard scores these differences should be neutralised.

Table 3.3 Means and standard deviations of standard scores on GH Scales by age and gender.

			Man Scale		Woman Scale		Self Scale	
Age level	Sex	N	Mean	SD	Mean	SD	Mean	SD
4	Boys	14	87	13	91	12	90	12
	Girls	20	89	25	83	22	85	15
	Combined	34	88	13	87	17	88	13
5 ½	Boys	10	103	12	97	14	98	13
	Girls	16	106	13	102	15	100	14
	Combined	26	105	12	100	15	99	14
7	Boys	8	102	14	106	14	100	16
	Girls	20	104	13	104	16	102	14
	Combined	28	103	13	105	15	101	15
8	Boys	10	109	8	103	13	108	10
	Girls	17	104	13	99	14	97	8
	Combined	27	107	11	101	13	102	9
Total	Boys	42	101	12	99	13	99	13
	Girls	73	101	16	97	17	96	13
	Combined	115	101	14	98	15	98	13

Note The combined means reported are unweighted.

Again, a 4(age) × 2(gender) ANOVA was performed for each scale yielding as expected significant main effects for age (MS: $F_{3,107} = 9.32$; WS: $F_{3,107} = 7.51$; SS: $F_{3,107} = 8.02$, all $p<.001$) but not significant main effects for gender and no significant interaction between age × gender with any scale. Looking at Table 3.3, all age groups yielded mean standard scores close to the expected deviation IQ score (100), and only the performance of 4 year olds was considerably depressed[13].

Turning to the study of scale differences by gender, a 4(age) × 3(scale) ANOVA with a repeated measures on the last factor revealed again a significant main effect of scale for girls ($F_{2,138} = 8.41$, $p<.001$) but not for boys. The pairwise comparisons with experimenterwise error rate of $p<.01$ showed that, contrary to the equivalent analysis on raw scores, girls obtained an overall significantly lower mean standard score drawing a woman or a girl than drawing a man (MS-WS: $t_{72} = 3.85$; MS-SS_w: $t_{72} = 3.51$, both $p<.001$). Also the interaction of age × scale was significant for boys ($F_{6,76} = 2.56$, $p<.05$) but not significant for girls ($F_{6,138} = 1.19$, $p>.05$). For boys, the pattern of scale differences across the age levels was mainly the same to the one obtained from the analysis on raw scores[14]. For girls, the superiority of the MS to the WS was significant for the two younger age groups (4 year olds, $t_{19} = 3.64$; 5½ year olds, $t_{15} = 2.90$, both $p<.01$) whereas the superiority of MS to SS_w was significant at the age of 5½ and 8 ($t_{15} = 4.03$; $t_{16} = 5.00$, respectively, both $p<.001$).

3.3.3 Raw scores on reduced item sets

It has been suggested that the usually observed gender differences in performance on GHDT may either reflect real differences in ability or might be an artifact of the scale's construction since there is some variability on certain types and on the total number of scorable items

[13] This was reflected at the post-hoc analyses (Newman-Keuls, $p<.05$) which showed that only the 4 year olds performed significantly lower from all other age groups.

between the MS and the WS (Scott, 1981). Thus, investigating gender as well as scale differences on a reduced set of items common to both scales (CM) would permit assessment of whether the scale construction contributes to the differential performance of boys and girls. Further, it has been also suggested that girls are more attentive to detail and figural elaboration (Ford *et al.*, 1974; Harris, 1963; Ortega *et al.*, 1988; Scott, 1981; Sinha, 1970) which might account for their superior performance, particularly with female drawings. To elucidate these issues the two developmental dimensions from Strommen's (1987) factor analytic study on the GHDT were thought to be appropriate. As has been already reported in the present chapter (see section 3.1.3), Strommen, analysing the structure of the item set common to both scales, identified two developmental dimensions; one consisting of the items assessing the presence of core HF features and another incorporating "elaborate" items which tap the degree of figural refinement and the depiction of complex spatial interrelationships (see Appendix A, Table S-D). Thus, obtaining a raw score on the item set loading on the core dimension (C) and another score on the item set loading on the elaborate dimension (E), evaluation of whether gender differences reflect true differential abilities can be explored.

The preliminary univariate 4 (age) × 2 (gender) ANOVA on each set (CM, C, E) separately for each scale yielded the following results. The main effects for age were significant for all the sets within scale[15], whereas gender failed to have a significant main effect and didn't interact with age. Post-hoc analysis ($p<.05$) revealed that for the set of elaborate and common items, all the groups were significantly different to each other, with substantial increments on raw scores with age. However, for the core items only the 4 year olds obtained significantly lower scores than any of the older age groups.

[14] Again, the MS yielded better scores to the other scales at the ages of 5½ and 8, whereas the WS was superior to the other scales at the ages of 4 and 7.

To investigate whether scale differences exist with the reduced sets of items a 4 (age) × 3 (scale) ANOVA with a repeated measure on the last factor was performed on each set (CM, C, E) separately for boys and girls. The only significant results from the within subjects analyses were obtained from the girls sample, with main effects of scale for the common and elaborate set (CM: $F_{2,138} = 4.14, p<.05$; E: $F_{2,138} = 10.91, p<.001$) and an interaction of scale × age only with the elaborate set ($F_{6,138} = 2.35, p<.05$). The scale differential however was in favour of MS as girls scored more common and elaborate items drawing a man than making a same sex drawing. Further, pairwise comparisons showed that the superior performance on MS was more pronounced when girls were making a self drawing (CM: $t_{72} = 2.87, p<.01$; E: $t_{72} = 4.51, p<.001$). The scale × age interaction for the elaborate set appears to be due to the fact that at the two intermediate age levels girls drew more elaborate features for MS than for WS whereas the performance of 4 and 8 years olds was comparable across these two scales.

3.3.4 Summary and discussion

From the descriptive analysis of children's performance it appears that the present sample as a group yielded mean standard scores comparable to the test mean (MS:101, WS & SS:98). Scott (1981) however, reviewing the studies on the GHDT, suggested that there is an upward bias in Harris's standardisation norms as most of the studies yielded scores approximately one standard deviation below the test mean. As it appears from the review of recent findings in Chapter II (see Section 2.1.1) this issue is unsettled; some either supported this assertion (Mehryar *et al.*, 1987; Phillips *et al.*, 1973) or refuted it (Oakland & Dowling, 1987; Sutter & Bishop 1986), whereas others demonstrated a moderate trend to this direction (Carvajal *et al.*, 1987; Fabry & Bertinetti 1990). Obviously, reporting a mean score for the total sample is not very instructive, and factors such as age range and groups considered, as well as sampling procedures can account for these inconsistencies.

[15] CM set: MS-$F_{3,107} = 64.7$, WS-$F_{3,107} = 66.11$, SS-$F_{3,107} = 70.94$; C set: MS-$F_{3,107} = 79.04$, WS-$F_{3,107} =$

Thus, as Scott (1981) remarked, Harris's standardisation sample was restricted to children attending regular school classes, whereas most of the subsequent studies reporting lower scores included non-institutionalised children. Similarly, while race was not rigorously controlled by Harris, some studies supporting the upward bias claim were carried out with ethnic minorities (Mehryar *et al.*, 1987) or with very diverse subject pools (Phillips *et al.*, 1973). Therefore, it is nor surprising that the present scores were higher than those of Harris, since the sample had uneven distribution of the sexes and was drawn from a population of normal children, where some lower social classes and ethnic minorities are likely to be underrepresented.

The results at each age level revealed that, similarly with Sinha's study (1970), the mean raw scores for children of 7 and 8 years of age were higher than the corresponding norms, resulting in unanimously standard scores of 1 to 7 points above the test mean for all the scales. On the contrary, the reverse picture was obtained for the 4 year olds, who obtained standard scores nearly one standard deviation below the test mean. At least, in the context of these findings, the need to revise the published norms is upheld only for the pre-school children. In fact, Harris (1963, p. 294) drew attention to the possibility of inflated norms for children younger than 5 years of age and recommended to be used as "tentative guides". The confirmation of his qualification by the present findings further underscore the necessity for the proper standardisation of the GHDT for pre-school children.

The investigation of gender differences with raw scores revealed that the performance of the sexes is comparable when drawing a man, but girls are superior to boys with female drawings. Apart from few studies which didn't find any significant gender differences (Strommen & Smith, 1987), or showed girls superiority either with the MS (Mehryar *et al.*,

---

58.39, SS-$F_{3,107}$ = 71.39; E Set: MS-$F_{3,107}$ = 98.13, WS-$F_{3,107}$ = 84.99, SS-$F_{3,107}$ = 98.33, all $p<.001$.

1987), or only when MS/WS are combined (O'Keefe *et al.*, 1971), the present results echo the prevalent view that noticeable differences are not present with the MS (Coyle *et al.*, 1977; Laosa *et al.*, 1973; O'Keefe *et al.*, 1971; Phillips *et al.*, 1973; Ward & Eliot 1987; for a review see also Scott, 1981), but girls outperform boys with the WS (Harris & Pinder 1974; Harris *et al.*, 1970; Ward & Eliot 1987). These findings give only partial support for the development of separate raw-to-standard score conversion tables for the sexes for both the MS and the WS. They also justify Scott's (1981) conclusion that Harris's standardisation method is liable to underestimate girls' abilities, particularly when assessed with the MS.

Yet, the absence of a gender bias from the analysis of standard scores proves that Harris, by suppressing the test score conversions for girls succeeded in balancing the performance of the sexes within each scale. However, another problematic aspect still persists concerning the obtained scale differential only in the performance of girls. In particular, depressing the standard scores for girls when they draw a female figure seems to substitute standard score scale differences in favour of MS for raw score scale differences in favour of WS and SS_w; what actually changes is only the direction of this difference. In view of these results, serious consideration should be given administering only one scale as a short version of the GHDT when assessing girls' conceptual abilities. If such a need appears, it seems that girls' abilities might be underestimated or overestimated depending on the sex of the figure requested to be drawn, the score that will be used (standard/raw) and possibly the age of the assessed girl (significant scale × age interaction with raw scores). Thus, it seems advisable to use both MS and WS or even the average of these two when testing girls. On the other hand for research purposes which require quick testing procedures and control over gender differences, it is reasonably reliable to administer just one scale and use standard scores.

The analysis of children's performance on raw scores for the reduced sets of items suggests the following. First, the absence of a gender bias on the item set common to both scales indicates that the obtained differences for the complete WS are likely to be an artifact of scale construction. Possibly, girls score more items peculiar to the WS than boys. The absence of gender differences on the core and elaborate dimension of the scales discounts the possibility that, at least for the common items between the scales, boys and girls demonstrate differential drawing abilities which can account for the gender differences found on the complete set of scales. Considering the comparable performance of the sexes for any set of items, it seems likely that the items specific to each scale are not tapping operatively equivalent abilities. Factor analysis suggests that boys are more attentive to proportion and inter-item spatial integration whereas girls are better with details (see Chapter II/Section 2.1.3). Since in Strommen's elaborate set there is a fair mixture of items tapping both skills, no gender difference was found which might exist in the item set peculiar to each scale, if the aforementioned abilities are unevenly represented.

Finding that scale differences are present for the elaborate and common set of items – of whose the former is a subset– is in line with the studies of Van Sommer (1984) on children's repetitive drawings. According to his results, salient (core) features are invariably present across consecutive drawings of one topic whereas elaborate features are less consistent. Two factors can be held responsible for finding that only girls showed differential performance across the scales in favour of the MS. Since the SS is always administered last, children are more likely to produce a less elaborate drawing towards the end of testing due to lack of motivation. This effect of test order should be pronounced for girls, particularly in respect with the self drawing, as they are making a same sex drawing in a successive order (M-W-S_W) compared to boys (M-W-S_M). In fact, finding that the superior performance of girls drawing a man is noticeable only compared to a self drawing corroborates the previous proposition. It is also possible that girls might try to differentiate their self drawings on more affective aspects or on features not common in both scales. Results from previous findings suggesting that the drawings of younger children are more likely to show stereotypy and those of older conventionality can also explain the age × scale interaction caused by girls' variable performance at intermediate age levels.

In view of the results, there is some indication that the gender differences on raw scores for the complete scales may be the artifact of scale construction. Further, the fact that scale differences were consistently found among girls across different types of analysis (reduced sets: elaborate, common, complete sets) further cautions against the use of the scales interchangeably. Since the direction of these differences was rather variable a number of factors can be held responsible; test order, scale construction, age of the artist. Thus, no firm conclusion can be drawn and more research is needed to determine the origin of the aforementioned differences. Analysing the type of items specific to each scale and counterbalancing the order of scale administration within sex seem to be worthy of exploration in future developmental research.

Finally, it is worth saying that the different pattern of developmental change for the core and the elaborate set corroborates fully Strommen' s finding (1987) that the GHDT is made of items assessing abilities of different developmental periods. Children of both sexes showed continuous developmental change on the items concerned with structural or content characteristics (elaborate), whereas the portrayal of core feature levels off by the age of 5½. Also, consistent with Strommen's observation, the two dimensions do not evolve with the same pace. In Figure 3.1, children's mean scores on each dimension have been plotted as a function of their overall mean score on the complete MS and WS. As the Figure shows, after an overall raw score of 28 on the complete set, the gain of points on the elaborate dimension relatively to the core dimension is precipitated. However the rate of inclusion of elaborate items is steeper for the MS than it is for the WS.

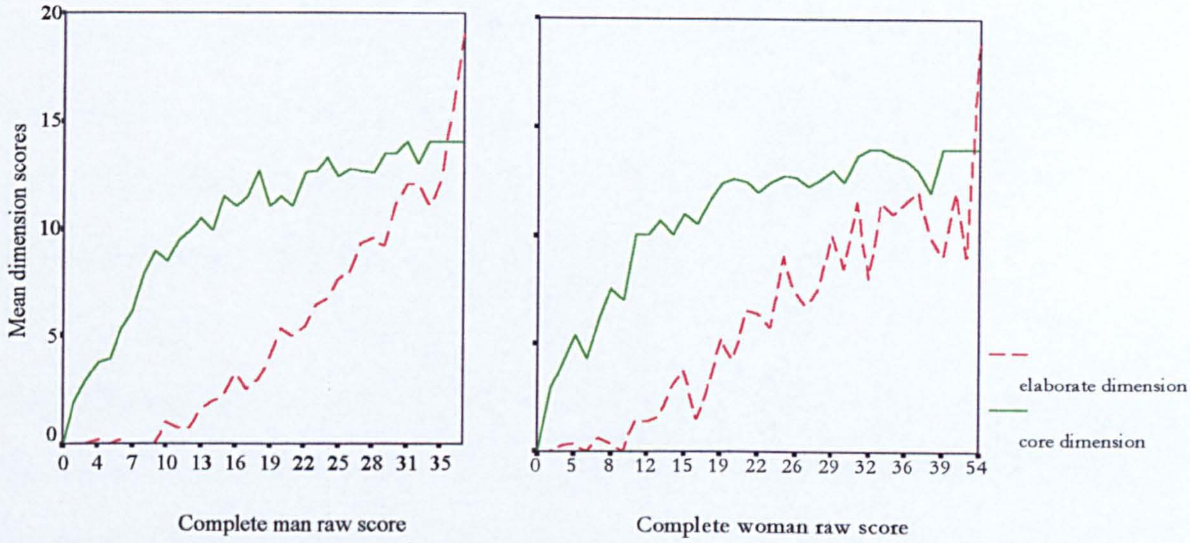


Figure 3.1 Mean dimension scores plotted as a function of the total GH raw scores

These findings might have certain implications for the development of an abbreviated version of the scales particularly for testing pre-school children (quick screening procedures). As previous research suggests, reducing the scales to a set of items assessing the presence of salient HF features does not hinder test validity and has diagnostic potential at the start of kindergarten (Goldman & Velasco, 1980; Simner, 1985). Having examined the overall performance of the GH scales and following the analysis on the reduced items sets, the properties of individual items will be investigated. The purpose of this study was to identify those items that reliably discriminate children of different abilities and characteristics.

3.4 Item analysis

Although the GHDT yields total score estimates of cognitive development, normative data for individual features of HFDs is rather sparse. Since the data from GH scales involved frequency counts, the number of children who scored on a particular item within each age group was calculated for each scale separately (MS-WS-SS). To be able to compare the results from this descriptive analysis with previous studies, the frequency counts were converted to percentages which are presented in the Appendix A from Table M_1 to S_{W4}. Data from previous research on features of HFD with maturational significance has been mainly gathered from drawings of a man or a person (Brown, 1990; Cox, 1993; Groves & Fried, 1991; Koppitz, 1968; Lark-Horovitz *et al.*, 1939, cited in Lark-Horovitz *et al.*, 1973; Ortega *et al.*, 1988; Simner, 1985; Snyder & Gaston, 1970). Thus, the main findings summarised here will be confined to the first two drawings; of a man and a woman.

The following analysis, apart from providing age-specific base rates (frequency of occurrence) for items specific to the GH scales, was undertaken in view of Koppitz's (1968) normative data. Koppitz, testing 1856 children between 5 and 12 years of age, identified 30 HF developmental items and classified them into four age-specific frequency categories, from *expected* and *common* to *not unusual* and *exceptional*[16] (see Appendix A, Table KDI). Since most of these items had derived from the GHDT, the degree of correspondence between the present data with that of Koppitz can be investigated in terms of the frequency of feature inclusion for the items common to both Koppitz and GHDT set[17]. Koppitz provides separate norms for the sexes at each age level. However, due to the relative small size of the present sample for this type of analysis, Koppitz's data were combined for the sexes, and an average rate of appearance was calculated per item against which the present data was compared.

[16] Koppitz required *expected* items to be present in at least 86% of the drawings, *common* items in 85-51%, *not unusual* in 50-16% and *exceptional* to appear in fewer than 15% of the drawings in whole-year intervals.
[17] Out of Koppitz's 30 developmental items, 24 from GH-MS and 23 from GH-WS are directly comparable.

3.4.1 Age differences

4-year-olds

Nearly two thirds of the total set of items in both GH scales (MS & WS) were never scored by the four year olds. From the 24 features which appeared at any time in the drawings of a man, 11 were rather rare (scored by <10% of the sample), 6 were present in one third, and 2 in half of the drawings, whereas only 5 were drawn consistently (approximately by ≥ 2/3 of the nursery children). Similarly, from the 25 item that were ever scored for the WS, 13 were scarce (<10%), 4 were present in one third, and 3 in half of the drawings, while again only 5 were consistently drawn (≥ 2/3 of the sample). The most consistently drawn body parts were the head and the legs, and the most common facial features were the eyes and mouth, while the fifth next common item for the drawing of a man was the arms and for the woman the nose. From the Appendix A, the data in Tables M_1 and W_1 suggests that nearly all the items which load on Strommen's core dimension of the scales were present (MS & WS: 12/14). However only 9 features from the MS and 7 from the WS scored from the elaborate dimension, all of which appear in less than 10% of the sample.

Although Koppitz does not provide normative data for the 4 year olds, the present results can be compared with Cox's (1993, p. 50-Table 3.1) and Groves and Fried's data (1991) who report percentages of feature inclusion, testing 74 and 108 four year old children respectively. The findings from the present sample are in agreement with these studies. Cox in particular also found the same five items to be the most frequently drawn. However, surprisingly enough, the outline of the head was drawn by fewer children in the present sample compared with the same age children in the aforementioned studies, and this difference was more pronounced in the drawings of the man[18]. Two inconsistencies were also observed between the studies with respect to hair and nose. Although the number of children drawing hair is comparable with that of Cox's data (45% *vs* MS: 38%, WS: 50%), it is much higher in the data of Groves and Fried (68%), especially for man drawings. Exactly the same pattern of results was obtained for the depiction of nose[19]. Possibly the inclusion of these two features depends on the sex of the HF requested, and is subsequently related to the slight difference observed for the presence of head outline in the drawings of man and woman. In particular, the contour of the head appeared more often in woman drawings along with these head features (nose, hair). Of course one would need to observe the sequence of feature production to be able to decide upon the nature of this relation[20].

Striking differences were obtained with Groves and Fried's findings relating to the depiction of clothing, and with Cox's results for the presence of body, whereas the occurrence of hands was markedly depressed in the present sample (9%) compared to both studies. While only two or three children included any item of clothing, all the children in Grove and Fried's study refined their drawings in this respect. This finding is rather puzzling and it might have resulted from differences in scoring criteria. Further, the sixth most commonly drawn feature in the present sample was the body which, consistent with Groves and Fried's findings, was drawn by nearly half of the 4 year olds. However, the depiction of the body was rare (5%) in Cox's sample. Considering the variability with which the body appears among the transitional drawers of this age under different contexts, such a big discrepancy might have resulted from instructions that didn't adequately prompt children to make their best drawings (see Goodnow *et al.*, 1986). Finally hands were an *exceptional* item here (1%) contrary to the other two studies where their occurrence was *not uncommon* (Cox: 20%, Groves & Fried, 26%).

---

[18] It is not uncommon in the drawings of young children that some features are not spatially integrated into a larger semantic unit.
[19] Cox data: 57%, present data: MS 47%, WS 62%, Groves & Fried: 77%
[20] Intending to draw more facial features could have made children to demarcate the head or alternatively having drawn facial features children might have seen this need later.

In summary, applying Koppitz's criteria for frequency categories, the inclusion of eyes is at an *expected* frequency (>85%) in the drawings of the 4 year olds along with the head for woman drawings, whereas for man drawings head is a *common* feature (85%-51%). *Common items* are also mouth, legs, body, whereas *not unusual* features (50%-16%) are pupils, hair, fingers and feet. Finally, the appearance of neck, brows, 2D mouth, and five fingers, well proportioned, is rather *exceptional* (<15%) in the present sample. The occurrence of arms and nose seem to be dependent on the sex of the HFD. Arms are *common* items among man drawing but fall into the status of the *not unusual* category for the woman drawings, while the reverse holds true for the nose.

<u>5-and-a-half year olds</u>

By the age of 5½ a dramatic increase took place in the appearance of GH scoring features in children's drawings. Specifically, 18 novel items were scored in all the drawings of a man, and 29 in all the drawings of a woman. Among these new items, slightly more were related to additional body parts for woman drawings than for man drawings (MS: 5/18, 28% *vs* WS: 11/29, 38%). On the contrary, relatively more of these new items were related to the realistic shape and the formal aspects of the HFDs of man than of woman (MS: 13/18, 72% *vs* MS: 18/29, 62%). Overall, however, if we apply Strommen's (1987) two developmental dimensions where the test items common to both scales load, these additions reflect the inclusion of items mainly from the elaborate dimension of the scales. In fact for the MS, 5½ year olds included twice as many items from the elaborate dimension compared to the 4 year olds (5½: 18/23 *vs* 4: 9/23), and for the WS three times more items (5½: 20/23 *vs* 4: 7/23)[21]. A series of chi-square tests were also performed to identify which items from the complete list of GH scales differentiated children from the two youngest age groups. Tables A_M and A_W from the Appendix A present the results for the MS and WS in respect. Five-and-a-half year

[21] The elaborate dimension contains 23 items in total.

olds passed significantly more items than the 4 year olds; 22 items from the MS[22] and 21 items from the WS, 7 of which were novel items for both scales. However, from the entire number of items which showed reliable developmental changes in the two scales, half were concerned with content characteristics (presence of body parts) and half with structural aspects (proportions, placement).

Besides the presence of significant differences in the rate at which the items appear, all the common items in the drawings of the 4 year olds increased in frequency, reaching the status of *expected* at the age of 5½. Consistently with Koppitz's observations (1968), 5½ year old children can be expected to include the following basic features: head, eyes, mouth, body, arms and some indication of clothing. There was some variation with respect to legs and feet depending on the sex of the HF. As in Koppitz's normative data, legs were expected to appear in the drawings of man, but here they were substituted by feet in those of woman. Probably idiosyncrasies in the scoring methods could account for this inconsistency[23].

Similarly with the results from the 4 year old sample, a lower percentage of 5½ year olds included a nose in their drawings which thus achieved the status of a *common* item in the present sample, but was an *expected* feature in Koppitz's sample. However, the figures for nose inclusion are comparable with other studies (Cox, 1993, p. 50; Brown, 1990), and although there is an inconsistency in the frequency category to which this facial characteristic belongs to with respect to Koppitz's data, the actual difference in percentages is moderate (6%-10%). The frequency at which hair appears seems to depend on the sex of the HF. In the drawings of woman it is an *expected* item whereas in those of man it is *common*. This reflects Koppitz's finding that when children were requested to draw a person and presumably

[22] Marginal differences were also found for 2 items (fingers/proportion and legs)
[23] What differentiates feet from legs are the proportions. Probably in the present sample short and elliptical contours might have been credited for feet instead of legs.

produce one of their own sex, hair was a *common* feature in the HFDs of boys but an *expected* one in those of girls.

Relatively more children drew hands and pupils in the present sample than in Koppitz's study, where they were classified as *not unusual* features. Although for woman drawings they still belong to the same frequency category, for man drawings they are *common* features. Further, as Cox (1993) had found, the delineation of neck was *exceptional*, contrary to Koppitz's results where it was ranked as a *not unusual* feature. Finally, relatively more children drew well-proportioned feet whereas fewer children included all five fingers, particularly in woman drawings, compared to Koppitz's data.

In general, the results are in agreement with Koppitz's finding on the developmental items present in the drawings of 5½ year olds. The features that are *expected* to be present in their HFDs are head, eyes, mouth, body, arms, an item of clothing and legs, especially when drawing a man, while feet were also drawn from the majority of children in the present sample. *Common* features are fingers and nose, which Koppitz found at the rate of *expected* frequency, while hair is *expected* in the drawings of woman and *common* in those of man. Also the representation of hands and pupils seems to be dependent on the sex of the HF. If it is man these items are rather *common* whereas in the case of a woman they drop into the *not-unusual* frequency category. Similarly with Koppitz's results, eyebrows, ears and arms-at-side are *not unusual* whereas nostrils and 2D lips are *exceptional.*

<u>7 year olds</u>

At the age of 7, the number of GH items that score at any time in the present sample moderately increased; 11 items appeared for the first time for the MS and 8 for the WS. Contrary to the previous age group, these additions involved both content and structural aspects of HFDs. Thus, from the total number of new entries, the items relating to shape and form were 6 for the MS and 5 for WS, whereas those involving embellishment were 5 and 3 respectively, and were mainly related to clothing (see Tables M_3 and W_3 in Appendix A).

When comparisons (chi-square test) were made between 5½ and 7 year olds on the total number of GH items, the following results were obtained. Contrary to the findings from the similar analysis with younger children, there was an overall smoother increase in the frequency of occurrence of GH items and notably fewer features yielded reliable age differences (MS: 4 items, WS: 7 items). The majority these features were related to the formal aspects of the drawings (see column 2 in Tables A_M and A_W, Appendix A).

Comparing the performance of the present sample with Koppitz's data, there was a striking correspondence in the type of items common to both sets (Koppitz and GH) that were found at an *expected* frequency in the drawings of 7 year olds. The data in Tables M_3 and W_3 (Appendix A) show that for both scales 9 items are *expected* to be scored consistently by the 7 year olds including head, eyes, mouth, nose, body, arms, legs, feet and at least one item of clothing. Again, as in the case of 5½ year olds, and consistently with Koppitz's results, the frequency at which hair scores depends on the sex of the HF; it is an *expected* feature in woman drawings and a *common* feature in those of man.

Fingers were *common*, as in Koppitz's sample, and hands and arms-at-side seemed to fall into the same frequency category. However, some marginal decrease in their rate of appearance in the drawings of man rendered them as *not unusual.* Considerable agreement was observed in the incidence of ears, eyebrows and correct number of fingers which were classified as *not unusual* features in both studies. Finally, the depiction of 2D lips was *exceptional*, yet a modest increase in its frequency in woman drawings assigned it into the *non-unusual* category in the present sample.

However, there were some inconsistencies relating to three items. The rate at which feet-in-proportion appeared in the present sample rendered it as a *common* feature, while it was classified as *not unusual* in Koppitz's data. The opposite applies for neck which was also found considerably more often in other studies compared to the present one (Lark-Horovitz *et al.*, 1939; Brown, 1990). Further, similarly with Brown's results (1990), pupils scored considerably more often here than in Koppitz's sample, and the difference was more pronounced in the drawings of man[24]. This pattern of difference for all three items was also present in the previous age group (feet-in-proportion, neck, pupil). However, the general picture is that the performance of the 7 year olds was comparable with the performance of the same-aged children in previous studies with respect to the most characteristic human features.

8 year olds

The results revealed a negligible increase in the number of novel GH items in the HFDs of 8 year olds compared to 7 year olds. For the MS, 56 items appeared in all the drawings of the sample as opposed to 52 items in those of 7 year olds, while for the WS there were 62 as opposed to 61. These additions were all related to formal aspects. Further, age differences in the number of children passing any given GH item compared to the previous age group were significant for 8 items in both scales, three concerned with content and 5 related to formal aspects of the drawings (see column 3 in Tables A_M and A_W, Appendix A).

From the items comparable to both Koppitz and GH set, the results were equivalent for all the human features that were anticipated at an *expected* frequency according to Koppitz (head, eyes, nose, mouth, hair, arms, legs, feet, body, some indication of clothing). However,

[24] Relative more children drew lashes/eyebrows and less drew pupils for woman drawings compared to man drawings. Possibly since both features compete for the same space, the appearance of one

it seems that the present group was more advanced than Koppitz's sample. With the exception of one item (eyebrows, also profile for the MS), when there was a difference in the frequency category to which a given feature was found to belong, it was due to higher base rates in this sample compared to Koppitz's. Specifically, fingers appeared at an *expected* frequency here whereas they were found at a *common* rate by Koppitz, and the same applies for hands, but only for the MS. Also, well proportioned feet and correct-number-of-fingers were found at a *not unusual* rate by Koppitz, but they were at a *common* rate here. Some inconsistencies were also observed for facial features which might be related to the sex of the HF drawn, or to underlying gender differences. Koppitz found that eyebrows, pupils and 2D lips scored at a lower frequency category by 8 year old boys drawing a person than by the same-aged girls[25]. Possibly, the fact that for the present purposes, Koppitz's data for each sex was combined could have accounted for the obtained inconsistencies in respect to these facial features.

3.4.2 Gender differences

Next, early claims that boys are superior than girls on items relating to the formal aspects of the HFDs, whereas girls show a greater concern for detail, were investigated. A series of chi-square tests were performed comparing the frequency at which boys and girls scored each item. The items where significant gender differences were found are presented in Table 3.4. Gender differences were not pronounced and were found slightly more often in woman drawings than in man drawings in the total sample and at each age level. From the total sample of drawings of both scales, the items relating to clothing (M_{55}, W_{43}, W_{52})[26] and the depiction of the pupil (M_6, W_6) were scored more often by girls than by boys. Also girls were

suppresses the frequency of the other, causing the bigger difference in the rate of occurrence for pupils only in man drawings.

[25] Eyebrows and pupils were *not unusual* features and 2D lips an *exceptional* one in the drawings of boys, whereas the former items were *common*n and the later item *not unusual* in the drawings of girls.

slightly better at attaching the limbs to the trunk when drawing a man (M_{45}), and had better motor control of the lines when drawing a woman (W_{64}) compared to boys. Comparisons within age by scale revealed gender difference in favour of boys only at the age of 5½ for the representation of chin and forehead in man drawings, and at the age of 8 for the representation of shoulders in woman drawings.

Table 3.4 Gender differences on GH items by scale and age level.

	Man Scale			Woman Scale		
Age	Item		*p*	Item		*p*
4				6: Eye (detail pupil)		.01
5½	14: Chin/forehead*		<.05	41: Attachment II		<.05
				43: Clothing		<.05
				52: Garb (feminine)		<.05
7	18: Hair I		<.05	2: Neck		.06
	55: Clothing I		.06	52: Garb (feminine)		.07
8				25: Shoulders*		<.05
Total	6: Eye detail (pupil)	$\chi^2 = 3.08$	.08	6: Eye detail (pupil)	$\chi^2 = 5.32$	<.05
	45: Attachment II	$\chi^2 = 3.46$	.06	43: Clothing	$\chi^2 = 3.35$	.06
	55: Clothing	$\chi^2 = 4.19$	<.05	52: Garb (feminine)	$\chi^2 = 4.93$	<.05
				64: Lines (quality)	$\chi^2 = 5.03$	<.05

Note *When N>40 for the cells on total, χ^2 tests with continuity correction were used for comparisons between the sexes. For gender differences within age level, Fisher exact test was used as 20≥n≤40.*
Items with an asterisk indicate that the difference was in favour of boys.

These results do not corroborate early claims that girls' HFDs are more detailed than boys (Ford *et al.*, 1974; Goodenough, 1926; Harris, 1963; Knopf & Richards, 1952; Koppitz, 1968). Although the observed differences were virtually in favour of girls at all ages, these differences were only on a limited number of items. Later research attributed gender differences on the basis that sexually differentiating features (head items & clothing) were used earlier by girls than boys (Brown, 1990; Papadakis-Michaelides, 1989; Willsdon, 1977)

[26] The letter denotes the scale and the number the location of the item in the scale, i.e. M_{55} = Man scale, item 55.

and that female HFDs are more sexually differentiated than male HFDs (Papadakis-Michaelides, 1989). The present findings also refute this assertion. There wasn't any gender-specific feature which appear reliably more frequently among girls than boys at all ages, and girls exceeded boys only in the representation of feminine clothing at intermediate ages.

3.4.3 Summary and discussion

Over the whole age range studied here, there were a number of items from GH scales which were never scored. Specifically, children of 4 to 8 years of age never passed 16 items from the MS and 8 items from the WS, five of which were common to both scales [27]. The preceding item analysis reveals important developmental aspects in the way HFDs change which goes beyond a simple accretion of features. At the age of 4, all the core characteristics of the human form appear variably in the total set of drawings, but only the barest minimum of features are expected to be consistently found. A dramatic increase, both in the appearance of novel items as well as in the frequency of item occurrence, takes place at the age of 5½. A growing concern for realism renders the additional items to pertain more to the elaborate and structural aspects of the HFDs than to the addition of essential body parts. Further, attention to proportion and placement is evident both for smaller (facial features) and larger (limbs, body) structural units, yet it is more pronounced with the latter. As a result of all these, a considerable number of GH items (MS: 21, WS: 24), assessing both the content and form of the composition, are available to discriminate this age group from the younger one reliably.

New features from the GH item set continue to be added to the list of elements present in the HFDs of the 7 year olds; yet these additions are fewer. A sudden preoccupation with clothing appears, as more items on outfit score here for the first time. However, the entry of new items relating to formal aspects – although it exemplifies children's continuous concern for realistic proportions and placement – is focused on smaller structural elements (ear, nose, jaw, breast line) compared to the former age group. Further, a considerably smaller number of GH items (MS: 4, WS: 7) were found to reliably differentiate children of 5½ and 7 year of age. Mainly these items are concerned with the structural aspects of the HFDs and had discriminative power for 4 and 5½ year olds also.

Finally, by 8 years of age the appearance of new items in the total set of drawings is negligible. Likewise, these novel elements add more structural balance to the human form rather than introduce fundamental parts. However, there was a considerable increase in the frequency of appearance of a large number of items present in the drawings of the 7 year olds. Thus, significant age differences are obtained for a variety of items with discriminative power for 7 and 8 years olds which are mostly different from those found in the former comparisons. These items are related to the structural aspects of smaller graphic units (nose, eyes, neck, fingers, etc).

The overall findings for Koppitz developmental items, although limited to features pertinent to the GH scales, corroborate her normative data. Inspection of Tables M_1 to W_4 from Appendix A reveals the following pattern: the number of *expected* items increases with age as the number of *exceptional* features decreases. Although certain inconsistencies exist at each age level yet they are not alarming. All these cases, except of two[28], didn't involve items previously classified by Koppitz as *expected* or *exceptional* ones. As Koppitz maintained (1968, p.11-13), from a developmental point of view, the absence of an expected item or the presence of an exceptional one for a given age group carry the most important implications

---

[27] These common items were: Hair IV (M_{21}, W_{22}), elbow (M_{34}, W_{27}), superior motor co-ordination (M-W_{65}) and directed lines/form for head outline (M-W_{66}) and for facial features (M_{69}, W_{71}).
[28] Only at the age of 5½ years, should nose and clothing have been classified as *expected* items whereas here they were *common* and *not unusual* respectively.

for children's abilities. Thus, the occasional misclassifications based on items' frequency do not seem to jeopardise the accurate assessment of children's level of development.

The few inconsistencies in the frequency of appearance of certain items might have resulted from sampling differences and from the fact that the results were combined for the sexes in the present study. The latter issue might have been primary responsible for the variation in the rate of appearance of certain facial features, where subtle or clear gender differences in favour of girls had been previous found (Brown, 1990; Ortega, *et al.*, 1988). In spite the finding from the analysis per item that there weren't pronounced gender differences either in the detailing or the formal aspects of the drawings, small differences could have resulted in the obtained inconsistencies. Also, since the present sample is biased towards girls, this can explain the finding that pupils, hair, 2D lips and eyebrows were at a relative higher frequency compared to Koppitz's data, especially among the older children. Another contributing factor could have been the fact that whereas in the majority of the studies base rates per items were obtained from children making a HFD of their own sex, here each child drew both a man and a woman. As a result, the frequency of certain items that are more masculine or feminine (see, Koppitz, 1968, p. 19) is expected to vary compared to previous studies.

It is surprising that the occurrence of nose and clothing was substantially depressed among the younger children, as well as the drawing of the neck up to the age of 7. Comparable results with respect to the neck were obtained by Cox (see Table 3.1, p. 50, 1990) and with respect to clothing by Snyder and Gaston (1970) testing first grade children. Obviously, well-controlled studies with a larger and more representative sample are required to investigate the generalizability of these results. A final comment needs to be made about two items which had been classified by Koppitz as exceptional up to the age of 8. Elbow and knee did not score by any children of the present age range although they appear in the lower 15th percentile of the range in Koppitz's sample. Although Groves & Fried (1991) found even higher rates than those of Koppitz – particularly for knees– in other studies (Brown, 1990; Snyder & Gaston, 1970) their occurrence was rare and considerably lower than Koppitz's norms. Thus, the findings are inconclusive for the validity of those items as developmental features.

3.5 Validity

3.5.1 Wechsler Scales

Means and standard deviations of Wechsler subtest standard scores, as well as VIQ, PIQ and FSIQ scores are reported in Table 3.5 for the total sample. A more detailed account of children's performance on Wechsler Scales is given in the Appendix A, in Tables I_1 and I_2, where results on both raw and standard scores are presented separately for each age group. From these results, it appears that the present sample as a whole is of average intellectual ability. Both the subtest mean scores are comparable with the standardised subtest mean (μ= 10, sd = 3) and FSIQ, VIQ and PIQ scores lie approximately half a standard deviation above the test mean (μ = 100, sd = 15). Table 3.5 shows that for the whole age range, GH standard scores (GHSSs) for all the scales where lower than Wechsler FSIQ, PIQ and VIQ mean scores. This difference was smaller for the MS and more pronounced between GHSSs and PIQ. To determine whether these differences were significant a 4 (age) × 2 (sex) × 2 (type of test score) ANOVA, with a repeated measure on the last factor, was performed separately for each pair of GH and Wechsler scores.

The within-subjects analyses revealed the same pattern of significant results; the main effect of test score was significant, with all GHSSs reliably lower than IQ scores[29]. The mean score

[29] Main effect of test score: FSIQ with M-SS, W-SS, S-SS and AMW-SS were $F_{1,107}$ = 21.58, 35.47, 49.12 and 30.35 respectively, all $p<.001$. Similarly PIQ with M-SS, W-SS, S-SS and AMW-SS were $F_{1,107}$ = 31.92, 48.6, 64.35 and 42.64, all $p<.001$, whereas VIQ with M-SS was $F_{1,107}$ = 5.86, $p<.05$,

differences ranged from 7 to 10 points with FSIQ, from 4 to 7 points with VIQ and from 9 to 12 with PIQ. There was also a significant age × test score interaction[30], showing that while these differences were substantial at the age of 4, they became to decrease dramatically at the age of 5½ until the test scores start to converge from the age of 7.

Table 3.5 Validity coefficients between GH Standard Scores (GHSSs) and Wechsler scores.

			rs for GH Standard Scores (GHSSs)			
Wechsler Scales (WISC-III^UK /WPPSI)	*X*	*SD*	M-SS X= 101	W-SS X = 98	S-SS X = 98	AMW-SS X = 99
Verbal Scale						
1. Vocabulary	11	3	.20^a	.21^a	.15	.22^a
2. Information	11	3	.14	.10	.11	.13
3. Comprehension/ Similarities	11	3	.14	.06	.10	.10
Performance Scale						
1. Picture Completion	11	3	.33^c	.31^c	.30^c	.33^c
2. Block Design	11	3	.29^b	.28^b	.35^c	.29^b
3. Picture Arrangement/Mazes	12	3	.33^c	.33^c	.28^b	.34^c
Verbal score (VIQ)	105	14	.19^a	.15	.14	.18
Performance score (PIQ)	110	17	.40^c	.39^c	.40^c	.41^c
Total score (FSIQ)	108	15	.36^c	.32^c	.32^c	.35^c

Notes M-SS = Man Standard Score, W-SS = Woman Standard Score, S-SS = Self Standard Score, AMW-SS = Average Man/Woman Standard Score. a = p<.05, b = p<.01, c = p<.001, N = 115. The correlations are partial, where age (CA) is controlled.

Previous research also suggests that the comparability of the GHDT with criterion measures of intellectual abilities varies at different levels of IQ estimates (Aikman *et al.*, 1992; Gayton *et al.*, 1971; Reisman & Yamokoski, 1973; Scott, 1981). To examine this claim, subjects were stratified into four categories of IQ estimates, ranking from below (<85) to above average (>115) intelligence. The concordance in the classification category that a child was placed on the basis of his GHSS and FSIQ was examined by hit rates. A detailed account of this analysis is presented in the Appendix A, Table V_1, separately for the three GH drawings, and the main results will be summarised here. Although this type of analysis was not pre-

with W-SS was $F_{1,107} = 13.11, p<.001$, with S-SS was $F_{1,107} = 18.77, p<.001$ and with AMW-SS was $F_{1,107} = 9.79, p<.01$.

[30] Age × Test score interaction: FSIQ with M-SS, W-SS, S-SS and AMW-SS were $F_{3,107}$ = 21.22, 19.63, 19.61 and 21.67 respectively, all $p<.001$. Similarly PIQ with M-SS, W-SS, S-SS and AMW-SS were $F_{3,107}$ = 21.38, 19.66, 19.65 and 21.66, all $p<.001$, whereas VIQ with M-SS, W-SS, S-SS and AMW-SS were $F_{3,107}$ = 13.01, 12.41, 11.59 and 13.28 all $p<.001$.

planned to ensure comparable numbers of children at each level of intellectual functioning, a clear picture emerged; the classification accuracy rates were rather poor. Only 32%, 28% and 26% of the total sample were classified in the same IQ category by the criterion measure and scores from the MS, WS, and SS respectively. A consistent pattern was also revealed; the best classification rate was achieved at the IQ range of 86-100 by all the GH scales whereas children of moderately above (101-115) and above average intelligence (>115) were those who were more frequently unfavourably classified by the GHDT.

Further, a quick inspection of Tables I_1 and I_2 in Appendix A revealed an unexpected pattern. Standard scores for Wechsler scales and subtests seemed to fluctuate with age, whereas there were some cases where children's raw scores tended to remain invariable or even to reduce with age. To investigate this matter, scores from all the measures (Wechsler scales and GHDT) were correlated with children's chronological age. This analysis was considered essential before any subsequent study of the between-instruments relationship was carried out. The analysis gave vexing results. Although standard scores from the instruments are correcting for the age factor, positive correlations were obtained between GHSSs and children's age (at mid 30s, $p<.001$), while on the contrary IQ scores from the Wechsler scales correlated negatively with age (around .30, $p<.01$, 001)[31]. What appears to be the major factor for these results is the performance of the 4 year olds whose GHSSs were substantially lower compared to the older children, while the reverse pattern was obtained from their IQ scores. Therefore, it was decided to account for these finding by controlling for age when the correlational analysis between the instruments was conducted on the total sample.

[31] We will come back to this issue in Chapter 12. Full details on estimates are provided in Table 12.2.

The estimates from this analysis have been presented in Table 3.5. The magnitude of validity coefficients for the overall sample was moderate but statistically significant in the majority of cases. The pattern of correlations was also virtually identical for all four GHSSs. Generally, the drawing scores correlated with FSIQ at mid 30s (<.001), yet performance on the GHDT didn't seem to correlate with the verbal component of the Wechsler scales. A poor relationship was obtained exclusively with the Vocabulary subtest, and of equal magnitude was only the coefficient between VIQ and the drawing score from MS. In contrast, stronger correlations were obtained with the performance component of the Wechsler scale. GHSSs correlated with all the performance subtest from the Wechsler scales around .30 and with PIQ at .40 (<.001). Since a significant correlation was obtained between GHSSs and PIQ, but not between GHSSs and VIQ, it was of interest to examine whether the difference between these correlation coefficients was substantial enough to support the convergent validity of the GHDT as a nonverbal test. The significance of differences was examined using *t*-tests for dependent correlations (Howell, 1982, p. 243). The analyses showed that except of the score from the MS, all GHSSs correlated significantly more with PIQ than with VIQ (for M-SS: $t = 1.49, p > .05$; W-SS: $t = 1.68$, S-SS: $t = 1.83$, AMW-SS: $t = 1.64$, all $p < .05$).

Table V_2 in the Appendix A reports validity coefficients by age and GH scale. In parallel with the results from the total sample, performance on the MS, WS and SS didn't correlate with performance on any of the Verbal subtests and subsequently with VIQ scores for all age groups except for the 7 year olds. In fact, only 7 year olds yielded correlations between GHSSs and Verbal subtest scores conspicuously higher than the other samples. In this age group, except of Vocabulary which didn't correlate with any GH scale, the remaining correlations were significant and relatively higher for the MS and WS compared with the SS, resulting in a reliable relationship between performance on these two scales and overall VIQ scores ($p < .01$).

A more variable picture was obtained when GHSSs were compared with Wechsler scores on Performance subtests at different ages. Thus, there was no reliable relationship at the age of 8, but 7 year olds yielded high and significant correlations for all the subtests. The only exception was the lack of a significant relation between the Picture Arrangement subtest and SS. Contrary to the 7 year olds, the performance of 5½ year olds on WS and SS didn't correlate with any of the Performance subtests and only Mazes – here substituted for Picture Arrangement – correlated moderately with MS. Finally, 4 year olds yielded the same pattern of correlation coefficients for the MS and WS but the reverse for the SS. Specifically, only Mazes correlated with performance on the former scales (Man/Woman, $p<.05$) whereas Mazes was the only subtest which did not correlate with performance on the latter scale. Finally, except for the 8 year olds, GH scores from all the scales correlated well with PIQ scores. The coefficients ranged from low 40s to high 60s with a median of 46. Only the PIQ score of 5½ year olds didn't correlate with WS. Finally, the validity coefficients with the FSIQ scores were significant only for the two intermediate age groups. Thus, FSIQ scores correlated considerably with performance on the three GH scales at the age of 7, however only moderately with MS and WS at the age of 5½.

3.5.2 PEFT/CEFT

Finally, performance on the GHDT was compared with performance on PEFT with the 4 year olds and with CEFT with the remaining age groups. Pearson product-moment correlations were computed on mean raw scores since standardised scores are not available for the two forms of the EFT (see Appendix A, Table V_1). The correlation coefficients for the total sample were partial, controlling for age, and ranged from .37 for WS and SS, to .40 for MS, all statistically significant at $p<.001$. The magnitude of correlations changed when examined by age and GH scale. Thus, performance on EFT correlated with MS at the ages of 4 and 7 (.43, $p<.05$) and with WS only at the age of 7 (.41, $p<.05$).

3.5.3 Summary and discussion

The results from the evaluation of the GHDT against criterion measures can be summarised as follows. The GHDT seems to underestimate children's conceptual abilities. For children of 4 to 8 years of age, GH scores from all scales were consistently and reliably lower than Wechsler IQ scores. Scott (1981), reviewing the studies of the GHDT concurrent validity, reports that for children of average intelligence (90 to 118), criterion IQ scores were underestimated by the GHDT approximately 12 standard score points. In fact the present results fully support this estimate, as the magnitude of test scores differential between FSIQ and GHSSs ranged from 7 to 10, with a median of 10 points for the total sample.

It has been suggested that the comparability of the GHDT with conventional measures of intelligence may depend on a number of variables, namely ranges of intelligence and age. Stratification of test results by IQ levels reveals that the classification accuracy rates for all the GH scales are low and that the GHDT tend to underestimate criterion IQ scores especially in the upper ranges of intelligence (>101). This finding is comparable with previous studies (Aikman *et al.*, 1992; Gayton *et al.*, 1971; Reisman & Yamokoski, 1973). Further, the size of the discrepancy between GH scores and criterion scores varied considerably across different age groups. Mean score differences for the 4 year olds in particular were sizeable, since the IQ estimates from GHDT ranged from mid 20s to low 30s standard score points lower than Wechsler estimates. Although still present, the test scores differential showed a steep decline at the age of 5½ and grew smoother from the age of 7.

This significant age × test scores interaction can be partly attributed to the fact that different Wechsler scales were used for the two younger and the two older age groups[32] (WPPSI for 4

[32] The mixture of subtests for these two Wechsler Scales differ slightly as well as item similarity for subtests common to both scales. Possibly, the nature of the construction of the Scales might have contributed to this finding.

& 5½, WISC-III^{UK} for 7 & 8). Although correlational studies on Wechsler Scales provide unequivocal evidence of their convergent validity, assessing schoolchildren with WISC-R and WISC-III showed that the latter yielded somewhat lower IQ scores, compared to the former (see Wechsler 1992, p. 88-91). Considering this, criterion IQ scores for older children might have been closer to GHSSs, attenuating test score differences between GH scales and Wechsler scales for younger children assessed with the WPPSI.

However, the fact that mean GHSSs for the 4 year olds were 15 points below the expected standard score of 100 offers another possibility in interpreting the obtained age × test score interaction. It seems that pre-school children in particular suffer more from the presumed upward bias in Harris's standardisation norms, resulting in relatively bigger discrepancies between GH and criterion scores for this age group. Nevertheless, another element which might have contributed to the fluctuation of test scores difference across age levels is that the GHDT might tap somewhat different abilities than the conventional tests of intelligence, and possibly the nature and the strength of its relationship with criterion measures of intelligence might change at different ages.

Correlational evidence from the entire sample corroborates previous claims that the drawing test is not closely related to conventional measures of intelligence. The validity coefficients fell in the mid 30s and are somewhat lower than those reported by the majority of previous studies (mid 40s). This can be due to the fact that IQ estimates were obtained by a short form of Wechsler scales and that the present sample was rather homogeneous and of a limited age range. Yet, except for studies with diverse subject pools or special populations, where validity coefficients for the GHDT are higher (Croake *et al.*, 1973; Gayton *et al.*, 1971 & 1974; Tramill *et al.*, 1980; White, 1979), in the majority of cases the correlations were quite modest and accounted for a small percent of the common variance (approximately <30% see Abell *et al.*, 1998; Aikman *et al.*, 1992; also Oakland & Dowling, 1983).

Notwithstanding this, considering that Wechsler Scales measure a wide range of abilities and that the GHDT uses a nonverbal medium, it was expected to correlate more with the performance rather than with the verbal component of the criterion measure. In fact, the results from the overall sample showed that the GHDT yielded significant and reliably higher correlations with PIQs compared with VIQs. Partial support for the convergent validity of the GHDT was also obtained from the significant correlations between the GHDT and EFT in agreement with previous studies (Saracho 1985). Since the latter loads on the Performance subtests of the Wechsler scale (i.e. Block Design, Picture Completion, Object Assembly) and on the flexibility of closure factor (a perceptual factor) (see Anastasi, 1990), the GHDT seems to assess abilities more closely related to perceptual, motor and spatial skills. Yet two qualifications need to be made; first, correlation estimates are likely to vary when the complete set of Performance and Verbal subtests is administered. Second, finding discernible mean score differences between GHSSs and PIQs challenges the strength of this relationship and implies that the GHDT also taps abilities not measured by the conventional instruments of intelligence.

The results from previous studies examining the issue of convergent validity are also inconclusive. Some have reported either a small trend or reliable differences supporting the view that drawing tests, such as GHDT, relate more to nonverbal than to verbal intelligence (Abell *et al.*, 1998; Dunn, 1967c; Fabry & Bertinetti, 1990; Oakland & Dowling, 1983; Short-DeGraff & Holan, 1992; Short-DeGraff *et al.*, 1989[33]; Sutter & Bishop 1986; White, 1979) whereas others obtained contradictory results (Dunn, 1967b; Laosa *et al.*, 1973;

Struempfer, 1971; Strümpfer & Mienie, 1968; Wisniewski & Nagliery, 1989). Nevertheless, the few studies which report separate correlations for Wechsler subtest agree in that the stronger relation exists between Block Design and GH scores (Dunn 1967c; Laosa *et al.*, 1973; Sutter & Bishop, 1986; White, 1979). In the present study all the Performance subtests from the Wechsler scales correlated comparably with GHSSs and the only subtest which consistently showed some moderate relationship with scores for all the GH scales was the Mazes from WPPSI and Picture Arrangement from WISC-III^{UK}.

Examination of the validity coefficients by age level yields a rather variable picture supporting earlier findings (Yule *et al.*, 1967; Datta, 1967; Phillips *et al.*, 1973; Sinha, 1970) that the nature of the relation between HFD tests and criterion measures of intelligence changes as a function of age. Harris (1963) seemed to acknowledge the possibility that the GHDT may measure somewhat different functions at different ages and that the significance of the test as a measure of mental abilities might change with development. However, he was reluctant to draw any firm conclusion on the basis of insufficient evidence. Twenty years later, Scott (1981), reviewing empirical literature, asserted that the frequently maintained assumption that there is a linear relation between HFD tests and conventional measures of intellectual maturity across age is inaccurate and instead she described it as curvilinear. The results, mainly form the Man scale and Self scale lend support to this assertion, as validity coefficients for the 4 and 8 year olds remained statistically not significant – reaching only the high 20s – whereas those of the two intermediate age groups were relative higher and attained significance.

Similarly with the results from the total group, the correlation coefficients obtained separately for the Verbal and Performance Scale with the GHDT at each age indicate that the

[33] Based on the General Information Verbal subtest of WPPSI, the authors obtained a non-significant

test is more allied to performance than to verbal abilities. This is primarily true for the two younger age groups which generally yielded substantially higher and statistically significant correlations between GHSSs and PIQs than with VIQs. This pattern was consistent for all the GH scales[34]. At the age of 8 performance on the GHDT didn't correlate with any of the two Wechsler scales whereas for 7 year old children, the GHDT correlated significantly with IQ scores on both Performance and Verbal scale[35].

It is rather precarious to draw any firm conclusion as to the specific mental processes and functions that the GHDT assesses at different ages just by looking at the correlational patterns with Wechsler subtests. It has been suggested that the GHDT relates more to tests of reasoning, spatial aptitude, perceptual speed and accuracy at older ages, and more to tests of numerical aptitude with kindergarten children (Anastasi, 1990). At least for the 7 year olds, the GHDT taps abilities which relate to the development and use of both perceptual-motor concepts and abstract-verbal concepts. On the contrary, there is some indication that for the younger children there is some partial overlap in the abilities assessed by GHDT and the non-verbal component of Wechsler scale related to visual perception and motor co-ordination skill (Mazes).

4. OVERALL CONCLUSIONS

In the present chapter a number of issues relating to the psychometric properties of the GHDT were rigorously investigated. On the basis of the overall results, the following conclusions can be drawn. All estimates of reliability were comfortably high. The re-scoring of the drawings by the same rater yielded remarkably consistent scores and shows that the

---

corrrelation with GHSSs.

[34] Only at the age of 5½, the correlation between W-SS with PIQ albeit higher with VIQ failed to reach significance $p = .07$.

[35] Only the correlation coefficient between S-SS and VIQ fell marginally out of significant level (p = .06).

scoring method is eminently credible. However, one should keep in mind that these reliabilities might be somewhat inflated by the fact that a large volume of drawings was scored, promoting subsequently the development of individual standards on each occasion. Items that could have passed on the first instance might not have passed the second time, despite the high consensus between any pair of absolute scores for each child. Examination of scorer's consistency item by item reveals systematic discrepancies that are not reflected in reliability coefficients, and fully corroborates the nature of these feature-specific ambiguities previously reported (Phillips *et al.*, 1973). These results warrant the revision of the scoring method for particular items to further enhance the objectivity and the precision of scores across time and different raters. Appendix A-Scoring, where these problematic aspects are elucidated and scoring suggestion are offered can thus be helpful in this attempt.

Estimates of internal consistency were comfortably high considering test's brevity, and comparable with those reported elsewhere. The ones obtained by the split-half method are in the high 80s, whereas inter-scale correlations lie 10 points lower, in the high 70s. Considering that partially different test components comprise the GH scales, the alternative-form reliabilities compare favourably with the split-half reliabilities, and warrant the administration of a single drawing when a quicker and cruder assessment is required. However, due to the nature of the scales' construction, they seem to measure a mixture of similar and different abilities. Consequently, and in keeping with Harris' suggestion, the mean of their standard scores (Man/Woman) should be used for greater reliability.

The fact that the magnitude of intercorrelations remained unchanged when the analysis was restricted to the set of items common to the scales rules out Harris presumption that the modest intercorrelations result from items peculiar to each scale. Two possibilities might account for this result. Perhaps the same item might acquire different representational significance in each scale. Although, this is truly speculative, it is not unlikely for example that the presence of eyelashes, common to all scales, might weight differently in the drawings of a man than of a woman or self. However, another explanation might be related to the fixed order of administration of the HF drawings; possibly children might gradually lose interest and produce less elaborate drawings. The analysis of raw scores from the common and the elaborate set of items showed that girls demonstrated differential performance across the scales in favour of the first drawing (Man).

The reliable scale differences obtained from all types of analyses, clearly caution against the use of the scales interchangeably, and this is particularly the case for the assessment of girls. The direction of scale difference was variable across analyses (complete/reduced set of items; raw/standard scores), and two factors seem primarily responsible for these inconsistencies. One has to do with the fact that girls are more affected by the test order, since they produced the same-sex drawing in a consecutive order (Woman-Self). Although this does not hinder the representation of core HF features, it reduces the inclusion of elaborate elements. The other is related to Harris' standardisation method and the depression of girls' raw scores when converting them to standard ones. Thus, girls' abilities are underestimated by the female drawings relative to the male drawing when the standard score is used but the reverse picture is obtained with raw scores.

Not only is the administration of a shorter version of the test not advisable for girls but also for children of 7 years of age. This follows from the finding that the significant age × scale interaction from the analysis of both raw and standard scores was mainly attributed to the pronounced differences in the performance of the 7 year olds across the scales. Therefore two issues need to be addressed in future developmental research with the GHDT. The first is the order of administration of the man, woman and self drawing in relation to the gender of the child and his age. It is essential to know if the fixed order of testing is in any case responsible for the scale differences observed with girls, and how it might affect children of different ages (auxiliary: practice effect, or inhibitory: loss of interest).

Second, the revision of Harris' published norms for girls seems imperative on the basis of the overall results on gender differences. In perfect correspondence with previous findings, gender differences in favour of girls are only present for the female drawings, rendering the depression of their raw scores in the conversion tables for the Man scale redundant, and their assessment with this scale dubious. Consequently, it is desirable to obtain two opposite-sexes HFDs from girls for the assessment of their intellectual abilities until the norms are corrected. The nature of the gender differences with the Woman scale however is vexing. Careful item analysis on both scales does not support the previous claims of girls' superiority in detailing and sexual differentiation of the HF, amply provided by the construction of Woman scale. Factor analytic techniques, applied on both scales, might be more instructive in unravelling the abilities and the extent in which these are tapped by the different scales.

In general the GHDT performed well as it effectively discriminated children of different ages and no evidence of the previously claimed upward bias in Harris standardisation norms was obtained. The only exception to the latter claim is the performance of the 4 year olds, who were markedly underestimated on the basis of the GH standard scores. Thus, it seems advisable for practitioners and researchers to use the raw scores. Since the test appears more promising and useful in the assessment of very young children, the revision of the norms for pre-school children is of vital importance.

So far, one can claim that the test appears to be a stable and a reliable measure. It is very effective in differentiating among children of 4 to 8 years of age and instructive in pinpointing the nature of the developmental changes in their graphic productions. This clearly follows from the discriminative analysis per item on each scale. On the other hand the results for its validity are not as satisfactory. Estimates of validity are variable depending on children's age and level of intelligence. They are modest, yet significant, for children of intermediate age (5½ & 7) but not significant for the 4 and 8 year olds, in agreement with previous observations that the relationship between the GHDT and criterion measures of intelligence changes over the years. Further the present study shows that the power of the GHDT as an estimator of criterion scores is weak in the upper ranges of intelligence (>101), again in harmony with the prevalent view. Finally, there is some indication that the test relates less to verbal rather than nonverbal intelligence, but strong evidence is lacking because of the small size of these correlations in the present results.

Does this suggest that these problematic aspects are inherent to the GHDT in particular or to the nature of the HF tests in general? The answer comes from the most recently developed drawing test. To meet the need to modernize, and to objectify the scoring system, as well as to update the norms developed more than 30 years ago by Harris, Nagliery (1988) developed the Draw-A-Person Test (DAP). Although the author maintains in the Manual that the test is launched as a revision of the GHDT, Kramer and Conoley (1992), assessing its utility in the assessment of intellectual abilities, conclude that "it is unfortunate that the revision stays so close to its predecessor" (p. 289). Data-based studies investigating its psychometric properties suggest that the two HFD tests perform overall comparably (Bardos, *et al.*, 1989; Naglieri, 1988; Smith 1987) and correlation estimates with criterion measures of intelligence were of the same range. (Kamplaus & Pleiss, 1991; Naglieri, 1988; Prewett *et al.*, 1989; Wisniewski & Naglieri, 1989, see also Motta *et al.*, 1993). According to Kramer and Conoley (1992) what seems to be responsible for the fact that the DAP test yields similar reliability and validity estimates to the GHDT is that Nagliery didn't adequately address the content and the purpose of the GHDT.

The use of human figure drawings as test instruments to assess intellectual abilities has a long history, yet the issue remains highly controversial. One has only to read the nine papers published in 1993 in the *School Psychology Quarterly* to realise the acrimonious nature of this debate. Kamphaus & Pleiss (1991) maintain that the mediocre validity coefficients of the HFD tests challenge the long-held view that they are measuring the same construct as intelligence tests. They add that Behaviour scales, which are not considered estimates of ability, yield correlations of the same magnitude as drawing tests with measures of intelligence. Motta *et al.*, (1993), from a congenial theoretical vein, argue that the convergent validity of the HFD tests as measures of nonverbal intelligence is dubious and their use as screening devices is superfluous. Their rationale is that scores on performance scales as well as short forms of standard intelligence tests (Weschler) correlate much higher with overall IQ scores than do HFD tests.

On the other hand, Naglieri (1993) rightly counterattack, claiming that a short form, or a subscale of an instrument such as the Wechsler intelligence test, yields higher correlations with the FSIQ than the HFD tests simply because these correlations represent a relationship between a variable and one of its own subsets. Besides, he claimed that he hadn't develop the DAP test as a predictor of Wechsler IQs. At this point it is worth recalling that Harris regarded the revised GHDT as a measure of conceptual maturity and not of intelligence. It is also fair to say that the drawing tests measure at least one type of intellectual maturity; that which is operationalised by visual-motor tasks of the Bender Gestalt variety (Aikman *et al.*, 1992; Oakland & Dowling 1983, Short-DeGraff & Holan, 1992). And it is also reasonable to expect rapid screening measures such as the GHDT not to perform as well as comprehensive batteries which operationalize intellectual maturity differently.

Notwithstanding the difficulty in defining the nature of the construct validity of drawing tests, such as the GHDT, it is preferable to adopt the constructive approach directed at improving rather than abandoning this instrument. The fact that it reliably differentiates children of successive ages and outlines the course of children's drawing development, adds to its practical utility in the detection of certain anomalies. Thereupon, it should be used in conjunction with more extensive measures of abilities. If one also considers the attractive properties of the HFD tests– economy and ease of administration, utility with linguistically impaired or inhibited children and its less culturally loaded component – it is not surprising that Kramer & Conoley (1992) predict that "despite the lack of 'hard' data to support their use, figure drawings will continue to be a popular part of many assessment batteries" (p. 289). Future developmental research aiming to improve the test's performance should revise the norms for the pre-school children, rectify those for girls and clarify the scoring for certain items. Factor analytic studies would also be instructive in unravelling the test component and measuring their contribution to the construction of the scales. Until then, one can only speculate about the construct that the test measures.

CHAPTER IV

THE DEVELOPMENT OF VISUAL REALISM:

THE CASE OF PARTIAL OCCLUSION

1. PREFACE

The main body of research on children's drawings has been devoted to how faithfully children convey information about the visual word and has traditionally related drawing development to intellectual maturity. The drawing theories of Goodenough (1926), Harris (1963), Piaget and Inhelder (1948/1967), and Lowenfeld and Brittain (1975) describe a sequence of developmental change culminating in one clear end; that of visual realism. In a congenial theoretical vein, Luquet (1913, 1927/1977) argues that children's drawings gradually lose their symbolic nature and acquire representational status. These theories – either implicitly or explicitly – take an ideal, visually correct projection as the standard against which children's drawings are assessed.

In the previous chapters, the representational status of human figure drawings was evaluated on the basis of the number of details and the delineation of intra-figural relationships. In that context, the developmental changes were viewed as an index of intellectual maturity. The following chapters approach drawing development from a different angle. The accomplishment of realism is addressed in terms of particular graphic milestones in representing spatial relationships between simple forms. The research purpose is not only to indicate when these skills are acquired, but also to describe the process, exploring children's graphic devices. In addition, the approach adopted is not limited to what children can or cannot draw, but *why* they perform in a certain way. Consequently, the experimental manipulations of various factors will allow the evaluation of the traditional approach to drawing development which makes inferences about children's intellectual development from their visually accurate or inaccurate representations.

2. INTRODUCTION

Scholars of children's art generally agree that their drawings often describe, rather than identify a depicted object by a specific view, and that the developmental change takes the form of moving from object/array-centred (or object/array-specific) to view-centred (or view-specific) representations (Cox, 1985; Ingram & Butterworth, 1989; Light, 1985; Light & Humphreys, 1981). As a research puzzle, this has been frequently translated into the question of why children draw parts of an object or array that are not visible from their viewpoint. This phenomenon is one of the few aspects of children's drawings (like tadpole forms) that has engendered a lot of research since it was first studied by Clark[1] (1897). This chapter will review the literature on the acquisition of a single viewpoint, with particular emphasis on the pictorial representation of the spatial properties of arrays of objects involving partial occlusion.

Looking in the visual field, objects rarely exist as separate entities; usually they are partially or totally masked by being inside, behind or next to other objects. Willats (1977) contends that the spatial differentiation of the pictorial space into foreground and background by means of partial occlusion is a device that contrary to others (texture gradient, foreshortening, atmospheric perspective) can be directly derived from the perception of the optic array. The child needs just to observe the discontinuities of the projective shapes of the visible surfaces and reproduce them. This probably explains the fact that Milbrath (1998), analysing children's drawings for the spontaneous use of perspective indicators, found that few children beyond the age of 10 used other more sophisticated devices to create the illusion

[1] Clark presented children with an apple through which a hat-pin was stuck. Young children made 'transparent' drawings with the hidden part of the hat-pin visible through the apple's contour.

of realism in pictorial space. Similar views are also held by other researchers (Hagen, 1985; Thomas & Silk, 1990, p.102).

Nonetheless, even if occlusion is more accessible than other devices, empirical findings suggest that it is not pictorially mastered until around the age of 8 (Cox, 1978; Freeman *et al.*, 1977). Children below this age tend to draw the hidden features of the occluded object which is represented with a complete form, initially spatially separate and subsequently united with the occluding form. Despite good experimental evidence for this developmental shift, the nature of this progression is still a theoretical problem. In the following section, the main theoretical approaches will be reviewed to establish the working framework from which different hypotheses later derived. Next, the determinants of visual realism will be discussed, presenting the empirical findings from the studies which systematically tested these hypotheses.

3. CONCEPTUAL FACTORS: CHILDREN DRAW WHAT THEY KNOW

Central to this theoretical framework is the notion that cognitive abilities determine children's perceptual processing and representational skills. Since representation cannot precede understanding, drawings are seen as reflections of ideas about objects, and of children's conceptualisation of space. The origins of this theoretical proposition are found in the early work of Kerschensteiner (1905) and are encapsulated in his axiom – later attributed to Luquet (1927/1977) – that young children draw what they know and older children draw what they see[2]. Kerschensteiner argued that young children, while drawing, use their concepts as models instead of the visual images of the perceived world. Further, their early drawing conventions, dubbed as *schemes*, embody the content and the form of their ideas, moderated by the limitations of the graphic medium.

---

[2] No translation of Kerschensteiner and Luquet's original works exist, but for the discussion of Kerschensteiner's ideas see Eng, (1931); Goodenough (1926); Lowenfeld, (1952), and for Luquet see Piaget & Inhelder (1948/1967); Harris, (1963); Freeman (1972).

Luquet, based on these ideas, postulated that children's drawings progress through four distinct stages, culminating in visual realism. Luquet ascribed *realistic intentions* to children long before they reach the stage of visual realism, since their drawings constitute honest attempts to depict (not just symbolise) a topic. However, young children's representational logic is underpinned by their intention to make their drawing recognisable in term of objects' true identity and not in terms of its momentary appearance. Such a representational attitude is characteristic of the stage of intellectual realism, where children draw from an *internal model* (modèle interne). Luquet stressed that this model is not directly synonymous to conceptual knowledge but is rather a generic exemplar of a subject class, as it comprises its essential defining aspects[3]. Young children's inability to omit a part of an object's shape or one of its characteristic features – concealed by its orientation or its relative position to another object – has been interpreted as an instance of this conceptually based informative bias and of the state of intellectual realism which lasts approximately up to the age of 8. Eventually, shaped by the conventions of the Western culture, children's representational intentions get visually realistic. Drawings become comprehensible only when they depict appearance and not just reality, and children start to render the various objects and their relationships in spatial perspective.

Piaget and Inhelder (1948/1967) incorporated Luquet's theory on drawing development into their model of spatial intelligence. Three premises are central to their theory. First, they construe drawing to be closely related to the construction of mental images by the child. Second, they postulate that spatial intelligence plays a vital role in the mental image the child constructs of the objects or the scenes which he tries to draw. Third, they maintain that the content of internal representations of spatial relations can be directly externalised in

[3] Luquet used the term *exemplarité* referring to the nature of this model and emphasised that it is dynamic because it contains both "criterial" features as well as peripheral, the use of which is dictated by the demands of a given situation. An example which illustrates its dynamic nature is a chimney with

drawings, and used them as one source to access children's knowledge of projective geometry and level of spatial reasoning. Thus, although both Luquet and Piaget treat drawings as translations of internal states (*internal model* or *mental image*), Piaget's approach is narrower in the sense that he focused on the spatial properties of these concepts.

Piaget related the stage of intellectual realism to the ability to understand and thus represent only the topological features of space, and the stage of visual realism to the ability to conceptualise its Euclidean and projective aspects. In this light, representing a partially occluding object with a complete form reflects an understanding of the topological relation of continuity, internal to the figure. Further, drawing it as separate, attached or overlapped by its occluder shows accordingly that the topological relationships of proximity, continuity and enclosure, internal to the array, are preserved. However, due to children's feeble capacity for spatial synthesis and integration, all these topological relationships are left uncoordinated. At the stage of visual realism, children begin to conceptualise space from a viewpoint, and construct an external spatial framework that regulates the mental representation of the spatial relationships between overlapping objects.

In the light of Piagetian theory, the interpretation of intellectual realism acquired a different meaning. Although Luquet ascribed children's disposition to draw concealed features to a conceptually based intention to communicate knowledge about objects or familiar relationships, Piaget interpreted it as a cognitive deficit, originating in the shortcomings of spatial reasoning. However, both disregard external factors (production problems, contextual cueing, cultural stereotypy) which can operate in the process of the graphic rendition of the mental image or model (see, Freeman, 1976; Kosslyn *et al.*, 1977). In fact, Luquet stressed the notion of intention, and recognised that there are occasions where a representational

smoke appearing in the drawings of a house done at winter and not at the summer time (for a relevant discussion see Freeman 1972).

intention, albeit visually realistic, may be difficult to "mobilise". Similarly, Piaget acknowledged that the depiction of a spatial relationship might appear quite some time after it becomes available in the child's reasoning (for Luquet see Freeman, 1972; for Piaget see Strommen, 1988). Yet, the factors interfering between mental and pictorial representation are not explicated, and the discrepancy between the availability and the accessibility of the conceptual structures is essentially underestimated.

4. PERCEPTUAL FACTORS: CHILDREN DRAW WHAT THEY SEE

Despite the differences among the perceptual theories, all attempt to explain children's drawings from the way that they process visual information. Gestalt theories (see Hochberg, 1978) and those of Gibson (1979) and Marr (1977, 1982) were the most influential in this framework (see also Arnheim, 1969; Gombrich, 1982). Gibson argues that the constantly moving eye unconsciously detects continuously changing sensations. However, perception is aimed to extract the structural invariant aspects of visual inputs. Children, striving to impose meaning and order, are less sensitive to the appearance of objects from a single fixed perspective, and their visual perceptions preserve the general and ordinary appearance of forms. Therefore, objects are perceived as wholes and not as projected surfaces. It is this aspect of perception which accounts for the visual inaccuracies in children's drawings.

Gestalt theory often views the child not just as a passive consumer but rather as a constructor of visual perceptions. The *Gestalt*, or what Hochberg described as the *canonical form*, refers to the *view* that best displays an object's characteristic structure (1978, p. 195). According to the Gestalt theory, this structure is defined by certain perceptual laws (e.g. contiguity, closure) which impose organisation on visual inputs. This conception is echoed in the work of Marr who proposed a more elaborate information-processing account. At the first level, we encounter view-centred, elementary information (visual primitives) which specifies the local properties of the visible surfaces. Although this process is sufficient for shape perception, it is not adequate for shape recognition which entails conceptual knowledge of objects (name, function and class). Form recognition is mandatory for pictorial representation, and both of them become possible at a second level of processing in which an object-centred perceptual mode is set in motion, transforming the visual primitives. The object-centred perceptual structures are based on the natural axis of an object, whereas the view-centred ones on the natural axes of the viewer. At the lowest level of processing, the object-centred representations include volumetric aspects of objects whereas later, fine co-ordination of information results in constructs that include depth, volume and spatial context.

Perceptual theories and Marr's model in particular have certain implications for drawing development. First, they emphasised that perceptions –after knowledge interference– bear only structural resemblance and not exact correspondence to visual inputs. Marr explicitly states that the perception of objects is defined by their natural axis and volume. From this contention it follows that these canonical and undifferentiated views have a privileged status within the visual system, and subsequently are more readily accessed. Willats (1987, 1992), adopting his model, maintained that the early graphic forms (graphic primitives) used by children denote objects as volumes and not as projected surfaces (occluding edges). Therefore, the undifferentiated drawings of young children originate from their perceptual limitations, and not, as cognitive theories have claimed, from their conceptual deficiencies.

Further, Marr's model illustrates how the intellectual factor renders the perceptual processes selective and constructive. Intellectual factors interfere long before the construction of graphic equivalents, at the time when visual impressions (representations) are formed. There is a whole range of theories endorsing the view that the child has lost access to his pure mode of vision (view-centred specification) before he attempted to draw, due to knowledge acquisition (linguistic symbols: Bühler 1930 & Sully 1895; experience: Lowenfeld, 1952 & Werner, 1948). Costall (1995, 1997), critically evaluating these theories, states that their basic premise is that children's "innocent" eye, once innately disposed to read visual inputs as forms, shapes and visual surface features, is later corrupted, and see them as categories and conceptual entities[4]. This is encapsulated in Wolf's (1997) beautiful motto "the eye is no longer a camera, but part of an embodied mind" (p. 191).

The exceptional artistic ability of autistic children (e.g. Nadia Chomyn) has been taken by some authorities as the most compelling evidence for the existence of the primordial sensory core, repressed by language and conceptual knowledge[5]. Lorna Selfe (1977, 1983, p. 86) claimed that it is precisely due to these children's conceptual deficiencies that the sensory core can be accessed, resulting in photographic realism[6]. This explanation has been heeded by other authorities (Arnheim, 1980; Cox, 1992, 194-200; Pariser, 1981; Thomas & Silk, 1990, p.136-139; Winner, 1982, p. 187).

Summarising the main ideas of the two theoretical approaches, it seems that for the perceptual theories children draw what they see, while for the cognitive theories children draw what they know. Yet, both seem to converge on the supposition that children's knowledge about objects give rise to concepts which include the defining features of an object's origin or function. These "internal descriptions" are readily accessed in young children, resulting in drawings, which represent categorical and invariant properties, rather than visual and variable ones. From the cognitive perspective, this is more a deliberate outcome, whereas from the perceptual one, it is an intrinsic disposition of the visual system.

[4] Costall (1995, 1997) calls these theories *repression theories* and maintains that there is no persuasive evidence for the existence of a sensory core which, once accessed, can lead to perspective art.
[5] An appeal to the concept of the innocent eye has been made to explain the paleolithic cave art and that of Bushmen and Australian aborigines (Bühler, 1930; Lowenfeld, 1952; Verworn, 1917)
[6] Pertinent is also the dual-coding theory of Paivio (1971), who assumes two separate representational codes; iconic and linguistic.

Still, their major limitation is the underestimation of the skills required to convert these internal representations into graphic shapes. This omission is amended by a third theoretical approach.

5. PERFORMANCE FACTORS: CHILDREN DRAW WHAT THEY CAN

This theoretical framework rejects the cognition/perception distinction. Instead it draws attention to the difference between competence and performance, *availability* of drawing devices and their *accessibility*. The first attempt to tackle the development of children's graphic planning strategies was made by Kellogg (1970), whose approach to the problem is perhaps the most daring. She attributed formal and aesthetic properties even to the scribble-pictures *(basic scribbles, diagrams, combines, aggregates)*. Later, several investigators examined systematically the graphic vocabulary of drawings (Freeman, 1980, *drawing devices*; Gombrich, 1972, *pictorial schemata*; Willats 1977, 1981, 1985, *drawing systems*), whereas others went beyond their "surface structure" and analysed the drawing process (Freeman, 1972, 1980; Goodnow, 1977; Van Sommers, 1984, Wallon & Baudoin, 1990). Research findings convincingly show that picture making is far more complicated than a simple translation of a mental image onto paper, and that drawing development is better understood in terms of the gradual overcoming of the multitude of problems inherent in the graphic production. These involve inventing the graphic means to map 3D information on the 2D pictorial surface, closely monitoring the drawing process, and overcoming strong response biases and executive constraints (for a review see Thomas, 1995). Freeman (1976, 1980), one of the strongest advocates of this approach, asserts that children's inability to draw what they see reflects inadequacies in planning and execution strategies rather than failures in perceptual analysis or conceptual understanding of spatial relations.

Freeman postulates that children can achieve view-specificity in drawing (partially) occluded scenes not so much by learning to draw what they see, but by learning *not* to draw what they can't see. The representation of (partial) occlusion in the 3D on the picture plane, requires the use of the hidden line elimination device (HLE) which captures the "negative" emphasis of his earlier suggestion. Until this device becomes acquired, children might be competent to conceptualise complex spatial relationships in three-dimensional space but they won't know *how* to represent them pictorially. The most compelling evidence in support of this claim comes from Hagen's study (1976; see also Beal & Arnold, 1990) which revealed that very young children (3 to 5) were accurate in matching pictures of partially occluding scenes to their own corresponding view, but failed to produce them graphically. Partial support has been also provided from studies which show that children prefer pictures which depict perspective and partially overlapping forms, but are unable to reproduce them (natural scenes: Golomb, 1983; Lewis, 1963, solid forms: Kosslyn *et al.*, 1977, natural forms: Fayol *et al.*, 1995; Goodnow *et al.*, 1986). Although children in some of these studies could not always give articulate explanations of their choices – based on the pictorial devices used to create the illusion of realism – the results point to the lag between perceptual awareness of an adequate representation and its realistic portrayal.

Freeman, along with other researchers (Freeman & Hargreaves, 1977; Goodnow, 1978; Goodnow & Friedman, 1972; Light & MacIntosh, 1980), stressed that the graphic medium has its own limitations and conventions which can mitigate view-specificity. Studies, using transparent models from which children had to copy a within- or a between-form spatial relationship, illustrate the operation and the effect of such restrictions. For instance, children of 5 to 6 years of age were unable to represent the spatial relationship of "behind", copying from a transparent array where a defining feature of an object (handle of a mug) or part of another (house-behind-beaker) was visible through the foreground contour (Davis, 1984;

Light & MacIntosh 1980). Since no concealment of information is involved, children's difficulty in producing a view-centred drawing can't be attributed to a cognitive urge to preserve the forms' completeness, rendering them recognisable. What seems to account for children's inaccuracies is that drawing within the contour of a foreground form is allowed to denote exclusively the spatial relation of *inside* (Freeman & Janikoun, 1972; Taylor & Bacharach, 1982). It is only if this drawing convention is violated that a view-specific representation can be accomplished. Young children, who are less flexible due to their limited graphic vocabulary, are more reluctant to tolerate this violation and its resulting ambiguity.

View-specificity can be impaired not only when a drawing convention needs to be transgressed but also when a drawing sequence needs to be modified. Cox (1992, p.112-116) and Van Sommers (1984, p.106-114) claim that early choices in compiling a drawing can determine the final result, and that some errors (or difficulties to adjust to the task demands) are the by-product of the procedural aspect of the drawing action (see also Goodnow, 1977; Johnson *et al.,* 1979; Karmiloff-Smith, 1990; Zhi *et al.*, 1997). Cox emphasised the representational significance of the subject matter in guiding sequential decisions. To draw a person behind a fence, fishing from a boat or wearing a coat, children spontaneously form a structural description of the scene where the person has temporal precedence. This descriptive order, albeit "logical", will set in motion an execution sequence where the occluded form will be drawn before the occluding one, and will result in a transparency. Only children experienced with the drawing medium will be able to anticipate the final product and modify this drawing process in favour of the graphic rule which prescribes action in the reverse order. Van Sommers (1984), arguing from a similar perspective, added another variable. He maintained that, when drawing from verbal instructions, the order of mention in the linguistic description affects the conceptualisation of the scene and subsequently the sequence of production. Thus, in a two-item array, if the occluded object precedes the occluding one in the verbal instructions, the act of drawing is likely to follow this sequential pattern, and subsequently result in a visually inaccurate representation of the array.

These examples suggest that there are conditions where the presence of certain antagonistic or conflicting factors contravenes the accessibility of the HLE device which might be within the graphic vocabulary of children. Freeman (1980), discussing the plethora of production problems, acknowledges the potent role of strong response biases in the organisation of a drawing. He also referred to the concept of canonicality, arguing that the generic, internally represented exemplars of objects or scenes interfere with the graphic deletions required in drawing a partially occluded object. A number of studies have shown that especially when children are asked to draw familiar objects or spatial relationships, they will automatically use over-learned graphic equivalents which capture their functional and idiosyncratic properties even though are hidden (Deregowski *et al.*, 1996; Goodnow, 1977; Ives, 1980; Ives & Rovet, 1979). Like the perceptual account, this approach interprets "errors of commission" as the accidental outcome of natural biases, contrary to the cognitive account which views them as the outcome of a deliberate choice to be maximally informative.

Research on production difficulties demonstrated the pitfalls of early competence theories and opened the way to the more systematic examination of children's drawings. While it was primarily concerned with technical aspects, trying to identify graphic devices and conventions, it demonstrated the within-age variability of children's performance. Subsequently, drawing research was placed on a more explicit and rigorous experimental footing, investigating a number of different performance factors which can interfere with the representation of view-specific information.

In spite of differences between the three theoretical perspectives, all of them interpret the invariant, object (array)-centred drawings as the result of a strong bias that gives representational precedence to the objective properties of forms and relationships in the world over the phenomenal ones in the visual field. Based on the different interpretations offered to account for this phenomenon, a number of studies have manipulated various factors to observe possible changes in performance. Either explicitly or implicitly, they have tried to promote information pickup by the visual system, by promoting attendance to the properties of the visual field and by repressing the overriding effect of conceptual knowledge. As the following section will reveal, contrary to traditional beliefs, the ability to resist the natural reliance on canonical internal representations is not totally dependent on age.

6. DETERMINANTS OF VISUAL REALISM

A drawing is a socially mediated object. To reiterate Costall's words from the introductory chapter, "pictures are made by people with people in mind" (1985, p. 28). In this line, a number of authors have criticised certain cognitive tasks for underestimating children's true abilities by providing very limited, or even misleading information, about what is required (Bryan, 1974; Donaldson, 1978; Donaldson, *et al.*, 1983). The socio-cognitive approach to drawing development asserts that the way that children interpret the task demands will determine how they will draw the requisite topic for the experimenter. In parallel with Olson's (1972) work on language, the drawer chooses the graphic forms – like a speaker chooses words – which will best serve his communicative purpose. Applying here Grice's (1975) conversational Maxim of Quantity, children should not show occluded material intrinsic to a given layout if the context explicitly or implicitly demands so. Thus, the selection of the relevant information will depend upon the clarity with which the task demands are presented to the child. A number of studies demonstrate that children are not always concerned to represent a notional view of a scene. Rather, their representational intentions are determined by the linguistic and non-linguistic context in which the task is presented.

6.1 Linguistic Content

6.1.1 Instructional clarity

Lewis *et al.*, (1993), referring to Grice's conversational analysis, remarked that the standard request in occlusion tasks "please can you draw this (point to the model) for me" is vague and violates his Maxims of Quantity and Manner. Considering that young children rarely draw looking at a model, this appeal neither carries sufficient nor precise information about what it is in the model the experimenter is alluding to. In an attempt to disambiguate the task demands, a number of researchers investigated the effect of verbal instructions on the graphic output. Empirical evidence suggests that rigorous specifications in verbal instructions to draw viewpoint (i.e. draw exactly what you see from where you are sitting) yielded significantly higher levels of view-specificity when drawing a single object (Beal & Arnold, 1990; Lewis *et al.*, 1993) or a spatial relationship between objects (Barrett & Bridson, 1983; Barrett *et al.*, 1985; Barrett *et al.*, 1988; Guthrie, 1994). However, there is a certain inconsistency between the studies as to the magnitude of the instructional effect at different age levels. While in the studies of Barrett & Bridson, (1983), and Barrett *et al.*, (1985), the enriched instructions promoted the incidence of partial occlusion only from the age of 7 years onwards, results from the remaining studies showed sensitivity to verbal cues at younger ages (5 and 6 year olds).

As Smith & Campbell (1987) note, this inconsistency can be attributed to the content differences and the level of specificity of instructions[7], as well as to the properties of the given scenes. Barrett *et al.*, (1985) also commented that their results might not extrapolate to other tasks with a different level of instructional clarity. However, consistent with more recent propositions (Bremner & Andreasen, 1997; Bremner & Batten, 1991), these findings overall imply that children prior to the age of 7 might have a preconception that drawing is actually about showing viewpoint when copying from a model. This possibility doesn't seem to conform to the abilities of children who, according to the stage theories, should be intellectual realists. Another implication is that children might not be as limited, as has been previously suggested, in their graphic skills, but can abandon a preferred graphic strategy in favour of the visually realistic one (HLE). Yet, it is fair to expect that depending on the level of graphic complexity of an array, children younger than 7 years of age will demonstrate differential ability to deploy the HLE device, even if they can accurately infer the task requirements from the verbal instructions.

6.1.2 Communication game

In a similar manner, it was thought that, alerting children to the communicative purpose which the resultant drawing is intended to serve, could direct children's attention to the visual properties of a given array and subsequently result in visually realistic representations. In the standard experimental paradigm, the drawing task is presented in the context of a communication game: children are told that their drawings will be used by someone else to find the viewpoint of the artist (child). Thus, if children's object (array)-centred drawings stem from a deliberate intention to make familiar objects or known relationships recognisable by the beholder, they might change their drawing attitude when informed that what renders a representation adequate is making *their* point of view of an array identifiable.

[7] Barrett and Bridson, (1983) identified three instructional levels, varying on the amount of information given to guide children's representational decisions.

Light and Simmons (1983), using a two-item array in depth, found that the game condition elicited significantly more occlusion responses from children of 7 to 8 years of age, but had no effect on younger children. On the contrary, in the study of Smith and Campbell (1987) the overall effect of the communicative setting was nonsignificant, and paradoxically enough, the standard condition elicited more view-specific drawings from children of 6 years of age onwards. Apart from the similarity of the materials used in both studies, there were procedural differences between the two studies which might have accounted for the irreconcilable results[8].

The game paradigm was also used by Vanessa Moore (cited in Freeman, 1980) to investigate its effect on the incidence of canonical errors with forms in unfamiliar positions. Moore presented children with a set of opaque cups in different orientations and dropped a sweet in the one which was turned so that its handle was totally occluded. In the game condition, children had to draw this cup so that another child could use the drawing to find which cup from the set had the sweet in. Paradoxically, this condition, far from promoting visually realistic responses, increased the incidence of canonical errors. 4 and 5 year olds drew the cup with its handle at the side significantly more often than in the neutral condition, where children were just asked to draw what they see. Davis (1983b, cited in Davis 1985) later replicated this result.

The general conclusion is that placing the drawing task in a communicative setting has a limited effect in the promotion of viewpoint encoding, while it can be counter-productive at younger ages. Researchers seem to adopt Freeman's suggestion (1980) that the

[8] In the study of Smith & Campbell, the duration of the test was longer as each child was tested twice, with a picture selection task introduced between the two trials. Part of the instructions given prompted children just to draw the shapes (balls) while the experimenter indicated *round* shapes kinaesthetically on the table. In the game condition, children were given either the specific role of "artist" or "detective" whereas in Light and Simmons' study these labels were not used. Possibly children might have interpreted negatively the role of a detective and didn't draw with a communicative spirit.

communication-game paradigm directs children's attention to the target model and inadvertently increases children's reliance on the invariant properties of within- and between-object relationships. Possibly, explicitly introducing a particular recipient of the communicative message exerts higher pressure to be informative and results in the error of commission[9]. Such an explanation implies that, when a visually realistic representation does not necessitate the concealment of defining features, the effect of the communicative context might be more pronounced. In fact, Light and McEwen (1987) revealed that, when 5 year olds were asked to draw L-shaped arrays of three bricks in various orientations which did not involve occlusion, they made significantly more view-specific responses if their drawings were to be used by another child to construct the array.

6.2 Properties of Materials

6.2.1 Familiarity of forms

A hypothesis, derived from the basic premises of conceptual and perceptual theories, was that, if view-specificity is impaired by the intrusion of conceptual knowledge, differences in performance should be observed depending on whether the materials used are familiar and figurative objects or not. The rationale was that the familiarity of the model can constrain information pickup by the visual system since its symbolic content might activate stereotyped (canonical) forms. This hypothesis gained support from a number of studies which demonstrated that meaningless or unfamiliar models favoured view-centred representations (Beyer & Nodine, 1985; Ingram & Butterworth, 1989; Krascum *et al.*, 1996; Moore, 1987; Phillips *et al.*, 1978; Reith, 1987, 1988; see also Dziurawiec & Deregowski, 1992). These findings suggest that children might encode the non-figurative and less-meaningful materials in terms of visual contours, while the familiar shapes are liable to be encoded in terms of conceptual classes.

[9] An alternative explanation which needs to be tested was offered by Light (1985). Using the findings from verbal tasks, he suggested that when communication fails, children younger than 6 years of age

Milbrath (1998) alludes to the possibility that if children are copying from a model array which involves familiar forms in a well-known spatial relationship, this can affect children's looking behaviour. Such a condition encourages children to look only long enough to classify the object or the structural relationship between objects and then to rely on their internal models –perceptual or conceptual – and their knowledge-centred schemata during drawing. Reith (1988) adds that the meaningfulness of stimulus materials not only influences the level of elaboration but also the execution of the drawing which can act as an obstacle to faithful depiction.

Two more procedural variables seem to trigger the representation of the objective rather than the phenomenal properties of familiar material; visual and manual inspection of a given array, and naming of the forms prior to the drawing act (Bremner & Moore, 1984, exp. 1; Guthrie, 1994; Krascum *et al.*, 1996; Lewis *et al.*, 1993; Moore, 1987). The results from these studies unanimously revealed that 5 and 6 year olds are capable of repressing hidden-features inclusion if they are not exposed to the non-visible parts of the array and if the identity of the forms is not mentioned; otherwise, children shift to object- and array-centred depictions. Comparable findings were obtained when the experimenter discussed the appearance of the scene involving partial occlusion, prior to the drawing act (Cox, 1986, p. 81). While this manipulation increased the incidence of separate configurations among children of 5 to 6 years of age, it succeeded in eliciting partial occlusion output only among older children.

Two explanations, not necessarily mutually exclusive, are offered. It is possible that in this context children might have misinterpreted the experimenter's intent as to produce drawings which inform about the object's genus and salient functional aspects, or inadvertently alert

often blame the recipient of the message and not the deficiency of the message itself.

them to what is not present in the visual field. The former is favoured by those who maintain that children can change their representational intentions in response to the communicative objectives specific to the task, and the latter by those who believe that these factors operate more at the level of the primary encoding of the structural features of the scene at hand.

6.2.2 Visual contrast between forms

A number of researchers systematically studied how the formal attributes of the objects present in an array influence the degree of realism of the representation of between-forms relationship. It has been suggested that when visual occlusion is involved in a two-item array, a pair of identical objects is more likely to increase the incidence of two complete and spatially separate forms. Cox (1985, 1986, p. 78-80), asking children to draw a model with one object behind another, found that children were better able to depict partial occlusion when the objects had a dissimilar shape rather than when they had a similar one. She also asserted that the greater the differences between the occluded and the occluder, the more likely is a view-specific response. In general, the dissimilar-object effect was more pronounced from the age of 6, but even children of 4 years of age demonstrated some degree of visual realism with material which differed on a number of attributes.

At first, it was suggested that layouts with dissimilar objects probably enhance the visual exploration of the array (scanning hypothesis: Cox, 1986), but recent evidence suggests that they facilitate the execution and the planning process (planning hypothesis: Morra *et al.*, 1996). The later proposition is that, when children encounter an array of similar shapes (i.e. two balls), they are more likely to encode it as a group of identical objects and subsequently draw them deploying the same graphic form (complete). Morra *et al.*, (1996) claim that such a case constitutes a misleading condition and necessitates a fair amount of attentional resources to overcome it. Young children, who are limited in this domain, are more likely to form an inappropriate internal description which will induce a visually inaccurate depiction of the spatial relation. It is also possible that the visual contrast between parts of spatially related objects draws attention to the junctions and stresses the discontinuities of visual surfaces.

6.2.3 Distinctive orientation of forms

Another factor which plays a crucial role in facilitating a view-specific approach when drawing a model, is employing forms with distinctive orientations. Although it has been already demonstrated that the canonicality or "best view" of individual items –or of an array as a whole– can impede visual accuracy, there is evidence that non-symmetrical and fronted objects may be particularly propitious for the representation of partial occlusion in a two-item array arranged in depth[10] (Light & Foot, 1986; Light & Simmons, 1983; Radkey & Enns, 1987; Teske *et al.*, 1992).

The argument that is put forward is that the most frequently used model array in occlusion studies is two balls one behind another. Yet, it offers a very limited range of views as balls lack a distinctive orientation, whereas for example a house behind a wall elicits many and perceptually unique viewpoints. As a consequence, in the latter condition, children will have to attend more to the model in order to select the precise vantage point from which the scene is being viewed. Thus, enhancing visual attention could result in increased incidence of realism.

6.2.4 Scenes with intrinsic hidden-features relevance

We have already examined the views of those who argue that children's drawings are made with an intentional communicative spirit and maintain that their object (array)-centred drawings reflect a comprehension rather than a cognitive deficit. We have also reviewed the

[10] This effect applies to conditions where children are copying from a 3D array rather than from a line drawing or from memory.

studies which, based on this contention, manipulated the linguistic content of the instructions by making rigorous specifications of tasks demands (explicit instructions, communication-game paradigm). Revisiting the same issue, a number of researchers tried to disambiguate the task objectives by manipulating the non-linguistic context in which the drawing is placed.

Using an approach reminiscent of the study of Hughes and Donaldson (1979) on perspective taking, Cox (1981, study 5) 'embedded' an occlusion task into the context of a 'cops and robbers' storyline, where a robber was partially hidden behind a wall from a policeman. Nearly half of the four year olds and the majority of children aged six and above used partial occlusion in this condition. Cox claimed that these findings show convincingly that children younger than 6 years of age have been underestimated on their cognitive and graphic skills required for view-specificity. These are evidently available, although not directly accessible at all times. Interpreting the determinants of her successful condition, she concluded that the materials used and the way they were presented cued children to the 'representational' relevance of hidden-feature omission, since the notion of hiding was an intrinsic property of the scene (Cox, 1985, 1986, p. 75-84).

The effectiveness of the "hiding scenario" in eliciting visual realism has been explored by subsequent studies, but the findings are inconclusive. In the study of Light and Foot (1986), 6 year olds used partial occlusion even without a compelling hiding story, whereas in the study of Ashton (1997), a comparable story had a facilitating effect in this age group but not in younger children. On the contrary, Arrowsmith *et al.*, (1994) revealed that 4 year olds were contextually sensitive to scenes involving the idea of hiding, and 5 year olds could even transfer their view-centred approach to less propitious scenes, presented at the post-test. On the basis of the overall findings, no firm conclusion can be drawn about the extent to which a hiding scenario can prompt view-specificity. However, children demonstrated some sensitivity to the task materials – even if they involved familiar forms – as long as the omission of the nonvisible features was an intrinsic property of the scene.

In another variant of contextual sensitivity, a number of researchers studied the potency of the canonical bias by manipulating the way a familiar object was presented in a non-standard orientation. In the classic experiment of Freeman & Janikoun (1972), children before the age of 8 ignored the visual appearance of the model and included the hidden handle of a cup in their drawings. However, a series of subsequent studies demonstrated that 5 year olds can resist the canonical error if this target cup is simultaneously presented with the same one in a standard orientation (handle at side) (Davis, 1983, 1984, 1985a). The sharp visual contrast between the forms promoted the understanding of what is about the array that the experimenter is alluding to and suppressed the tendency towards hidden-feature inclusion. Further, Davis and Bentley (1984) showed that the contrast effect is powerful not only in a within task but in a between task condition as well. Five-and-a-half year olds were able to resist the canonical error if beforehand they had copied a cup with the handle in sight. Similar results has been reported also by Bremner (1985) who succeeded in promoting view-specificity from the same age group, using various layouts arranged in depth without involving occlusion.

In conclusion, the empirical findings from the studies reviewed in this section indicate that young children are not rigidly anchored to a certain depiction bias. Even children who are expected to make 'intellectual realist' responses can be resourceful and adapt their drawings to suit the informational demands of a given situation if these are readable from the context.

6.3 The Model's visibility

Although pre-school children are less likely to draw copying from a prescribed scene, the provision of a model can guide children's drawing decisions under some conditions. Cox (1978), criticising the study of Freeman *et al.*, (1977), asserts that young children, limited in their understanding of verbal terms, might be particularly constrained when they are asked to draw from memory a depth relationship involving occlusion. A number of other researchers maintain that studying spontaneous drawings at transitional stages of development is less likely to reveal how competent children are. Their argument is that a model can help children to visualise a scene from a particular viewpoint and to extract the pictorial devices needed for translating its projective appearance on the pictorial surface (Chen, 1985; Phillips *et al.*, 1978).

Three studies systematically manipulated the model's visibility for the pictorial representation of partial occlusion of a two-item model array. Radkley and Enns (1987) argued that the developmental changes in drawings of occlusion reflect general cognitive changes in perspective taking ability, since children before 9 years of age would have difficulties selecting one perspective and maintaining it consistently during production. Employing an aperture device which restricted children's view, they obtained an increase in total occlusion scenes by 5 and 6 year olds, and in partial occlusion scenes by 6 year olds, compared to the free-viewing condition. Two recent studies, comparing visible and screened models, yielded overall equal numbers of partial occlusion drawings (Klaus, 1992; Morra *et al.*, 1996). Yet, while Morra examined the performance of 8 year olds, Klaus showed that the benefit of a visible model was different at various age levels[11]. It should be noted however that there is no strict equivalence between Radkey and Enns' study (1987) and the two recent

[11] Five year olds were totally unaffected by this manipulation and 7 years olds only moderately affected. Visibility of the model affected the 9 year olds by improving performance when drawing on an horizontal surface (paper on table) but inhibiting it when drawing on a vertical surface (paper on wooden board).

ones with respect to the restricted viewing condition. Whereas in the former the model was visible but through a peephole, in the latter ones it was totally hidden.

Naturally, some inconsistencies exist as to the extent to which a model can be facilitating at different ages depending on the properties of the scene, the nature of skills required and the degree of visual realism expected (quantification of shapes' dimensions, relative spatial position). The fact that even 9 year olds have difficulty utilising all the aspects of the projective shape (Reith, 1994) suggests that copying from a line drawing or a photograph of a scene should be easier than copying from a 3D model (see Mitchelmore, 1985 and Cox, 1992, chapter 9). In general it's fair to say that given the opportunity to attend to how things appear in the visual field, some children might resist reliance on internal schemes – especially when scenes involve familiar objects – and can be guided in their subsequent drawing decisions.

A number of influences on children's drawing performance were demonstrated by the studies reviewed here. The empirical findings suggest that interpreting children's drawings in terms of a clear-cut developmental progression from intellectual to visual realism cannot adequately explicate the flexibility and resourcefulness that children of 5 to 7 years of age often reveal. Wolf (1997) reflecting upon this suggests that "a vision of growth should be pictured not so much as a ladder of ascending stages but as the development of a repertoire of choices" (p. 189). Since drawing performance depends upon a number of factors (conceptual maturation, perceptual abilities, graphic skills and procedural knowledge), a systematic analysis of task demands is required in order to offer a satisfactory account of the resources children have at their disposal. This task is undertaken in the following chapters.

CHAPTER V

OCCLUSION: STRUCTURALLY SEPARATE OBJECTS

1. INTRODUCTION

The studies carried out in the present and following chapter were designed to examine the way in which children can combine very common shapes to represent graphically simple spatial relations involving partial occlusion. Having reviewed the literature on occlusion, it is evident that, despite good experimental evidence for the developmental shift from 'object-centred' or 'array-specific' representations to 'view-centred' ones, the nature of this progression is still a theoretical puzzle. The following series of tasks were designed with due attention to the kinds of problems children face and the strategies they use to deal with the loss of integrity of the occluded object.

The first concern was to disentangle the production difficulties from the conceptual difficulties by varying the demands on graphic competence. In certain tasks, children were asked to draw a nominated scene from memory, whereas in others they had to copy it from a line drawing. In the latter case the graphic device (HLE) was provided at hand without a graphic translation of three-dimensional information being required. Further, some arrays had the occluding form pre-drawn on children's papers, alleviating also the planning problems which operate during production (see Cox, 1992, p.112-116; Karmiloff-Smith, 1990; Van Sommers, 1984, p.105-114). Thus, this design enables us to evaluate the impact of conceptual knowledge on graphic output when technical competence and organisational demands are reduced.

Trying to identify other possible causal variables affecting the output of visual accuracy, factors relating to children's attentional resources and to the task's communicative clarity were also explored. The former aspect relates to empirical findings suggesting that arrays with enhanced visual contrast between the occluded and the occluding forms elicit higher occlusion rates (Cox, 1985 & 1986; Morra *et al.*, 1996). By varying the degree and the nature of inter-item differentiation across the tasks, the merit of the 'dissimilar-material effect' was assessed as well as its underlying mechanism. In addition, the communicative clarity of the task was examined with respect to its social intelligibility; a factor identified by theorists adopting a socio-cognitive approach. Thus, enriching the contextual salience of certain tasks by means of the concept of 'hiding', its presumed facilitating effect was evaluated. The general research interest was not only to identify any systematic variation in performance across tasks but also any age-related differences in children's susceptibility to these manipulations.

The structural characteristics of an array of objects were also investigated in relation to the pictorial presentation of partial occlusion. Partial occlusion in the visual field can result from a variety of spatial relations. Some of these relations are dependent on our line of sight (in front/ behind) whereas some others remain unaffected by the stationary point we adopt (inside of/next to). This feature consequently influences the extent to which the partial concealment of an object by another is an integral property of the array or just a matter of viewpoint. A number of authors, endorsing a cognitive approach to children's drawings, argue that their graphic strategies reflect the structure of the cognitive representations they form of the scenes at hand (Crook, 1984 & 1985; Light, 1985; Nicholls & Kennedy, 1992). If this view is correct, we should expect arrays involving a different spatial relation between forms (i.e. behind or inside) to be encoded in a different way. This will affect the incidence of partial occlusion output in general, and the relative frequency of the non-occlusion strategies (array-specific) in particular, across these scenes.

The studies are divided into two parts. In the first part, covered in the present chapter, the partial occluded scenes involve objects spatially distant from each other, whereas in the second part, presented in the following chapter, the target forms are structurally integrated.

2. GENERAL METHOD

2.1 Participants

From the 115 children who were tested with the standardised instruments in Chapter III, 110 took part in the all the experimental tasks on partial occlusion. The present sample consisted of 39 boys and 71 girls divided into four age groups: 30 children with mean age of 4 years and 2 months (range 3:7-4:7 years), 25 children with mean age of 5 years and 6 months (range 5:3-6:2 years), 28 children with mean age of 7 years (range 6.8-7:5 years), and 27 children with mean age of 8 years and 3 months (range 7:8-8:8 years, SD = 3.31). However, due to sporadic occurences where children either did not produce any data or their drawing couldn't be scored, there are occasional fluctuations in the sample sizes across the tasks.

2.2 Materials

Nine tasks were administered in total. Five of these tasks followed the paradigm of 'one object behind another', and these will be presented in this chapter. For all the copying and in some of the completion tasks, children were presented with a line drawing picture of a partially occluded scene. The pictures were presented on A6 cards and were simplified representations of the scene with no contextual elements. Children were given a standard test booklet where each scene had to be drawn on a separate page. With the exception of three tasks, each page had the outline of the occluding object pre-drawn. The drawing paper was sufficiently thick so as to prevent previous drawings and subsequent pre-drawn forms to be visible during drawing. A more detailed description of the task material will be given when discussing each task separately.

2.3 Procedure

Children were tested individually and completed their drawing without time constraint. The data from the nine tasks were collected in two or three drawing sessions, depending on children's age and level of concentration. The experimenter was always sited to the left of the child and presented him with one task at a time. For the tasks where children were drawing from a model, the stimulus cards were placed on the table directly above children's paper and remained in view throughout the drawing. The children were asked to look carefully at each card and copy the drawing (or complete their own drawing) so that "it looks exactly like this" (point to the card). There was no explicit reference made to the particular spatial relationship between the items of the array (behind, next, inside). All drawings were done freehand with black pencils. On completion of each array, the stimulus materials were removed and children were instructed to turn to the next page of their booklet. The order of presentation of the tasks was fixed and equivalent to the order followed here[1]. However other experimental tasks were frequently administered between the various occlusion tasks in an attempt to confine the practice effect between successive items assessing similar skills, and to sustain children's motivational level[2].

2.4 General scoring method

Earlier research by Freeman *et al.*, (1977) and Cox (1985) provided a basis for the construction of the scoring scheme and the selection of the drawing criteria. First, the drawings were classified into those that depicted partial occlusion and those that did not. The main criterion for classifying a drawing as showing partial occlusion was the omission of the

[1] For instance, in the assessment of children with arrays of one form behind another, the Apples task was always administered first with the Bottles next, followed by the Sunset and the Balls tasks. The latter task had two conditions which, due to their similarity, were presented in a counterbalanced order.
[2] Some cards were multipurpose as they contained features for the assessment of the pictorial representation of partial occlusion as well as of the spatial axes (Chapters VIII-XI). Such cards were the ones used for the Sunset, Bottles and Houses tasks. In these cases, children were asked to draw first the vertical (tree-Sunset, chimney-Houses) or horizontal feature (water level-Bottles) that was missing from their paper and latter to complete their drawing by drawing the partially occluded item.

concealed part of the occluded object. The graphic device associated with this pictorial representation was *hidden line elimination* (HLE). As a consequence, all drawings were divided into those depicting the occluded object with the HLE device and those with a complete contour.

At a later stage of data analysis two more refined criteria were adopted in order to assess the structural and spatial arrangement of the forms and to investigate the various solutions that children used in a more detailed way. Occlusion and non-occlusion drawings were sub-divided according to the proximity and the orientation of the contour of the occluded object relative to the occluding one. Some researchers (Cox, 1985; Ingram & Butterworth, 1989; Willats, 1977) maintain that in a successful partial occlusion configuration the interrupted boundary should be firmly attached around the complete one. Applying the criterion of proximity, Figure 5.1 presents the various drawing solutions children can adopt representing a scene, using as an example an array where one ball is behind another[3].

Seven types of structural arrangements can be identified, progressing from those where the (further) object is totally eliminated (A) to partially exposed (B & C) and fully represented (D-G). Among those drawings that depict only the visible part of the object, two strategies were distinguished: partial occlusion (B) and separated occlusion (C). What differentiates the two strategies is that in the first case the interrupted boundary of the occluded form is anchored on the complete outline of the occluder, whereas in the second case it is suspended and separated. As a result, in all the drawings using the HLE drawing device, the ability to omit lines was assessed in relation to the skill to form junction points, and the degree of the realistic portrayal of the spatial relation in question was thoroughly examined.

[3] Not all strategies were pertinent to all the experimental tasks

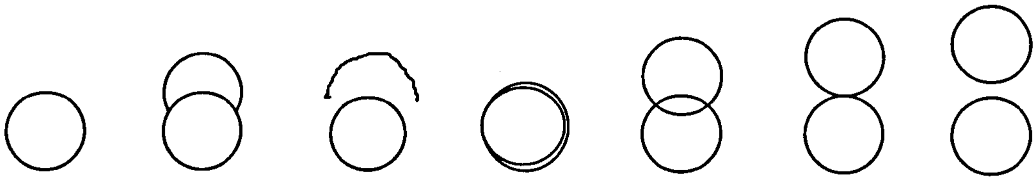


Figure 5.1 Drawing categories of structural arrangement on the basis of relative proximity between the forms. Key: A: Total occlusion, B: partial occlusion, C: separated occlusion, D: enclosure, E: overlap, F: attachment, G: segregation.
Note: Apart from D, the series represent a progressive rising of the point of view of the observer from immediately in front to directly above. An alternative spatial alignment (horizontal) could result in a successive displacement of the line of sight from immediately in front to laterally beyond.

Among those drawings which depict the occluded form with a complete boundary, four structural arrangement patterns evolved. Children either spatially separated the two complete outlines (G) or spatially united them by producing outlines which are concentric (D), mutually overlapping (E) or attached at a point (F). As a consequence, within both perspective-appropriate (B & C) and inappropriate drawings (D-G), segregated and unified structural configurations were found.

To investigate how faithful was the depiction of the spatial arrangement of the two objects to the stimulus array, the alignment of forms relative to the spatial axes was also examined. Three arrangement patterns were identified: (i) horizontal alignment where the forms were presented side by side, (ii) vertical alignment when they were placed one on top of the other and (iii) oblique when the items were in diagonal relation to each other[4]. With the exception

[4] This latter alignment pattern was pertinent only to one of the occlusion tasks (Apples task). For this case, the line drawn from the centres of the items had to fall within 21° to 69° from the horizontal axis for the alignment to be oblique, and within 20° from the horizontal and vertical axes to be classified as

of total occlusion and enclosure, all the previous structural arrangements in Figure 5.1 could be presented with one of these three spatial alignments. Figure 5.2 demonstrates the alignment patterns using the partial occlusion paradigm.

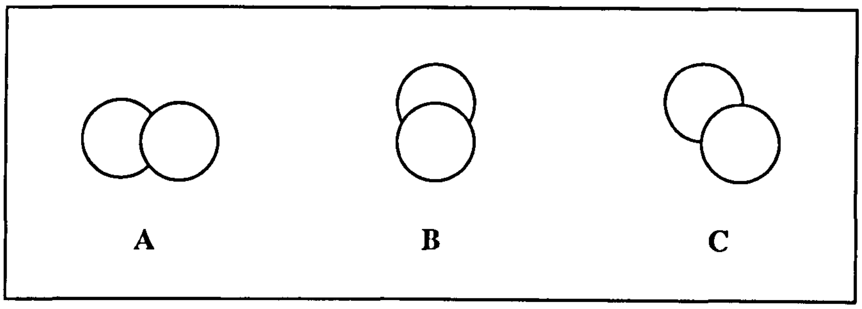


Figure 5.2 Drawing categories of spatial arrangement on the basis of the relative alignment of the forms to the axes. Key: A: Horizontal alignment, B: Vertical alignment, C: Oblique alignment

3. THE APPLES TASK

The first task of the series was designed to study how children attempt to draw a depth relationship between two identical 'nonfronted'[5] objects with minimal internal features and no contextual elements. Although this type of material has been much employed in previous research (Arrowsmith *et al.*, 1994; Ashton, 1997; Chen & Holman, 1989; Cox, 1978 & 1981; Freeman *et al.*, 1977; Radkey & Enns 1987; Smith & Campbell, 1987) with mnemonic or 3-D copying tasks, there is empirical evidence suggesting that similar objects elicit lower levels of partial occlusion than dissimilar ones (Cox 1985 & 1986). Recent work by Morra *et al.* (1996) has found that the 'materials effect' operates at the level of planning. In the present study, the demands on planning were reduced by providing a 2-D model and the occluding shape pre-drawn on children papers. The aim was to observe the effect that this manipulation would have on the representation of view-specific and array-specific information.

horizontal and vertical in respect. In any other task, the drawings were classified as vertical (or horizontal) when the pair of objects was drawn within 45° from the vertical (or horizontal) axis.

[5] This term is used here to refer to forms without a distinct view (i.e. apples, balls, blocks)

3.1 Apparatus

Children were presented with a line drawing of two apples, 3.00 cm in diameter each, placed diagonally one behind another with the far apple to the left of the near one (Figure 5.3-a). The stalk was the one feature which defined the shapes. All children were provided with the near apple pre-drawn on their paper (b) and were required to draw the apple that was missing in order to reproduce the array exactly as it looks on the stimulus card. After each child had completed the drawing, he was asked to point to the card and match his drawing to the equivalent forms on the stimulus card.

a b

Figure 5.3 Schematic illustration of stimulus material.

3.2 Results

3.2.1 Partial occlusion

A gross categorisation was applied dividing the drawings into those using the HLE drawing device and those using a complete contour to depict the further apple (Table 5.1). Over the entire age range, 60 (54.5%) children failed to produce partial occlusion, whereas 50 (45.5%) succeeded. The use of the HLE device was dependent on age [$\chi^2(3) = 28.99, p<.001$] with a significant increase in the preponderance of partial occlusion responses as children were getting older. Binomial tests were used to examine the relative frequency of the two drawing modes (complete and HLE) at each age level. As Table 5.1 shows, children of different ages demonstrated a different pattern of device dominance. Four year olds drew significantly more complete forms, whereas eight year olds produced more partial occlusions. On the contrary, the responses of the two intermediate age groups did not reveal any representational preference between complete and incomplete contours. Comparing the distribution of drawing responses between successive age groups, the difference was significant between the two younger groups [$\chi^2(1) = 8.56, p<.01$ with continuity correction].

Table 5.1 Distribution of drawing devices by age.

Age groups	N	Drawing device: Complete	Drawing device: HLE	*z*	*p*
4	30	28	2	-4.75	<.001
$5^{1/2}$	25	14	11	-0.60	*NS*
7	28	10	18	-1.51	*NS*
8	27	8	19	-2.12	<.05

3.2.2 Structural drawing strategies

Drawings were further classified on the basis of proximity between the contour of the further apple to that of the pre-drawn one. Three main arrangement patterns were identified. Children either drew the visible part of the contour of the further apple meeting the near apple at junctions (partial occlusion) or drew a complete contour either spatially distant (segregation) or proximal to the occluding apple at varying degrees (unification: enclosure, overlap and attachment). The use of the three structural arrangements was age dependent [$\chi^2(6) = 42.75, p<.001$]. The partial occlusion output increased in frequency with age as the incidence of separate arrangements decreased, while unification tendencies were equally salient across all age groups. These results confirm the developmental pattern reported by Freeman (Freeman *et al.* 1977) where apparent age-related trends in the use of structural devices were obtained with partial occlusion and segregation, but not with unification strategies.

The distribution of the structural strategies within age levels revealed the following pattern: four year olds produced mainly segregated (57%) and unified arrangements (37%), while the responses of 5½ year olds did not show any dominant depiction strategy (partial occlusion 44%, segregation 28%, unification 28%). Seven and eight year olds produced primarily partial occlusions and structurally united configurations, with partial occlusion becoming the dominant choice only among the oldest of the groups (7 year olds: partial occlusion 64%, unification 36%, 8 year olds: partial occlusion 70%, unification 26%). To examine in a more detailed way the developmental sequence in which these structural devices emerge, the unification strategy was broken down into three sub-categories. Table 5.2 presents the frequency distribution of the drawing responses within each age group.

Table 5.2 Distribution of structural strategies by age.

		Age groups				
Structural strategy		4	5½	7	8	Total
Partial occlusion		2	11	18	19	50
Segregation		17	7	0	1	25
Attachment		4	0	2	0	6
Overlap		3	6	6	7	22
Enclosure		4	1	2	0	7
Total		30	25	28	27	110

Note The horizontal alignment is used as an example for the illustration of structural devices.

The Table above clearly shows that the spatial segregation of apples is the dominant response mode for the 4 year olds, with a shift in favour to partial occlusion among the rest of the age groups. Also, the tendency to draw the apples separate declines at the same time as the number of partial occlusions increases with a crossover at the age 5½. With respect to the unification devices, the use of enclosure and attachment was rather negligible and mainly found amongst the youngest of children. Overlap was as frequent as any other of the unification strategies for the 4 year olds, whereas it was preferred to enclosure and attachment by all the other children. However, no reliable age-trend was revealed in its frequency.

3.2.3 Spatial alignment patterns

Next, the relative placement of the requested apple with respect to the spatial axes was considered. The frequency data revealed that the majority of children mainly chose only between two ways of aligning the further apple with the pre-drawn one. Across the entire age range, 70 children produced an oblique (diagonal) arrangement, 30 a horizontal, 3 a vertical and 7 produce an enclosure configuration. It has been suggested that children usually reserve the vertical dimension to indicate depth (Bremner, 1985a; Ingram & Butterworth, 1989; Light & Humphreys, 1981; Teske *et al.*, 1992; Van Sommers, 1984, p.110). However, since the oblique perspective was a feature of the scene they had to copy, children had the flexibility to abandon a preferred dimension in favour of another that allowed a better encoding of view-specific information. Figure 5.4 presents the developmental progression in the use of the three spatial alignment patterns.

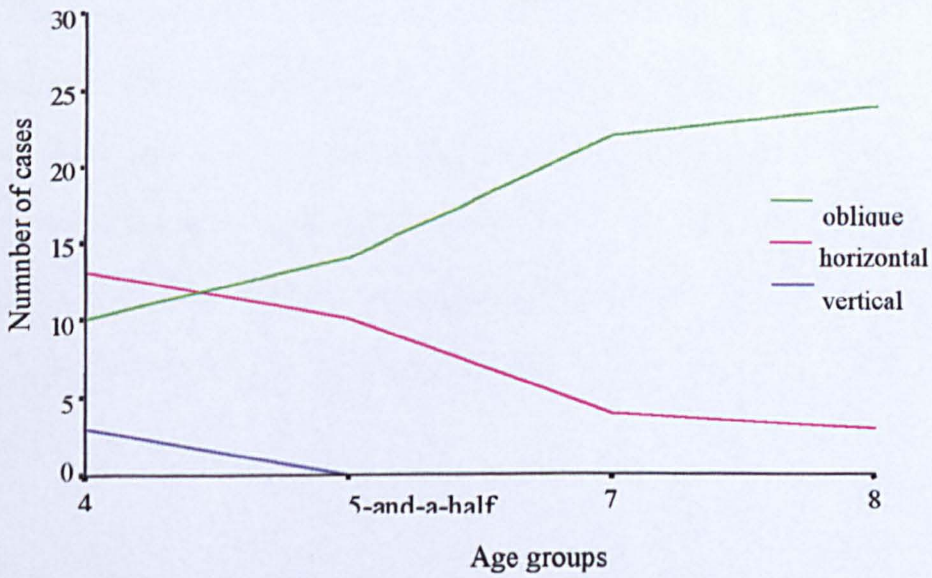


Figure 5.4 Age-related shifts in the relative frequency of the three alignment patterns

Vertical arrangements were only found among the 4 year olds while the use of the other two dominant spatial strategies was age dependent [$\chi^2(3) = 16.49$, $p<.001$]. Children up to 5½ years of age seem to oscillate between horizontal and oblique arrangements with a shift in favour of the latter at the age of seven. In general, the results of this study suggest that there is a clear age-related shift towards oblique alignment with a corresponding decrease in the use of horizontal representations. Also, the prevalence of oblique arrangements over vertical ones – even among the drawings of the 4 year olds – indicates that, although young children could not use partial occlusion as a depth cue, their drawings demonstrate some rudimentary aspects of perspective taking. Also, among the children who did not use partial occlusion, a test of association between alignment strategy and drawing device revealed that using a spatial arrangement to decode depth (vertical or oblique) was related to the most array-specific device. Thus, vertical or oblique alignment was more frequent with spatial overlap, whereas horizontal alignment with segregation and attachment [$\chi^2(2) = 11.79, p<.01$].

3.2.4 Identification trial

When children were requested to draw the apple that was missing from their paper, they had to make an inferential judgement about its spatial position on the basis of the pre-drawn one. Because in the present array the items were identical, children had to assume that, since the pre-drawn apple is at full view on their paper, its spatial equivalent is the near apple at the stimulus array, and subsequently to draw the far one. On completion of their drawing, children were asked to point to the stimulus array and indicate which of the two apples they had drawn on their paper. Table 5.3 presents the distribution of children's choices in this identification trial.

Table 5.3 Distribution of children's representational choices by age.

	Apple drawn			
Age groups	Far apple	Near apple	*z*	*p*
4	14	16	-0.36	*NS*
$5^{1/2}$	16	9	-1.40	*NS*
7	22	6	3.02	<.01
8	19	8	2.12	<.05

Nursery children were equally likely to choose to draw either the far or the near apple. Older children were better at deducing the spatial position of the requested apple in the array, and this ability became established at the age of seven. However, a test of association between age and accuracy of inference fell short of significance [$\chi^2(3) = 7.0, p>.05$].

Considering the type of inference children were making, the data were re-examined on the basis of the alignment patterns children adopted in placing the apple relative to the pre-drawn one. This analysis enables identification of the drawings that preserve the internal spatial relation of the forms whilst bear little resemblance to the model. For example, if children have correctly chosen to draw the further apple, they should have placed it above or behind and/or to the left of the pre-drawn apple. Any other placement pattern would result in an inconsistency between their inferential drawing choice and their subsequent alignment strategy (i.e. choosing to draw the further apple but placing it in front or to the right of the pre-drawn one). Alternatively, if children have chosen to draw the near apple, they could produce a configuration faithful to the internal spatial relations of the stimulus array if they have placed it below and/or to the right of the pre-drawn one. Again any other representation (i.e. placing the near apple above or/and to the left of the pre-drawn one) would result in a lack of equivalence between the apple they have chosen to draw and the alignment strategy they used. These latter instances were simply defined as errors.

Table 5.4 presents the distribution of drawings across the types of alignment patterns by drawing choice and age. The identification of errors was rather precarious in the case of an enclosure configuration. In four drawings, the far apple was enclosing the near apple, whereas in three, the near one encircled it. Due to the difficulty defining which of these arrangements represented a line of sight construction, enclosure drawings were excluded from further analysis.

Table 5.4 Frequency of different alignment patterns by age and apple drawn.

	Oblique				Horizontal				Vertical	Enclosure		
Age groups	Near	Far	Errors	N	Near	Far	Errors	N	Errors	Near	Far	N
4	2	3	5	10	8	2	3	13	3	1	3	4
$5^{1/2}$	4	9	1	14	4	5	1	10	0	0	1	1
7	5	17	0	22	1	3	0	4	0	0	2	2
8	6	18	0	24	2	1	0	3	0	0	0	0
Total	17	47	6	70	15	11	4	30	3	1	6	7

With respect to the spatially accurate drawings, when children produced an oblique alignment, they were more likely to choose to draw the further apple above and to the left rather than the near one below and to the right of the pre-drawn one (far: 47 vs. near: 17). On the contrary, children aligning the apples side by side did not demonstrate any bias in their choice, and were equally likely to draw the far apple to the left as they were to draw the near apple to the right of the pre-drawn one (near: 15 vs. far: 11). Looking at the prevalence of these preference patterns at each age level, the following observation can be made; seven and eight year olds were more likely to draw the far apple in an oblique alignment, whereas four year olds were more likely to draw the near in a horizontal alignment.

Considering the erroneous spatial arrangements on the basis of children's representational choice, the following results were obtained. First, making such errors does not seem to be related to the type of spatial alignment. Although the three drawings where the children aligned the apples vertically were all inaccurate in this respect, in general these errors constituted only a small portion of the total drawings within each alignment pattern. Thus, the majority of children (81%) produced drawings that preserved the true spatial relationship between the objects in the array, despite the fact that some of them were not visually accurate. Nonetheless, making such errors seems to be associated with age. Pooling errors across the three alignment patterns (oblique, horizontal, vertical), they accounted for more than one third of the drawings of 4 year olds (11 out of 26), they were rare among the 5½ year olds (2 out of 24) and were never encountered among the older age groups. These results are in agreement with those of Freeman *et al.*, (1977) where reversal errors in the representation of one apple behind another were not found from the age of 6 years.

Moreover, children seem to favour a different structural strategy for each type of alignment pattern. Looking at the total occurrence of each strategy within the two dominant spatial categories in Table 5.5, there was a strong tendency to use partial occlusion when the apples were obliquely aligned, whereas the dominant device was segregation when the apples were horizontally arranged [columns on total, oblique: $\chi^2(2) = 35.17, p<.001$; horizontal: $\chi^2(3) = 12.9, p<.01$].

Table 5.5 Distribution of alignment patterns by structural strategy and apple drawn.

	Oblique			Horizontal		
Strategies	Near	Far	Total	Near	Far	Total
P. occlusion	0	46	46	1	3	4
Segregation	4	3	7	10	6	16
Overlap	16	1	17	4	1	5
Attachment	0	0	0	2	3	5
Total	20	50	70	17	13	30

In addition, oblique drawings, unlike horizontal ones, revealed a reliable relation between the apple children chose to draw and the device they used to depict the array. In particular, as it would be expected, children used partial occlusion with an oblique arrangement only when they chose to draw the far apple. What was though interesting is that for the same type of arrangement they used exclusively overlap when they decided to draw the near apple. Horizontal arrangements did not show any clear association between the structural strategy children employed and the kind of apple they chose to draw. In general, regardless of what kind of alignment children produced, they properly used partial occlusion only when drawing the far apple (except for one case), they mostly used overlap when drawing the near apple (22/24), and they demonstrated no depiction bias in their representational choice when they used either segregation or attachment.

Summarising the results from Tables 1.3-1.5, it can be suggested that, in order to reproduce the array, making the correct inference about the apple that is missing is a precursor for the use of the perspective-appropriate strategy. Over the total sample, 71 (64.5%) children succeeded in making the correct choice by drawing the further apple. Of those children, 47 (66.2%) produced the most spatially accurate configuration, aligning the further apple above the pre-drawn one along the vertical-oblique axis, 11 (15.5%) produced horizontal arrangements, faithful to the internal spatial properties of the model array, whereas the remaining 13 (18.3%) made spatial configurations which were either ambiguous (enclosures) or incompatible with their choice to draw the further apple. Of the 47 oblique drawings which were the most structurally informative, 46 were also view-specific, depicting the further apple partially occluded by the nearer one. On the other hand, of the 11, less informative horizontal drawings only two depicted the further object partially occluded [$\chi^2(1) = 39.67, p<.001$].

3.3 Discussion

From the analysis of results, we can comfortably make the following observations. The pressure towards itemisation of the constituent forms within the array undermines the pictorial representation of partial occlusion. Consistent with previous research, the developmental change towards visual realism is mainly attributed to the decline of separate arrangements and the ensuing increase in the use of the HLE device (Chen & Holman, 1989; Cox, 1981; Freeman *et al.*, 1977). This type of shift takes place at the age of 5½, whereas partial occlusion achieves the status of the preferred modality at the age of 8.

Comparing these results with studies using equivalent materials, the use of partial occlusion in 3D copying tasks fluctuated from 0% to 19% up to the age of 6 (Arrowsmith *et al.*, 1994; Ashton, 1997; Chen & Holman, 1989; Cox, 1978; Smith & Campbell, 1987), and was not even well established (54%) by the age of 9 in mnemonic tasks (Freeman *et al.*, 1977). On the basis of these findings, we can confidently claim that the provision of a 2D model can activate covert competence in the accurate representation of an object partly concealed by an identical one at a younger age.

Another possibility that could have attributed to the facilitating effect of the present task is the absence of colour. As was ascertained by Bremner (1985a) and Light & Foot (1986), a number of 'copying-from-a-model' studies with low partial occlusion rates have used two identical, non-canonical objects, differentiated by colour and allowed to be drawn with coloured pencils (Cox, 1978 & 1981; Light & Humphreys, 1981; Light & Simmons, 1983). Their argument, fully supported by the present results, is that children given the opportunity to differentiate the two objects by colour are less likely to explore other more appropriate graphic means (height in the page, partial occlusion).

Turning to the alternative devices and to the transparency mode in particular, enclosure was surprising rare and primitive in the coding of depth relations. The variable placement of the further apple either within or around the near one implies that children are rather dubious about the clarity of their representational statement. On the other hand, overlap was a stable competitor to the other alternative modes from the age of 5½. These results are virtually the opposite from the ones obtained by Freeman *et al.*, (1977), using the same materials in a mnemonic task. Probably, providing children with a model might have facilitated a better conceptualisation of the scene.

What is thought to be rather interesting is that essentially all the children who used overlap in the present task claimed to have drawn the near apple. This has a significant implication. If children, when drawing a whole array, show a sequential preference involving the coding of the nearer object temporally earlier, this procedural strategy is in conflict with the representational logic of overlap which requires the nearer object to be drawn temporarily later. Thus, the low frequency of overlap in Freeman's study might have been the by-product of the procedural aspect of drawing. Asking children to complete rather than reproduce the whole array seems to have relaxed the conflict and subsequently promoted the use of overlap.

The arrangement of apples along the vertical-oblique axis in the majority of the sample shows that children can retain fidelity to the array-specific relations in this respect. Also, the distribution of responses essentially between diagonal and horizontal placements confirms the findings of Braine *et al.* (1993), and generally indicates that children could use the oblique drawing system (Hagen, 1986; Willats 1977) to code depth at a young age copying from a model and latter spontaneously drawing from memory. There is also an association between the ability to reproduce correctly the spatial alignment of the apples in the array and the ability to preserve their structural integration. Nearly all the children who used the HLE device to denote depth also used the vertical/oblique dimension, whereas among the alternative strategies, unified arrangements were more likely to be found with vertical-oblique alignment and segregated configurations with horizontal alignment.

Some researchers (Light & Humphreys, 1981; Teske *et al.*, 1992) have claimed that there is no way of knowing whether depictions are truly faithful to the array-specific relationships from the drawings of 'nonfronted' and identical objects. However, the identification task along with the recording of spatio-temporal drawing sequence allowed close scrutiny of the data. The finding that in 81% of the total drawings there was left-right and near-far orientation correspondence between the apple children decided to draw and its equivalent in the stimulus array further substantiates children's adherence to the model array[6]. As in the study of Freeman *et al.* (1977), 'errors of reversal' were made primarily by the youngest age group suggesting that approximately up to 5 years of age children's attention to the spatial properties of the context is more limited.

Finally, although the nature of the task (copying task, 2-D model, lead pencils) promoted the use of depth cues (spatial axis, HLE) at a younger age than previously reported, we cannot comfortably say that generally children demonstrated an apparent appreciation of the task requirements in the present task. There were a considerable percentage of children in the total sample (35%), or even in the oldest of the age groups (30%), who failed to accurately infer that the apple missing from their papers was the far one. One plausible explanation for this result might have been the degree of similarity between the items of the array. Perhaps the fact that the near and the far apple could be differentiated only in terms of their completeness might have impeded the accuracy of judgement. Since making the correct representational choice is mandatory for the use of the view-specific structural and spatial device, this proposition was investigated in the following study.

4. THE BOTTLES TASK

In the present task two variations were introduced in comparison to the previous one. Again, the array was made of two objects of the same kind, and children had to complete the unfinished array on their paper by drawing the further one. However, choosing objects with a discrete orientation facilitated differentiation on this attribute. Thus, the item at the front had a non-canonical orientation, whereas the item at the back had a canonical one. Since children are sensitive to orientation differences, they should be more likely to accurately infer the task

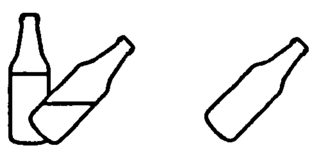
[6] Even for the children who chose to draw the front apple, in spite of the fact that their drawings were consequently less visually realistic representations of the stimulus array, the majority of them (82%) were drawn faithfully to the internal spatial relations of the forms within the array.

demands. If this is the precursor for visual realism, then the present task should elicit more partial occlusion output.

Further, height in the picture was a property of the scene in the Apples task, and the majority of children preserved it while encoding at the same time depth. In the present task the nearest of the items was shown at eye level. Thus, if children want to use the vertical dimension as a compromise solution to encode depth, then they should violate viewpoint-fidelity. Presenting children with a conflict, it is possible to examine the priorities children give to the representation of various information and find at what stage of development a shift in their representational precedence takes place.

4.1 Apparatus

Children were presented with a picture outline of a pair of bottles, 3.00cm tall and 1.00cm wide placed one behind the other in a straight line relative to children's viewpoint. The bottle at the back had a vertical orientation and was partially occluded by the near one which was tilted 45° degree from the vertical clockwise (Figure 5.5-a). Children were provided with the outline of the tilted bottle drawn on their paper (b) and were requested to look carefully at the model and draw the bottle that was missing from their paper to make their drawing look exactly as the one on the card. On completion of their drawing, children had to match each bottle on their paper with the corresponding one on the stimulus array.



a b

Figure 5.5 Schematic illustration of stimulus material

4.2 Results

4.2.1 Partial occlusion

Table 5.6 categorises the drawings into visually accurate ones, where children drew the partial occluded bottle by the successful implementation of the HLE device, and visually inaccurate ones where children produced complete forms. Over the entire age range, 30 children (28%) used the perspective-appropriate drawing device and 78 (72%) employed an inappropriate one. The distribution of drawing responses across the age levels revealed that the use of the appropriate device was age dependent [$\chi^2(3) = 40.02$, $p<.001$]. Specifically, children up to the age of 5½ never used partial occlusion –with only one exception–, whereas a sudden increase in its use emerged by the age of 7 and became established amongst the oldest children. Comparisons between successive age groups yielded a significant difference in performance only between the two intermediate age groups [5½ with 7 years old: $\chi^2(1) = 7.48$, $p<.01$, 7 with 8 years old: $\chi^2(1) = 3.12$, $p = .08$ with continuity correction].

Table 5.6 Distribution of drawing devices by age.

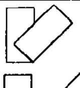
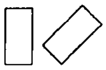
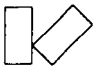
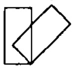
		Drawing device			
Age groups	N	Complete	HLE	*z*	*p*
4	28[a]	28	0	-5.29	<.001
$5^{1/2}$	25	24	1	-4.60	<.001
7	28	17	11	-1.13	*NS*
8	27	9	18	-1.73	<.05

a Two children produced unrepresentative drawings.

4.2.2 Structural drawing strategies

To identify possible developmental shifts in the use of drawing devices other than the HLE one, the non-partial occlusion drawings were subdivided into segregating arrangements, where the two bottles were spatially distant, and unification arrangements, where the bottles shared some common space. In the overall sample, drawings were equally distributed across the three structural patterns with partial occlusions accounting for the 28% of all drawings, segregations for the 43% and unifications for the 29%. Examining the distribution of drawing responses by age, a developmental sequence in the use of the three structural devices clearly emerged [$\chi^2(6) = 65.5$, $p<.001$]. The results are set out in Table 5.7 where the unified arrangements are sub-divided into attachment and overlap.

Table 5.7 Distribution of structural strategies by age.

Structural strategy		Age groups 4	5½	7	8	Total
Partial Occlusion		0	1	11	18	30
Segregation		26	14	5	2	47
Attachment		1	9	10	2	22
Overlap		1	1	2	5	9
Total		28	25	28	27	108

Note *The horizontal alignment is used as an example for the illustration of structural devices.*

Spatial separations prevailed in the drawings of 4 year olds whereas those of 5½ year olds were mainly distributed between segregated and unified configurations (14 *vs.* 10). Thus, although the first two age groups were equally ineffective in omitting the occluded part of the further bottle, they varied in the treatment of the structural properties of the scene. By the age of 7, children were still guided by a binary choice with unified arrangements remaining a stable competitor to another alternative device which now was replace by partial occlusion (12 *vs.* 11). Children of 8 years of age seem to have resolved this bimodality and demonstrated a clear preference towards the use of the HLE device. On the whole, the results from Table 5.7 suggest that the tendency to spatially separate the forms gradually decreases with age, while the use of partial occlusion increases with a crossover at the age of 7. With respect to the two unification strategies, attachment was preferred to overlap and was mainly used by the two intermediate age groups, while overlap was at low frequency with a modest rise among the oldest children.

Due to the low incidence of visually accurate drawings, children's placement patterns were further explored. Thus, the relative distance between the bottles was measured. With this analysis the degree of the between-forms spatial overlap was examined independently of the specific structural devices and regardless of the degree of visual accuracy of children's compositions. The inter-object distance between the two closest points of bottles' contours was recorded in all drawings with values ranging from positive to negative. A positive value indicates that the occluded upright bottle was drawn spatially divergent from the tilted one while a negative value suggests spatial convergence between the forms with zero denoting attachment at a point. Since the projective overlap of the two bottles was 5.00 mm edge to edge, the magnitude of deviation of the obtained values from the expected one could be readily detected.

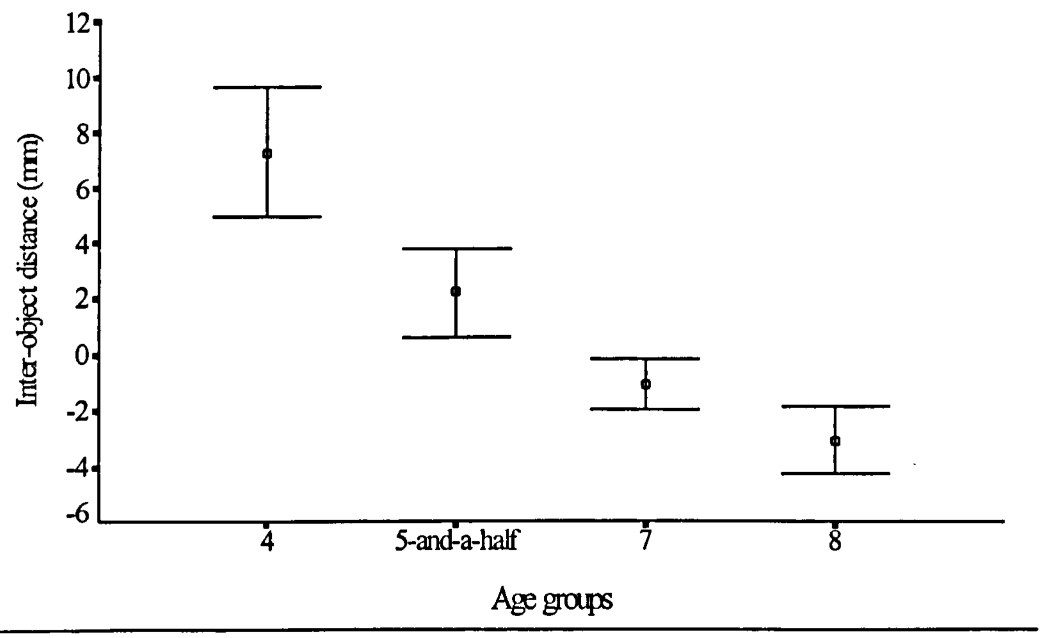


Figure 5.6 Mean inter-object distance by age level

The Figure 5.6 presents the mean of inter-object distance by age. The data suggests that with age the placement of the requested bottle gradually approaches the optimum distance. The analysis of variance yielded a significant effect of age ($F_{3,107} = 33.7$, $p<.001$) and the follow-up study revealed reliable differences in the means between all age groups except among the two older ones (Newman-Keuls post-hoc analysis of difference, $p<.05$). Therefore, there was a developmental progression not only in the use of the appropriate depth device but also in the depiction of the internal structural properties of the array (spatial overlap).

4.2.3 Spatial alignment patterns

The ability of children to represent view-specific information was next examined with respect to the spatial axes they used to align the far bottle. Since on the stimulus picture of the array the far bottle was drawn adjacent and at the same plane with the near one[7] (see Figure 5.5), children should choose the horizontal alignment if they were to preserve both array-specific and view-specific information. In fact, the majority of children in the total sample demonstrated a clear preference towards the horizontal alignment (z = 3.17, p<.01, two-tailed) as 71 (66%) arranged the missing bottle next to the pre-drawn one, and 37 (34%) drew it along the vertical dimension.

A developmental shift was also obtained in the use of the appropriate alignment mode. These data are presented in Table 5.8. Comparing the total number of horizontal and vertical output at each age level, it seems that children up to the age of 5½ had no representational preference, while from the age of 7 they mostly used the appropriate alignment strategy [$\chi^2(3) = 15.76, p<.01$]. Freeman (Freeman *et al.*, 1977) identified two graphic conventions for the pictorial representation of an in-depth relation in a two-item array; the structural device of partial occlusion, which deals with the between-items spatial relation, and the placement device of vertical alignment, which tackles the relation between the objects and the context. To investigate whether children used both conventions concurrently or whether there is a lawful relation between the alignment patterns and the structural devices, horizontal and

[7] One of the edges of the interrupted contour of the vertical bottle at the back is anchored to the lower bottom corner of the tilted one at the front. So if the near bottle was to be rotated to its canonical orientation, the far bottle would have been totally occluded

vertical drawings were re-examined on the basis of the strategy children used to deal with the concealed information of the occluded bottle.

Table 5.8 Distribution of spatial alignment patterns by structural device and age.

	Horizontal				Vertical			
Age groups	Occlusion	Separate	Unified	Total	Occlusion	Separate	Unified	Total
4	0	11	0	11	0	15	2	17
5½	0	8	7	15	1	6	3	10
7	9	3	12	24	2	2	0	4
8	16	1	4	21	2	1	3	6
Total	25	23	23	71	5	24	8	37

As Table 5.8 (bottom row) shows, the relative frequency of the three structural devices appears to be dependent on the type of spatial alignment [$\chi^2(2) = 11.00, p<.01$][8]. What the data suggests is that the use of a strategy that integrates information about the spatial overlap of the forms –no matter whether it is the least (unification) or the most visually accurate (partial occlusion)– renders the use of verticality as a depth cue redundant. On the contrary, when children are using the least informative depiction strategy, namely the spatial segregation, half of them give representational precedence to the fact that both rest on the same horizontal surface and separate the bottles horizontally, while the other half use height in the picture plane to denote the in-depth relationship and separate the bottles vertically. From Table 5.8, it seems that the use of horizontal and vertical separation is not age-related due to the fact that from the age of 7 the use of separate arrangements declines dramatically.

4.2.4 Identification Task

Nearly all children (93 out of 108, 86%) inferred accurately the task demands and claimed that they were drawing the far bottle. Only a small proportion (5%) of 4 year olds inappropriately drew the far bottle to the right of the pre-drawn one. Choosing to draw the near bottle was found primarily among the 4 year olds (10 out of 15, 67%). However, these

[8] Same results were obtained when the analysis was performed on the four structural devices with unification strategy broken down to attachment and enclosure, $\chi^2(3) = 12.95, p<.01$.

few children drew the incorrect bottle with fidelity to its properties, representing it tilted and placing it below and to the right of the pre-drawn one.

4.3 Discussion

Consistent with our prediction, the orientation difference of the bottles within the array forces children to attend carefully to the items and interpret accurately the task requirements. This is mainly true for the children of 5½ years of age and above, as 4 year olds appear impervious to this manipulation. Either these children are insensitive to orientation differences, or their deductive abilities are too limited, resulting in simply drawing the item closer to them.

However, contrary to our prediction, helping children to understand that they have to draw the far bottle doesn't foster visual realism. The partial occlusion output is suppressed in the overall sample from 46% in the Apples task to 28% in the Bottles task, and the developmental change in the use from complete to partial forms is displaced from the age of 5½ years in the former, to the age of 7 in the latter task. In fact, the present task hinders most the performance of 5½ year olds if one considers that 44% passed the Apples task while only 4% of the same children succeeded with the bottles.

To unravel the underlying factors responsible for the decline of partial occlusion responses, we should turn to the distribution of alternative strategies. Two crucial observations can be made. First, the Bottle scene promoted segregation, which rose from 23% in the Apples tasks to 43.5% in the Bottles tasks, and second it caused the substitution of overlap (20% in Apples task) by attachment (20% in Bottles) in the whole sample. What these results imply is that the interplay between partial occlusion and segregation denotes that in essence the tendency to separate the items is the only true competitor to the use of the HLE device. Second, children make different strategic choices determined by the structural encoding of each particular scene.

Looking at the performance difference within each age group, the present task eradicated children's competence to resist segregation at intermediate stages of development (5½ & 7 years of age)[9]. So, it seems that the youngest children have limited abilities to radically change their approach across tasks and the oldest have well-established skills to apply them consistently. On the contrary, transitional children are more susceptible to the structural properties of each scene and are sufficiently flexible to alter their representational priorities.

Two features in the present scene might have led children to encode it as 'two distinct objects separated by distance'. The first is the degree of spatial occlusion between the forms and the second their orientation difference. It is only reasonable to expect that the smaller the area concealed by the object at the front the more likely is for children to perceive the further object as integral. Since the projective overlap is smaller in the Bottles than in the Apples task, children might be more compelled to retain the completeness of the further object in the former rather than in the latter task. Besides, children might have not seen the need to distinguish the bottles in terms of their spatial relationship (and formal integrity) since they could do so in terms of their orientation. To take the argument a little further, knowing that children are relational coders (Bryant, 1974; Freeman 1980, chapter 6) and also responsive to visual contrast (Bremner, 1985a, Cox, 1985; Davis, 1983 & 1985a; Davis & Bentley, 1984), it might have been more likely to have noticed (or consider relevant for representation) the more pronounced orientation difference between the items than their formal difference. Turning to the unification strategies, the substitution of the transparency mode for attachment can be also explained by the decrease of the overlapping space between the bottles in

[9] Five and a half year olds produced twice as much separate arrangements in the Bottles task compared to the Apples task, whereas for the 7 year olds segregations raised from 0% to 18%.

comparison to the apples. Another contributing factor is also the fact that here the children were less likely to draw the item at the front which consequently reduces the possibility of superimposing it on the further.

Finally, considering the spatial axes children used to place the forms, the majority of them retain fidelity to their view of the array, aligning them on the same horizontal surface. The vertical dimension is called forth only by the younger children (4 & 5½), lacking the skill to represent the depth relation by the HLE device. Half of them gave representational precedence to the fact that one bottle was behind another using height, whereas the other half preferred to show that both bottles rest on same surface. This distribution of responses also exemplifies the presence of a conflict, which is resolved by older children who suppress verticality as visually non-veridical, and turn to partial occlusion as the only appropriate cue to denote depth. Further, the shift of the alignment strategy from vertical-oblique, in the Apples tasks, to horizontal at the present one suggests that children's devices are dependent on the structure of the scene and are not generic representational modalities. Empirical evidence from the Apples and Bottles tasks also suggests that the use of the appropriate alignment pattern becomes well established by the age of 7.

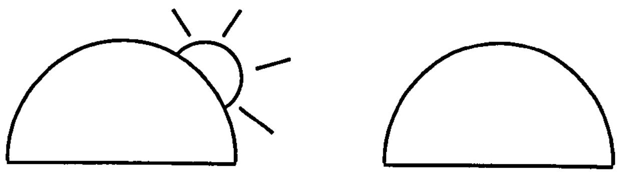
5. THE SUNSET TASK

Up to now, consistent with previous research, it can be claimed that, even under some facilitating conditions, it is still difficult to trigger visually realistic responses in children younger than 6 years of age (Barret *et al.*, 1985; Cox, 1986, chapter 4; Davis 1984; Light & Simmons, 1983; Radkey & Enns, 1987). Many researchers, approaching children's drawings from a socio-cognitive perspective, maintain that the paucity of partial occlusions in the younger age groups results from their limited comprehension of what they are expected to communicate. A number of studies (Arrowsmith *et al.*, 1994; Cox, 1981 study 5; Cox, 1986, p.75-83), carried out under this conception, found that the more richly contextualised tasks are more successful in eliciting partial occlusions than the 'standard' tasks. Cox (1981) endorsed the view that the notion of 'human sense' and the idea of hiding, conveyed by the task materials, cause the children to attend more and subsequently to represent realistically the appearance of the scene.

All the previous studies investigating the effect of the hiding notion used a human character behind an occluder (Arrowsmith *et al.*, 1994; Ashton, 1997; Cox, 1981, study 5; Cox, 1986, p.75-83; Light & Foot, 1986). If it is assumed that the human form has special significance for children, it could have contributed to the readability of the scene, limiting the generality of the hiding effect. Further, in some studies, the hiding scenario was explicitly stated at the preamble, whereas in all the studies, as Light & Foot (1986) first noticed, the items of the array had a canonical orientation (i.e. man-wall) and were different on a variety of attributes; features which might have inadvertently boosted the propitious effect of the hiding notion (see Chapter IV, section 6.2.2 & 6.2.3). The following task was designed to assess the merit of Cox's interpretation, controlling for the aforementioned factors. Consequently, the idea of hiding was inherent in the scene, but it was not explicitly stated during instructions, whereas the array neither involved an occluded human figure nor structurally distinctively different items.

5.1 Apparatus

Children were presented with a line drawing picture (Figure 5.7-a), depicting a mountain with a closed semicircle, 6.00 cm in diameter and 3.50 cm high from the base to its tip. A half-hidden sun, 2.00 cm in diameter, was placed diagonally at its right slope. Children were provided with the outline of the mountain (b) and were asked to copy what was missing from their paper to make their drawing look exactly as the one on the stimulus card.



a b

Figure 5.7 Schematic illustration of stimulus material

5.2 Results

5.2.1 Partial Occlusion

The drawings were classified into partial occlusions, where children used the HLE device, and into non-occlusion ones, where they depicted the sun with a complete outline. Three strategies were identified among the non-occlusion drawings; segregation, attachment and overlap. Table 5.9 presents the number of drawings in each drawing category by age level.

Table 5.9 Distribution of drawing devices by age levels.

		Age groups				
Strategies		4	5½	7	8	Total
P. Occlusion (HLE)		10	23	27	27	87
Non-Occlusions						
Segregation		9	0	0	0	9
Attachment		2	2	1	0	5
Overlap		9	0	0	0	9
Total (complete)		20	2	1	0	23

Over the entire age range, drawings of partial occlusion prevail over non-occlusion ones, as 87 (79%) children successfully omitted the non-visible part of the sun, whereas only 23 (21%) drew the sun with a complete outline. The analysis of visually accurate and inaccurate drawings by age showed that there was an abrupt increase in partial occlusion output from the age of 4 to 5½ [$\chi^2(1) = 17.19$, $p<.001$ with continuity correction], with performance stabilising from this age onwards, as partial occlusion became the dominant depiction device.

Also, although the use of complete contours occurred mainly among the 4 year olds, its prevalence is not well-established due to a noticeable incidence in the use of the HLE device amongst the youngest children (binomial test, $z = 1.82, p = .07$ two-tailed, $p = .03$ one tailed). Turning to the non-occlusion strategies, four year olds oscillated between segregating and unified configurations (9 *vs*. 11 respectively), with overlap being the modal device among the latter ones. On the other hand, the three children who used complete contours in the intermediate age groups tended to avoid the least informative strategy of segregation and the ambiguous use of overlap, and all used attachment.

A number of researchers (e.g. Cox, 1985; Ingram & Butterworth, 1989; Willats, 1977), elucidating the scoring criteria for a partial occlusion drawing, underscore the importance of the two contours making contact at appropriate occluding points. A closer examination of the present data revealed that there were certain cases which were defined as incomplete partial occlusions, where the 'ends' of sun's visible contour were not firmly anchored on the mountain's outline. Inaccurate placement resulted in separated occlusions, with the visible part of the sun appearing as free floating, and interposed occlusions with the visible contour breaking the mountain's surface[10].

Of the total number of partial occlusions, only 12 drawings (14%) fell into the class of incomplete occlusions. Half of them were made by the 4 year olds and four by the 5½ year olds, with equal numbers of separated and interposed occlusions in each age group, while the two older groups made one separated occlusion each. This data indicate that there was no difference in the relative frequency of the two types of incomplete occlusions which were rather scarce overall. It also suggests that their occurrence is mainly attributed to the conspicuous increase of partial occlusion output in the younger children whose dexterity is less developed, and that children with age eventually gain mastery of all the aspects of a partial occlusion configuration.

In the previous two tasks (Apples and Bottles), the drawings were also examined in terms of the spatial axes children used in aligning the requested form to the pre-drawn one. This type of analysis was not pertinent to the present study since in all cases children drew the sun considerably higher (diagonally) from the horizontal base of the mountain.

5.3 Discussion

Summarising the main results from this task, the performance of the 4 year olds appears somewhat varied with a considerable increase of partial occlusion output compared to the previous tasks, and comparable numbers of separate and overlapping arrangements. No developmental pattern was obtained beyond this age level as partial occlusion becomes the dominant device for representing the in-depth relation. The striking performance difference of children younger than the age of 8 in comparison with the previous tasks is in line with the socio-cognitive approach. Children are not rigidly fixed to certain graphic strategies but can be persuaded to change their depiction bias to convey information pertinent to a given situation (communicative purpose: Light & Simmons, 1983; visual contrast: Davis 1985a; hiding notion: Cox, 1981). The present task, by contextualising the idea of partial occlusion, made children to appreciate its relevance and, consistent with Crook's cognitive account (1985), to change the way they structurally encoded the scene.

The study may perhaps be regarded as a test of Cox's hypothesis (1986, p.82) that the presence of a human character has privileged status in conveying the idea of hiding or adding contextual salience in partial occlusion scenes. The present results revealed that it is not the

---

[10] Lenient scoring criteria were adopted for both categories of incomplete occlusions in order not to

identity of the occluded object but the familiarity of the scene and the relevance of the hidden-features omission, inherent in the scene, which can trigger partial occlusion rates, even higher than less convincing manlike scene [man-block: Cox & Simm, cited in Cox 1986, p.82; boy-girl: Ashton, 1997).

The lack of verbal reference to the idea of 'hiding' confirms the findings of Light & Foot (1986) and suggests that the social intelligibility of the materials renders the explicitness of the instructions redundant in eliciting partial occlusions. However, contrary to their claim, the dissimilarity between the forms, used previously in the hiding scenes, cannot be held totally liable for the visually realistic drawings. It was found that in the absence of a compelling 'hiding' story, the presentation of a pair of different objects (salient or neutral) with canonical orientations and asymmetrical features can promote visual realism form the age of 6 (Cox, 1986; Light & Foot, 1986). However, in the present scene, although the two items had a less pronounced dissimilar structure (sun-mountain: two semi-circles varying in size), they elicited partial occlusion from the age of 4.

Finally, a result that is worth mentioning is that, with respect to the two unification strategies, overlap was preferred to attachment (markedly among the youngest of children). This finding seems to support the argument in the discussion section of Bottles tasks that the greater the visual overlap between the forms the more likely for the child to produce a transparency. The pattern of fluctuation in the incidence of overlap is consistent with the structure of the scenes; higher in the Apples and Sunset task, where half of the further item was concealed, and lower in the Bottles task where its hidden part was much reduced.

---

penalise pencil slips and poor muscular control. For drawings to score in these categories, the occluded edges should either cross mountain's outline or depart from it by more than 3.00 mm.

6. THE BALLS TASKS

So far, the occluding form was always pre-drawn on children's papers. This design was employed to evaluate readily the impact of conceptual knowledge on the graphic output when technical competence and organisational skills are reduced. In the remaining tasks, the demands on planning ability and graphic skill were increased by asking children to draw the whole array of a pair of identical objects.

Besides the fact that this type of stimulus material was most often used by previous research so as to allow direct comparisons, it has been suggested that particularly these arrays are more likely to cause planning problems and subsequently impede visual realism. The argument is that they promote group-encoding by similarity (Cox, 1986, p.80-81; Morra *et al.*, 1996). Given that young children have difficulty analysing a configuration into its constituent parts (Shepp & Schwartz, 1976; Smith & Kemler, 1977), when they have to reproduce it as a whole, they are more likely to call forth one figurative scheme; that of the complete object and simply repeat it twice. By keeping the materials operatively equivalent with the previous tasks where the occluding item was provided (Apples, Bottles), the merit of the previous claim and the impact of production factors could be assessed at different age levels.

Second, the production patterns often reported in occlusion research (Cox, 1981, study 1; Light & Humphreys, 1981; Ingram & Butterworth, 1989; Van Sommer, 1984, p. 106-114) were also examined. It was of interest to see how dominant these sequential regularities are at different age levels, and if there is an interplay between the graphic devices – originated from the final spatial arrangement of items – and the spatio-temporal strategies – obtained from the order of production of the nearer and further item. Finally, in the following task children were tested under two conditions where the degree of occlusion between the items was systematically manipulated. Given that there was some indication from the previous studies that the degree of inter-item overlap can influence children's drawing strategies, it was expected that enhancing the degree of concealment of the further form in one condition should promote the omission of its hidden part [11].

6.1 Apparatus and procedure

Children were presented with two line drawing cards, each depicting a pair of vertically arranged balls, 3.00 cm in diameter, placed one behind the other (Figure 5.8). The stimulus arrays varied in the degree of partial occlusion. In one card the front ball hid half of the rear ball (medium occlusion, Figure 5.8-A) and in the other it concealed more than two thirds of its shape (fine occlusion, Figure 5.8-B).

A. Medium occlusion (MO) B. Fine occlusion (FO)

Figure 5.8 Stimulus arrays with varied degrees of partial occlusion.

Children were presented with one card at a time. Due to the similarity of the two tasks, the order of presentation was counterbalanced between children at each age level; half children were given first the card A followed by the card B, and the other half was give the tasks in the reverse order. The preamble and the instructions given were directing children's attention to the drawings of the two balls with no reference to their spatial arrangement, and were prompting them to make a drawing that would look exactly like the one on the card. As children were drawing, a note was taken of the temporal order used. On completion of each drawing, children were asked to match the forms on their paper with the equivalent ones on

[11] This manipulation would increase between-forms differentiation and visual contrast, and subsequently affect the way children encode the structure of the array.

the stimulus card. This procedure allowed the referent of the forms in the drawings to be checked, and made it possible subsequently to assess the spatial accuracy of the configuration and examine the temporal sequence of the drawing.

6.2 Results

6.2.1 Partial occlusion

All the drawings fell into classes of identifiable devices, except of one drawing with the MO condition and two with the FO one, all of which were made by 4 year old children. Using the same generic criterion as in the previous studies, the drawings were first divided into those depicting only the visible part of the further ball with the HLE drawing device and into those representing the ball with a complete contour. There was also one drawing of total occlusion in each of the two conditions where the further ball was completely deleted in the graphic output, and it has been eliminated at this stage of analysis. Table 5.10 presents the frequency distribution of drawing devices broken down by age level and condition.

Table 5.10 Frequency distribution of drawing devices by age and condition.

		Drawing devices				
Age groups	Cond[1]	Complete	HLE	Total	z	P^2
4	MO	22	6	28	–3.02	<.01
	FO	16	11	27	–0.96	*NS*
$5^{1/2}$	MO	8	17	25	–1.80	*NS*[3]
	FO	4	21	25	–3.40	<.001
7	MO	4	24	28	–3.78	<.001
	FO	0	28	28	–5.29	<.001
8	MO	1	26	27	–4.81	<.001
	FO	0	27	27	–5.19	<.001
Total	MO	35	73	108		
	FO	20	87	107		

Note 1 Cond = Condition, MO = medium occlusion task; FO = fine occlusion task
Note 2 The probability values given are for a two-tailed test
Note 3 The difference is significant at p<.05 for an one-tailed test.

Over the entire age range, for the MO condition, 35 (32%) children failed to produce partial occlusion, whereas 73 (68%) succeeded, and for the FO condition 20 (18.5%) children produced complete forms, while 87 (80.5%) drew only the visible part of ball's contour. Chi-square tests were used to compare the distribution of the complete and incomplete drawings by children of different ages. The results revealed that the ability to omit the concealed properties of an object was age dependent [MO $\chi^2(3) = 41.59$, FO $\chi^2(3) = 42.0$, $p<.001$].

Binomial tests were used to examine the pattern of device dominance at each age level. Except of the 4 year olds, the use of the HLE device was well-established in all age levels for both conditions with no substantial variation in its relative frequency between these older groups. On the contrary four year olds performed significantly differently from the older groups since they showed a representational bias towards complete forms in the MO condition and no representational preference in the FO one [comparison between the first two age groups, MO: $\chi^2(1) = 9.84$, FO: $\chi^2(1) = 8.52$, both at $p<.01$ with continuity correction]. The pattern of results was essentially the same in both conditions, apart from the fact that, for all age groups, HLE was more common in the FO task regardless of the order of presentation which had no significant effect [χ^2 (1) = 1.13, p>.05].

To investigate further any possible task effect, we compared the overall performance of children at both tasks and within each age group. In general, children were very consistent in either using a complete contour or the HLE device across conditions. Over the total number of children who produced representational drawings in both conditions, 89 (83%) children used a single representational style, while 18 (17%) changed their drawing strategy. The direction of change among the inconsistent responders revealed a reliable pattern, as 16 children used the more advanced drawing strategy for the FO condition and only 2 for the MO one [McNemar's test for change $\chi^2(1) = 9.39$, $p<.01$]. When the changes in drawing style were considered in each age group, the pattern remained the same, but it was not statistically reliable due to the small number of shifts in response at each level[12].

In general, the analysis of the present results suggests that the use of the HLE device is achieved at the age of 5½ and that children drew similarly in the two conditions. However, when children change their drawing style, they use more often the perspective-appropriate device with the FO task than with the MO one. This result suggests that in an array of two similar objects, the smaller the visible part of the occluded object is, the more likely is for children to omit its concealed portion. On the contrary, the more exposed the structure of the further object is, the more likely is to be drawn as complete. This issue will be explored in the following section where the drawing devices will be examined in more detail.

6.2.2 Structural drawing devices

The drawings portraying the further ball with a complete outline were subdivided as separate and unified on the basis of its proximity to the near one. Figure 5.9 plots the relative frequency of separate, unified and partial occlusion drawings by age and task. The general pattern of results for the two conditions is that at the age of 4 children produce equivalent numbers of drawings with the three structural devices. The frequency of the representational styles changes abruptly at the age of 5½ where we observe a steep decline in the use of spatial separation and a smoother decrease in the use of spatial unification with a corresponding gradual increase in the partial occlusion output [MO: $\chi^2(2) = 13.16$, FO: $\chi^2(2) = 10.32$, both at $p<.01$]. From the age of 7, separate arrangements cease to occur and performance stabilises as children produce mainly visually accurate drawings. Over the present age range, the developmental progression in the use of the perspective-appropriate

[12] For all cases of inconsistency, performance on FO was better than on MO task within age level: 7 vs. 2; 4 vs. 0; 4 vs. 0 and 1 vs. 0 for each age group respectively.

graphic device involves a linear decrease in unified arrangements which gives way to a steady increase in the incidence of partial occlusions.

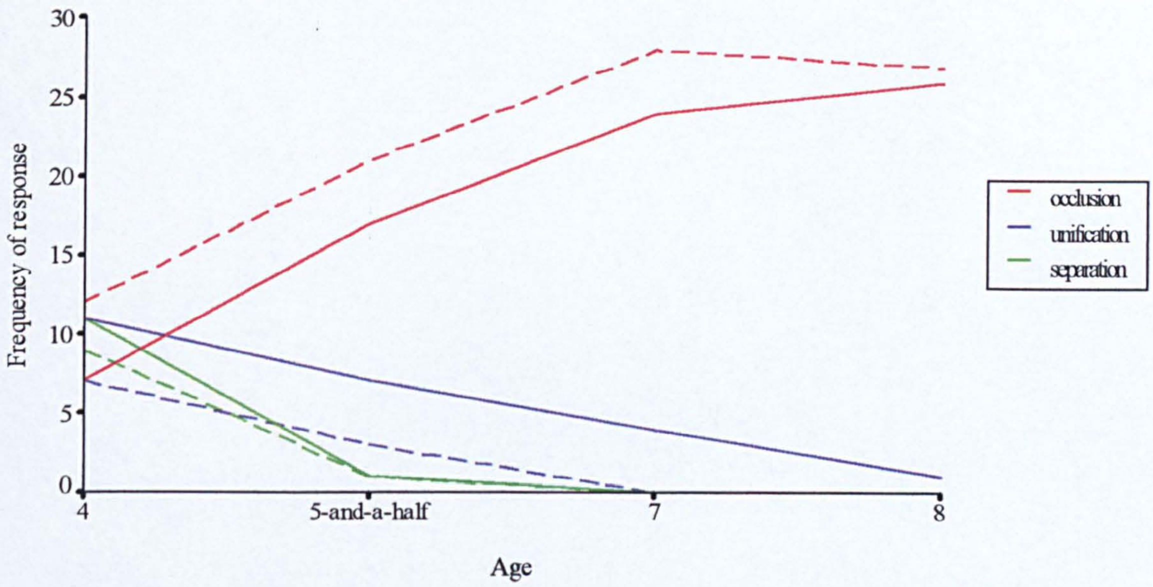


Figure 5.9 Age-related shifts in relative frequency of structural devices of one ball behind another. Key: Medium occlusion —, Fine occlusion ----

The data in Figure 5.9 also elaborates the task effect in performance, found in the analysis of results at the previous section. While the FO task suppresses the use of inaccurate arrangements, particularly the unified ones, fostering at the same time the use of the HLE device, the exact reverse pattern applies to the MO task at all age levels. Thus, the spatial overlap between the two balls in the FO task prompts the younger of the children to abandon the unification device in favour of partial occlusion.

For a more detailed account of the structural devices, the unified arrangements were subdivided into the drawing categories of attachment, overlap and enclosure. Due to the occurrence of a substantial number of drawings that did not comfortably meet all the criteria of partial occlusion, the drawings using the HLE device were also differentiated into those of partial occlusion and separated occlusion if the incomplete form of the further ball was separated from the occluding ball. Table 5.11 presents the distribution of the drawing strategies across age and condition.

Table 5.11 Distribution of drawing strategies by age group and condition.

Strategies		Cond	Age groups 4	5½	7	8	Total
Occlusions							
Total		MO	1	0	0	0	1
		FO	1	0	0	0	1
Partial		MO	5	14	24	26	69
		FO	5	16	23	26	70
Separated		MO	1	3	0	0	4
		FO	6	5	5	1	17
Total		MO	7	17	24	26	74
		FO	12	21	28	27	88
Non-occlusions							
Segregation		MO	11	1	0	0	12
		FO	9	1	0	0	10
Attachment		MO	4	2	1	0	7
		FO	3	1	0	0	4
Overlap		MO	7	5	3	1	16
		FO	2	2	0	0	4
Enclosure		MO	0	0	0	0	0
		FO	2	0	0	0	2
Total (complete)		MO	22	8	4	1	35
		FO	16	4	0	0	20

Notes *Cond = Condition, MO = Medium occlusion task, FO = Fine occlusion task.*
The vertical alignment is used as an example for the illustration of structural devices.

All children drew two balls except for one 4 year old child who produced a total occlusion drawing by completely omitting the further ball in both tasks. The data in Table 5.11 can be summarised as follows. Turning to the non-occlusion drawings, as it has been already established, segregation had a low frequency of occurrence in the overall sample for both conditions, and its use was age-related as it was abandoned at the age of 5½. In total, unified arrangements were found more often in the MO task (21%) than in the FO one (9%). Due to their rare incidence in the latter condition and their predominate use by the 4 year olds (7 out of 10), there is neither a reliable developmental shift nor an apparent preference in the use of any of them. On the contrary, the relative frequency of different unified devices varied in the

MO task. Children seem to prefer overlap to attachment (16 *vs* 7 respectively), and this pattern, albeit no statistically reliable, was consistent within all age levels, whereas drawings of enclosure never occurred.

Task differences were also observed in the distribution of partial and separated occlusions. Among the children who used the HLE device for the depiction of the further ball, only 4 children of a younger age (5%) drew it distant from the contour of the near one in the MO task. On the contrary, the number of separated occlusions increased to 17 (19%) in the FO task and were of comparable frequency up to the age of 7 years. Thus, although the majority of children who were able to use the HLE device as a depth cue were able to apply it faithfully to the view-specific properties of the array, the FO task increased the incidence of separated occlusions relative to the MO task.

Concluding, partial occlusion was the dominant strategy, with a clear developmental pattern, whereas overlap and separated occlusion constitute the second most often used device for the MO and FO task respectively. However, because they were few in number in the overall sample, no age-related trends could be revealed, and their use reflect more local representational decisions on the basis of the properties specific to the array in display. Finally, the separate arrangement of two complete forms was the third relatively common strategy in the overall sample for both tasks, but was exclusively restricted to the 4 year olds.

6.2.3 Spatial alignment patterns

In a further analysis, the data was examined on the basis of the relative spatial position of balls. All drawings were classified into vertical and horizontal arrangements except of the two enclosure configurations of the nursery children in the FO task which were excluded from this analysis due to their ambiguity. Horizontal arrangements were negligible, and therefore there was no developmental change in the distribution of the two spatial placement patterns. Over the entire age range, verticals constituted 93% and horizontals 7% of all the drawings in the MO task, whereas for the FO task the distribution was 95% and 5% in respect. All horizontal arrangements, except one, were produced by the 4 years old, yet the prevalence of verticals indicates that the use of the vertical dimension to graphically denote depth is well established even in children as young as 4 years of age.

The low frequency of horizontal alignment of balls did not allow the systematic examination of the distribution of alignment patterns by the type of structural device used. Besides, the data can be described as follows; in the FO task, 4 year olds made two horizontal alignments with partial occlusion and three with separation, while all unified arrangements occurred with vertical alignments. In the MO task, the modal arrangement of partial occlusions was the vertical in all the drawings, while horizontal arrangements were equally divided between separate and unified compositions.

6.2.4 Sequence of production

Horizontals and verticals were also analysed in terms of the temporal and spatial ordering used[13]. Having recorded the drawing sequence and pinpointed the near and far ball in children's drawings from their matching response, the spatial position of the forms on the paper was examined in relation to the temporal order used in the sequence of drawing. Eight children of the twelve who produced horizontals for both tasks drew the near ball first with no apparent left-right sequential bias in the way the front and back ball were drawn[14].

When the spatio-temporal order of drawings was considered for the vertical arrangements, children demonstrated structural accuracy and a preferred order in which they drew the forms

[13] Considering the two enclosure arrangements obtained in the FO task only, they were constructed in the reverse way. One child made the further ball first and drew the outline of the near ball around it, whereas the other child drew the near ball first and enclosed it within the further ball.

(Table 5.12). First, for the vast majority of vertical drawings, the spatial arrangement of the balls corresponded with that of the scene, depicting the nearer object in the front plane and the further ball in the upper plane (MO: 96, FO: 97). Only eight drawings (MO: 5, FO: 3) showed a front-back inversion of which seven were produced by the younger age group. In general, all the children, using the vertical dimension to encode depth, produced drawings displaying orientation correspondence with the forms of the stimulus array.

Table 5.12 Distribution of vertical drawings by age and by the relative location of object drawn first across conditions.

	Conditions			
	MO		FO	
	First drawn ball		First drawn ball	
Age groups	Near	Far	Near	Far
4	12 (3)	5 (2)	12	6 (2)
$5^{1/2}$	12	12	13	11 (1)
7	20	8	22	6
8	17	10	20	7
Total	61 (3)	35 (2)	67	30 (3)

Within each cell, the figure to the left denotes accurate reproduction of the spatial position of the balls in the array (front ball-lower, back ball-further) and the one in parenthesis represents a reversal (front ball-higher, back ball-lower). MO: Medium occlusion, FO: Fine occlusion.

In addition, vertical drawings –similarly to what the small proportion of the horizontal drawings showed– revealed a reliable association between the ball that was first drawn and its spatial position within the stimulus array. Over the total number of verticals, children employed the 'nearer first' rule, drawing more often the front ball and working in a bottom→top drawing sequence [binomial one-tailed test MO: $z = 2.59, p<.01$, FO: $z = 3.30, p<.001$]. As the data in Table 5.12 suggests, although only the 5½ year olds were equally likely to start their drawing with the near as well as with the further ball, there was no age-related trend [MO: $\chi^2(3) = 2.85$, FO: $\chi^2(3) = 5.29$, both $p>.05$] or task difference in the preferred order by which the objects were drawn. This indicates that it is the spatial position

[14] In the MO task, the occluding ball was placed to the left in two drawings and to the right in five drawings. In the FO task, the occluding ball was placed to the left three times and to the right two.

of the objects within an array that determines the temporal order by which they will be drawn.

The spatio-temporal order of the vertical drawings is inconsistent with results from studies of executive constraints (Glenn *et al.*, 1995; Goodnow, 1977, pp.59-65; Goodnow & Levine, 1973; Kirk, 1981; Pemberton, 1985; Simmer, 1981; Wallon & Baudoin, 1990). A number of researchers have identified the so-called starting position principle which prescribes graphic execution from the top to the bottom of the page. Considering that a drawing of partial occlusion will be more likely to be constructed starting first with the occluding ball and fitting around it the elliptical form of the further ball, the order of construction was examined by the drawing devices used. This analysis will make it possible to determine whether the preference for drawing front to back is guided by representational forces or from formal constraints underlying the graphic devices. Table 5.13 presents the spatial position of first drawn form in the array by the three structural strategies.

Table 5.13 Order of construction by strategy within condition.

	Conditions			
	MO		FO	
	First drawn ball		First drawn ball	
Strategies	Near	Far	Near	Far
Partial occlusion	51	22	63	24
Segregation	7	5	3	7
Unification	11	12	5	5
Total	69	39	71	36

Notes *MO = Medium occlusion, FO = Fine occlusion.*
Drawings of total occlusion were naturally excluded.

Only partial occlusion drawings demonstrated a fixed sequential pattern in the ordering of the near and further ball. The majority of children started the drawing with the item at the front instead of the one at the back, avoiding anticipated embedding which demands more pre-planning and motor control. On the contrary, children producing segregated and unified configurations displayed no starting preference in the drawing sequence of the near and further ball. This pattern of results was practically the same when the same analysis was confined to the drawings with vertical alignment[15]. Thus, it seems that working from the foreground form first in representing pictorially an in-depth relationship does not necessarily entail that the front object has representational precedence in the order of array's construction. When the drawing strategy obliterates underlying executive constraints (as in segregation and unification), no preference for drawing the near ball first can be observed, regardless of the alignment pattern used.

6.3 Discussion

The most striking feature of this study is the high incidence of partial occlusion rates in the overall sample and within each age level, in both conditions. Visually accurate drawings were generated even at the age of 4, whereas from the age of 5½ partial occlusion became the dominant drawing strategy for the representation of the in-depth relation. Previous studies, testing children with a model array of two balls along with a variety of facilitating procedures, (explicit instructions: Chen & Holman, 1989; Cox, 1978 & 1981, study 2, communicative context: Light & Simmons 1983; Smith & Campbell, 1987, viewing conditions: Morra *et al.*, 1996) yielded substantially lower output of view-specific drawings, particular from the younger children.

The obvious difference, which this inconsistency can be attributed to, lies in the features of the array. While in the previous studies children had to work from a 3D model, where the balls were differentiated by colour, and were allowed to draw them with colour pencils, in the present one they had to copy a 2D, black and white line-drawing. According to the production deficit hypothesis (see, Freeman, 1980), children have difficulty finding the

[15] For the MO task, the number of times the near ball was drawn first compared to the far ball is: Partial occlusion 51/22, segregation 5/3, unification 8/12. The frequency distributions for the FO task are 61/24, 2/5 and 4/4 respectively.

graphic means to convey a depth relation in a 3D scene on the 2D picture surface. Thus, providing children with the graphic device at hand (HLE) and forcing them to differentiate the items spatially and not on another attribute (colour), appear to be a more facilitating condition in promoting partial occlusion than any other testing procedure.

This finding implies that children's difficulty in drawing a partial occlusion is not primarily due to mental operations that they are not conceptually ready to undertake; i.e. partially deleting a complete figurative scheme. Confronted with a 2D model, they can appreciate the pertinence of the graphic system to denote the occluded form and readily suppress their conceptual knowledge of its canonicality. Nevertheless, what is really puzzling is that the performance of children changed radically from the Apples and Bottles tasks to the Balls tasks, all of which involving copying a 2D array of similar objects. The rate of success increased from 45.5% in the Apples and 28% in the Bottles task, to 68% and 81% in the Balls tasks with medium and fine occlusion in respect. Even if we recognise the differences between the Bottles task and the Balls tasks (more pronounced orientation of items and visual contrast in the former task), still these findings are vexing.

What seems to account for the task differential in performance is that, in the Balls scene, children had to copy the whole array, whereas in the Apples and Bottles tasks they had to complete the array by drawing the further object. Contrary to the initial hypothesis, having the occluding item pre-drawn seems to hinder visual realism for arrays made of similar forms. A plausible interpretation for this effect is that this condition presents what Morra *et al.* (1996) describe as a misleading situation. This situation possibly accentuates the completeness of the target form and activates the conceptual schema that preserves its canonicality; a case to which the younger children are highly susceptible. Only older children who have adequate attentional resources to make the requisite inferential judgement seem to be able to inhibit this conceptual knowledge[16]. To complete the argument, asking children to reproduce the whole array seems to draw their attention to both items (occluded and occluding), while copying from a 2D model seems to inhibit their encoding by similarity.

Turning to the non-visually realistic devices, as in the Sunset task, the facilitating nature of the Balls tasks – and in particular the fine occlusion version – suppresses all the alternative strategies from the age of 5½ and makes the responses of the 4 year olds rather variable. Consistent with the results from the previous occlusion tasks, the developmental change is mainly attributed to the decrease of separate arrangements and the subsequent increase of partial occlusions. The shift in the relative strength of these strategies takes place at the age of 5½ here. Further, consistent with our predictions, reducing the visible part of the further ball (FO) prompts children to produce more partial occlusions in this condition. Perhaps, since the array is made of similar forms, by doing this, the between-items visual contrast is also increased, inhibiting their encoding as a group of the same objects (two balls). This factor could have also contributed to the obtained task differential across arrays of similar forms.

The more interesting finding from this study is the increase of overlap from the FO to the MO condition which suggests that although the use of separation is age-dependent, the use of overlap is array-contingent. Considering this result along with those from the previous tasks, it appears that children are very sensitive to the magnitude of visual overlap between two forms. When this is moderate, they are more likely to preserve the integrity of the occluded form and use overlap. This assertion is also supported by the findings from the Apples task.

[16] Since the scenes involved similar object pairs, it would have been better if children had to match the pre-drawn form to its equivalent one in the stimulus array *before* starting to draw. However, first it was important to examine how competent children were in interpreting the drawing unaided, and second it was originally thought that the elimination of this parameter would have simplified the task.

However, at the ends of the range –either when the concealment is too small (Bottles) or too large (fine occlusion in balls)– overlap ceased to be a preferred mode.

Looking at the alignment strategies and the production sequences, children demonstrate array-fidelity and some rudimentary aspects of viewpoint-fidelity. Although Freeman *et al.* (1977) reported that the spontaneous use of the vertical dimension as a depth cue emerged at the age of 7, in agreement with Cox's assertion (1978), younger children can appreciate its relevance when provided with a model. Also, the fact that vertical alignment was the dominant device within all age levels corroborates Light and Foot's (1986) claim that the use of coloured balls and the provision of coloured pencils in previous studies (Cox, 1981; Light & Simmons 1983; Smith & Campbell 1987) inhibit the exploration of graphic convention to encode depth[17].

Turning to the spatio-temporal order of drawings, the initial analysis of all the drawings confirmed previous findings that the near-far dimension can be mapped onto the sequence of production (Cox 1981, study 1; Freeman, 1980, p.230; Ingram & Butterworth, 1989; Light & Humphreys, 1981; Light & Simmons, 1983; Smith & Campbell 1987). However, the actual tendencies appeared to be overstated when the analysis was restricted only to the non-occlusion drawings. The finding that, among the vertical arrangements, the 'first-front-at-bottom' sequential bias is contingent on the executive constraint of partial occlusion along with the absence of a 'left-to-right' order among the horizontal arrangements, suggest the lack of such spatio-temporal regularities in the present study. However, the results are rather inconclusive due to the low incidence of non-visual realistic drawings. Perhaps in a more challenging task, these production strategies can be appropriately evaluated.

7. THE BUS TASK

The present task was designed to examine the development of compositional skills required for the creation of pictorial depth. So far, children were presented with a model of a scene, depicting two whole objects one behind another. The aim of the present task was to find out how children would represent the depth relation among parts of a single object, working from memory. The drawing theme was rather common, and the figural embellishment required was minimal and simple to assess children's abilities without placing a cognitive overload. Children spontaneously draw cars and buses in their drawings but they do not necessarily represent the third dimension by partially occluding the far wheels. The purpose of the study was to investigate the developmental trends in the graphic solutions that children unaided adopt in the representation of depth as they relate parts of a figure to each other, to the proximal structure of the figure and to the distant spatial frame. It was also expected the demanding nature of the present task to suppress visual realism relative to the previous tasks and consequently allow the examination of the spatio-temporal orders of production more systematically.

7.1 Method

7.1.1 Apparatus and procedure

Children were presented with a line drawing of a bus in side view, pre-drawn on their paper. Its shape was simplified and undifferentiated, with no directional cues (rear/front) or contextual features (windows, wheels) added. The contour of the bus was a rectangular block 6.00 cm long and 2.50 cm wide. To enhance its representational intelligibility for the younger children, the word 'bus' was written within its enclosed space. Children were asked to complete the picture by adding the four wheels that were missing. The ground-line was not

[17] Bremner (1985a) made the same claim but discarded it on the basis of his finding that, in the absence of coloured pencils, a two-item array was reproduced with horizontal arrangements. Possibly, the use of a 3D model instead of a picture might have contributed to this result.

pre-drawn on children's papers so as not to confine the drawing space available. Emphasis was placed only on the number of wheels that had to be drawn without making explicit demands on their spatial differentiation in terms of the back and the front wheels. No further guidance was given. If on completion of the drawings, children had only drawn two wheels, they were prompted to add the other two.

7.1.2 Scoring criteria

Drawings were scored for the present of explicit and implicit depth cues. Children could either use partial occlusion to demarcate spatially the near/far wheels or they could anchor each pair on two different sides of the bus (see Figure 5.10). With the former technique, spatial differentiation of the near-far wheels is accomplished by varying the completeness of the equivalent graphic forms (front wheels with a complete circle *vs.* back wheels with an incomplete circle), while with the latter approach spatial differentiation is achieved by using distinct (parallel) spatial axes in aligning them with the body of the bus. Subsequently, these two solutions yield different degree of visual realism.

Figure 5.10 Diagrams of strategies for wheels' spatial differentiation

Drawings were also scored for the arrangement of wheels along the bus's edge. Attention was drawn to the spacing and distribution of forms relative to each other and to the main structural unit of the bus. This analysis was carried out to identify those cases in which the application of various compositional principles implied the presence of a spatial relation between the requisite items, without this been specified by the use of explicit depth cues. In addition, the temporal order of production was monitored to investigate any sequential constraints and possible ordering principles of representational significance not evident from a simple inspection of the finished product.

7.2 Results

7.2.1 Depth cues

A. Spatial reference axes

The drawings were first classified on the number of the reference axes children used in aligning the wheels to the edges of the bus. Nearly all children (101/110: 92%) aligned the four wheels with one edge of the bus. Of those children, all except one[18] drew them on the lower boundary of the bus. These children seem to have interpreted the given configuration as a side view representation of a bus, and, by using its lower horizontal axis to align the items, created a plane perpendicular to the surface of the page.

There was also a small number of children (8/110: 7%) who drew the wheels in pairs along two parallel edges, creating what Cox describes as a 'folding-over' arrangement (Cox, 1992, p.137-141). In six of these cases they used the upper and lower edge of the bus, and in two drawings the side edges. Both of these placement patterns are elusive, yet interesting solution in depicting in-depth relations between parts of a figure. Their difference lies in that in the first case, which was relatively more popular, children encoded depth along the vertical dimension, thinking of one pair of wheels behind the other *across* one's line of sight. On the contrary, in the second case children used the horizontal dimension, thinking of one pair behind another *along* one's line of sight. Both types of arrangements were restricted to the two younger age groups while the use of the parallel 'side-edges' was found only among the 4 year olds.

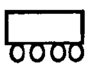
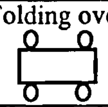
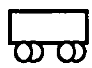
[18] There was only one drawing by a 4 year old child who placed all four wheels at the upper edge of the bus and produced a top-bottom inversion.

Although the 'folding-over' solution demarcates spatially the two pairs of wheels, it is a trade off between over-inclusiveness and view-specificity. It preserves the completeness of each wheel and the spatial relationship of each element to the other and to the immediate spatial frame of the bus. However, all these relations are uncoordinated with the external spatial framework resulting in an impossible view. It is not surprising that these arrangements are only seen from the younger children who are often very informative at the expense of visual realism. Finally, there was a single case of a 4 year old child who filled all the sides of the bus's contour with wheels; a strategy which reflects more a rudimentary disposition towards space-filling rather than a futile intention of spatial differentiation. This last case was excluded from any subsequent analysis.

B. Near-far wheels: Completeness of form

Since the majority of children did not spatially differentiate the near from the far wheels by the alignment axes, it was examined if they did so by means of partial occlusion. Of the 101 children using one reference axis, 78 children drew the back wheels intact and 23 eliminated the part occluded by the front ones. All partial occlusions were horizontal except of one case of a 5½ year old where it was vertical with the incomplete contour placed above the front wheels within bus's outline. Table 5.14 presents the various spatial solutions.

Table 5.14 Distribution of drawing devices by age.

		Spatial representation		
		Linear	Folding over[a]	Partial Occlusion
Age groups	N			
4	29	26	3	0
5 ½	25	17	4	4
7	28	21	1	6
8	27	14	0	13
Total	109	78	8	23

[a] *The top-bottom edges were used here for the illustration of the 'folding-over' device.*

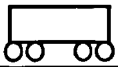
Of those children who aligned all four wheels along the same boundary, the ability to use partial occlusion instead of linear arrangement was age-dependent [$\chi^2(3) = 17.7$, $p<.001$]. However, this ability is not yet fully developed even among the 8 year olds. Also, its acquisition appears to be very smooth as none of the comparisons between successive age groups revealed a significant shift in response.

C. Compositional strategies

Although age-related differences were found in the use of the HLE device, the fact that a significant number of children within each age group used a complete contour to represent the back wheels called for a closer examination. In all of the previous tasks, non-occlusion drawings were re-examined on the basis of the proximity of the complete contour of the further object to that of the front one. Similar criteria were applied in the present study to be consistent in the treatment of the data across the occlusion studies. Thus, if the ordering of the wheels was done with no attention to placement patterns, it constituted a mere enumeration of elements with no thematic structure. Only when the ordering of forms was done with attention to the distance and spacing of the requisite items, the relationship between constituent parts was pictorially implied – albeit explicit depth information were lacking. At the next stage of analysis, the drawings with a linear arrangement of wheels were analysed for the use of compositional grouping principles. The purpose was first to identify the ordering principles in the drawings with linear arrangements, and subsequently to differentiate arbitrary from rule-governed spatial compositions in order to investigate age-related trends.

Two types of compositional arrangements were identified; children either dispersed the wheels along the full length of bus's near boundary or clustered them in pairs. The figural overlap between any of the front wheels with its corresponding back one suggests spatial proximity between the forms. As a consequence, paired compositional arrangements are more structurally informative than dispersed ones since they convey some information about the contiguity of wheels separated in depth. In addition, within each of the two compositional arrangements, various degrees of relative proximity between the forms were observed. Each item can be either spatially separate or united with the following one of the rank or of the pair. Table 5.15 presents the frequency distribution of the two gross compositional arrangements broken by wheels' relative proximity and age levels.

Table 5.15 Distribution of compositional arrangements by proximity of wheels and age.

	Compositional Arrangements					
	Paired Composition			Dispersed Composition		
						
Age groups	Separation	Unification	Total	Separation	Unification	Total
4	2	0	2	19	5	24
$5^{1/2}$	2	4	6	9	2	11
7	6	5	11	10	0	10
8	5	5	10	4	0	4
Total	15	14	29	42	7	49

Notes *This analysis applies only to the drawings where all four wheels were arranged linearly.*
The schematic illustration of the two gross compositional arrangements uses as a paradigm the spatial separation between adjacent wheels.

A test of association between the type of compositional arrangement (columns on totals) and age was statistically significant [$\chi^2(3) = 18.8$, $p<.001$]. The data shows that children of 4 years old prefer to spread the items along the bus's boundary, while at the age of 5½ children start to group the wheels in pairs which is preferred by 8 year olds.

Further, the relative degree of contiguity between adjacent wheels seem to be dependent on the type of compositional arrangements (bottom row) [$\chi^2(1) = 9.04$, $p<.01$, with continuity correction]. When children produced paired arrangements, they either spatially separated the wheels within each pair or united them by attachment or overlap. This pattern was present within each age level. On the contrary, in dispersed compositions, the majority of children drew them as separate units, avoiding any contact between them. Moreover, among paired configurations, in all but one drawing, the unification of wheels was achieved by attachment, whereas among dispersed configurations, which were exclusively drawn by younger children (4 & 5½), all unified arrangements were achieved by overlap. This suggests that the representation of contiguity of wheels by the young children is unusual and its occurrence is more the result of poor motor control and lack of graphic planning rather than an intentional attempt to portray the spatial overlap between the components.

7.2.2 Structural integrity: wheels-bus

Although the scoring of drawings mainly focused on the relative placement of wheels, coincidentally some interesting cases came to light which deserve a brief mention. First, the data revealed that in 15% of the total drawings children also used partial occlusion to represent the spatial overlap between two other figural boundaries; that of the wheels with the boundary of the bus. In those cases, children omitted the upper part of the wheel concealed by the frame of the bus. Embedding the wheels to the lower boundary of the bus brings the body of the bus to the foreground and the wheels at the background, and creates structural integration.

Although the overall frequency of such drawings is too low to allow meaningful analysis, the following observations were made: First, they were found relatively more frequently among the 5½ and 8 year olds, and their occurrence did not reveal any clear developmental pattern. Second, this advanced strategy seems to be triggered by the use of partial occlusion for the back wheels. More than half (14/23, 61%) of the children who drew the back wheels partially occluded by the front ones were also able to draw all four wheels embedded to the bus frame. On the contrary, this device was rarely found in the drawings where no such differentiation was made.

The examination of the quality of anchoring also revealed a small number of cases (15%) where children totally failed to spatially relate the wheels to the bus edge, as they were drawn distant and unsupported, floating in undetermined space. However, these drawings constitute nearly half of the total drawings of the 4 year olds (13/29, 45%) whereas by the age of 5½ children could firmly anchor them on bus's edges [$\chi^2(1) = 5.6$, $p<.05$, with continuity correction]. This result suggests that young children lack the skill to coherently unite the constituent elements of a figure to its structural frame which either results from their meagre motor control and graphic planning, or from their disposition to keep each feature to its own secluded space.

7.2.3 Spatio-temporal order of production

The sequence of production was investigated to further explore the principles underlying the compositions. Five sequential patterns were identified in the drawings of wheels; one in the folding-over arrangements and four in the drawings where all the wheels were drawn along one axis. Since folding-over drawings constitute only a small portion of the total sample, a brief mention will be made to that first. In these drawings, children dealt with each spatial axis separately, drawing first the pair of wheels resting on the same plane and proceeding to the next one, with a left-right (→) sweep for compositions with wheel anchorage on the 'side-edges', and a bottom-top (↑) sequence for those with anchorage on the long edges. Since geometric constraints are rather minimal in these configurations (space availability at each boundary and relative proximity between parallel edges), the resulting fixed order of production is more likely to have resulted from the prevalence of representational (symbolic) forces over formal ones. The collinear graphic sequence across pairs suggests that children start and complete the items that rest at the front before moving to those lying at the back.

The types and frequency distribution of sequential patterns for 'one-axis drawings' are presented at Table 5.16 by age group. Drawing the wheels in an uninterrupted linear succession (serial) was the most popular production sequence (44%) followed by the inward order (26%), where wheels at the ends of the rank were drawn before the ones in the middle. Close in preference to the inward sequence was the mixed sequence (23%), whereas children rarely drew the wheels at the middle before the wheels at the ends of the rank (outward sequence 7%). The frequency distribution of the types of production sequences revealed exactly the same pattern when partial occlusion drawings were excluded, in which idiosyncratic geometric constraints operate.

Table 5.16 Frequency distribution of production sequences by age.

Sequential Order		Age groups 4	5½	7	8	Total
Serial	•⟶	10	5	15	15	45
Inward	⟶•⟵	9	11	2	4	26
Outward	⟵•⟶	4	2	1	0	7
Mixed	········	3	3	9	8	23

Notes *The dot in the diagrams denotes the onset of production sequence.*
Folding-over arrangements are not included in this analysis.

Analysis of the data in Table 5.16 also revealed that different age groups demonstrate different sequential preferences [$\chi^2(9) = 24.22, p<.01$]. With the exception of 5½ year olds, all the age groups favoured serial order. However, while inward and outward sequences were mainly found up to the age of 5½, mixed sequences were used from the age of 7. Again, confining this analysis only to the line-up arrangements, the same developmental pattern was obtained.

These findings suggest the following: The ordering of wheels in an unbroken sequence prescribes movement to the nearest item along the trajectory and constitutes a production routine which is the most economical. The fact that it is the most popular within linear and partial occlusion drawings considered in isolation, further supports the contention that it has no representational significance but evolves from formal principles (proximal relocations).

The outward sequential pattern appears to be the accidental outcome of a non-deliberate representational attempt to draw four wheels. In particular, all the outwards orders were produced by the younger children (4 & 5½) who, intending to draw primarily two wheels, had no available space to accommodate the other two when prompted, and simply added one at both ends of the array.

On the other hand, the age-related differences in the use of inward and mixed orders reveal sequential patterns of different representational logic. The fact that the inward sequential pattern deviates from the serial production order and is mainly used by the young children[19] (4 and 5½) – who would have been the most susceptible to a graphic routine– indicate that it is a rule-governed sequence of a symbolic significance. These children, even if they didn't spatially differentiate the front from the back wheels by means of partial occlusion, did so by the temporal order of production. Drawing first the end components and fitting latter the other two within the intermediate space, separate temporally the wheels lying at the front to the ones lying at the back, and prescribe movement from the front plane to the back one.

Considering the mixed orders, a closer inspection of the data revealed that such sequential patterns occurred in the drawings of young children (4 and 5½) with dispersed compositional arrangements, whereas they appeared more often in the drawings of older children (7 and 8) only when they produced paired compositions (paired non-occlusion and partial occlusion drawings). This result indicates that a mixed order can be the result of lack of planning in young children who are guided simply by the availability of space in the placement of features. On the contrary, a mixed order in the drawings of children above 5½ years of age could have arose from the formal constrains imposed by the use of partial occlusion as well

[19] This pattern constitutes 81% of non-occlusion drawings and 78% when non-occlusion and occlusion drawings are collapsed.

as from the increased graphic competence which allow them to break from a long exercised graphic routine (serial order).

Finally the drawings were analysed for the starting position and the ordering of items with respect to the left and right. The drawings which qualify for the present analysis were those where the children traced the wheels linearly moving along the horizontal dimension. Analysis was confined to the delineation of the first two wheels for matters of simplicity. In the total of 107 drawings, where the first two wheels were horizontally traced, 79 (74%) pairs were drawn from left to right. Developmental differences were observed only in the youngest of the children whose responses did not demonstrate any preference in the starting position and direction of movement (left→right: 17, left←right: 11).

A final remark deserves to be made concerning the embellishment of wheels with secondary features, such as spokes and hubs. Although, figural elaboration was not requested, some children spontaneously drew additional features. The data showed that children tend to add more features with age [$\chi^2(3) = 12.0, p<.01$]. The change towards the inclusion of secondary features was rather marked from the age of 4 to 5½ years [$\chi^2(1) = 9.28$, $p<.01$, with continuity correction], whereas from this age onwards this tendency was rather stable.

7.3 Discussion

The present study was designed to investigate how children render the three-dimensionality of a single form, portraying the depth relationship between its constituent parts. The results reveal clear developmental trends both in the use and the 'clarity' of pictorial cues for the representation of the structural and spatial features of the object. In general, the present task yielded a lower percentage of partial occlusion output relative to the previous studies. This is mainly true for older children who overall produced more advanced drawings copying from a model, whereas the removal of a 2D model did not affect the performance of younger children. This result seems to suggest that children, being in a transitional stage of acquiring the requisite graphic device, can appreciate its relevance when provided with a model, while its propitious effect is abolished at a stage where the graphic competence is very limited.

The majority of children in each age group produced side-views of the bus of varying sophistication by drawing the wheels below the bus contour, while mixed-views were rare and mainly found among the younger children. Four year olds mainly drew the wheels as complete and separate entities, leaving their spatial relations undetermined. The dispersal of wheels below the bus was also occasionally done with disregard of the proximal framework. The lack of anchorage and of explicit reference lines in the drawing of pre-school children have been regularly reported in the previous studies (Golomb & Farmer, 1983; Leeds *et al.*, 1983; Golomb, 1987). These features elucidate what Piaget describes as *synthetic incapacity* and are symptomatic of children's preoccupation to keep each element in its own secluded space.

From the age of 5½, the constituent elements attain solidity and the composition gains structural unity. Children also begin to apply ordering principles, grouping the forms on the basis of the spatial proximity between the front and the back wheels. The 'folding-over' arrangements, albeit perceptually non-veridical, exemplify such an attempt. Although the developmental change in the use from one to two spatial axes is not pronounced here, it is in the expected direction. These arrangements, first described by Kennedy (1978), have been found in young children's drawings of geometric objects (Phillips *et al.*, 1985; Chen, 1985) and naturalistic scenes (Cox, 1992, p.137), and interpreted as object-centred topological representations. Also, this solution is indicative of the way children of different ages decode the pre-drawn outline of the bus. In accordance to Willats' view (1985), younger children are more likely to have taken the enclosed region for the whole volume of the bus rather than for its visible surface. As in the former case each edge stands for a surface, children, willing to distinguish spatially the pair of wheels, subsequently anchor them on different axes.

Children at the age of 7 are totally competent to form a correct description of the scene and accordingly unite all the features in a coherent spatial framework, revealing a single point of view. Although they still oscillate between paired and separate arrangements, it seems fair to say that the former is the precursor of partial occlusion. With the emergence of the deliberate structuring of available space, drawings gradually attain view-specificity, and by the age of 8 children mostly drew paired arrangements, half of which depict the back wheels partially occluded by the front ones in each pair. It was interesting to find that, with the increased awareness that edges stand for occluding contours, children can represent simultaneously two types of partial overlap; that of the bus's edge with the wheels and that of the front wheels with the back ones. In general terms, the obtained developmental pattern in the use of various graphic devices is in agreement with the features identified by descriptive theories (Lowenfeld, 1952; Piaget & Inhelder, 1948/1967) and validated by empirical findings (Golomb, 1987; Golomb & Farmer, 1983; Leeds *et al.*, 1983) as characteristic of different stages in the organisation of pictorial space.

Finally, consistent with the initial prediction, the demanding nature of the present task shows the presence of spatio-temporal strategies which were reliably used by those limited on their graphic skills. Although the required completion of the bus's form favours an automatic sequential order, when younger children break away from it, they mainly have a representational reason. The prevalence of inward sequences at the ages of 4 and 5½ proves that it is the spatial position of the wheels which dictates the temporal order of production. With respect to the other non-linear sequences, their examination along with children's graphic devices allowed us also to assess the contribution of various forces in the ordering of wheels. Age-related differences were found in their use which, among the older of the children, are governed by formal constraints (partial occlusion) whereas among the younger are more of an accidental outcome due to lack of advanced planning.

In the present chapter, the factors underlying children's ability in omitting parts of an object not present in their visual field were examined with arrays of two-items aligned in depth. The study of partial occlusion will continue in the next chapter where scenes of different structural properties will be employed. Finally the research in this area will be completed by assessing the overall performance of children across all the tasks and at different age levels, and general conclusions will be drawn.

CHAPTER VI

OCCLUSION: STRUCTURALLY INTEGRATED OBJECTS

1. PREFACE

In the previous chapter we examined the developmental progression in the pictorial representation of partial occlusion when one form was behind another. All the scenes in these tasks involved two discrete objects whose structural blending was dependent on the viewpoint of the observer, and was not an integral property of the array. Specifically, a partially seen object behind another is spatially distant in reality and appears complete simply by moving to a different position. If children's reluctance to represent the spatial overlap by means of partial occlusion stems from a conceptual bias to draw what they know about the actual internal spatial relation of the forms, some of their inaccurate solutions (structural segregation), albeit non-visually realistic, can still be array-specific. A way to investigate the nature of these representations, is to present children with occluded scenes where the structural integration of objects is intrinsic to the array and not relative to our line of vision, for example a flower in a vase.

The present chapter involves three tasks where one form is partially occluded by virtually sharing a common boundary or region with another form. If among the non-visually realistic strategies, the tendency to spatially segregate the forms has a conceptual origin and represents a drawing convention specific to the spatial characteristics of the scene, it should be substituted for structurally unified arrangements in the present series of tasks. On the contrary, if the drawing devices, identified in the previous chapter, are impervious to the relationships internal to the array, then they will manifest general modes of approaching a graphic task involving partial occlusion.

Thus, the present work aims to unconfound the use of partial occlusion as a depth cue and to investigate its relative frequency in representing spatial adjacency (next) and partial enclosure (inside). At first, the data from the two new types of partial occlusion will be considered, presenting each task separately and comparing performance as we go along. A comprehensive look at performance across all the partial occlusion studies will follow in the next chapter where systematic comparisons will be made across the structurally different scenes, and the mechanism underlying the pictorial representation of view-specific relations will be discussed.

2. OVERVIEW OF THE STUDIES

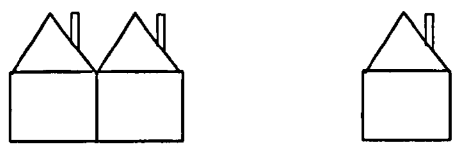
The first task was designed to investigate the representation of spatial contiguity of two identical forms, aligned along the horizontal axis. The remaining two tasks involved arrays of dissimilar items where one form was partially embedded in another. The tasks also made varied demands on the figural transformation of the occluded form to accomplish visual realism. Some drawing themes yielded a greater loss of the form's integrity (man-in-boat task), whereas others affected the structure of the target object minimally (terraced houses and flower-in-vase tasks). Thus, if a certain informative attitude undermines the pictorial representation of occlusion, then we should obtain variation at occlusion output across these scenes. In addition, if the alternative drawing strategies capture the structure of the scene, their relative frequency should also change between drawings of a spatially adjacent form and an enclosed form.

Children were presented with a model array for the Houses task, whereas for the other two they had to work from imagination. This variation on task demands (copy from a model *vs.* draw from memory) across the structurally different scenes resulted from the difficulty in finding the verbal description that would adequately prompt the children to conceptualise and subsequently draw two forms adjacent to each other[1]. The instructions were either directing the children to make a drawing exactly as the one on the stimulus card (Houses task) or requiring them to complete the drawing by adding the target form, without explicitly referring to the spatial relation between the items (Vase task & Boat task). The occluding form was always pre-drawn on children's papers to alleviate the executive constraints and to control for errors resulting from drawing biases (drawing the occluded form first). Although the findings in Chapter V showed that this manipulation surprisingly hindered visual realism, it was mainly attributed to the similarity of the constituent items in the arrays used. Since in the present arrays the items varied substantially, this manipulation wasn't expected to interfere with performance. The order of presentation was fixed, with the Houses task first followed by the Vase and the Boat task. However this order was not consecutive. These tasks were administered as part of a battery of experimental studies where no one task is preceded or followed by any other task that assesses exactly the same ability (in this case partial occlusion in structurally integrated scenes).

3. THE HOUSES TASK

3.1 Apparatus

Children were presented with a stimulus array, depicting two houses on the horizontal axis sharing a common boundary (Figure 6.1-a). All children were provided with the left house of the array, pre-drawn on their paper (b), and were requested to draw the one that was missing to make their drawing look exactly the same as the one on the stimulus array.



a b

Figure 6.1 Schematic illustrations of layout (a) and pre-drawn form (b)

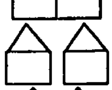
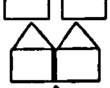
[1] Labels such as "next to" or "near to" were thought that vaguely described the structure of the scene. and the term "adjacent" were beyond the comprehensive skills of the younger of the children.

The drawings were scored for the graphic means children employed to connect the requested house to the pre-drawn one. To be able to reliably differentiate the cases where children eliminated the shared boundary (wall) from those cases where they drew a complete contour and immaculately superimposed its jointed side to the pre-drawn one, the order of direction of drawing strokes was recorder.

3.2 Results

The drawings were analysed applying the criteria of the contour's integrity and its relative distance to the pre-drawn form. Three basic solutions were found in the treatment of the shared boundary. Children applied the HLE device, omitting the overlapping boundary from the graphic schema of the house, drawing an open rectangle meeting the pre-drawn form at junctions. Alternatively, they drew a complete outline with various degrees of proximity to the pre-drawn house. Among the latter solution, children either preserved the closeness of the houses with no inter-item gap or segregated the forms. Only one drawing of a four year old child could not comfortably fit to any of the aforementioned drawing solutions, and was excluded from the analysis. Table 6.1 presents the distribution of connection techniques across the age groups.

Table 6.1 Distribution of drawing strategies by age groups.

		Age groups				
Strategies		4	5½	7	8	Total
HLE		4	11	22	26	63
Segregation		22	13	4	1	40
Attachment		2	1	2	0	5
Enclosure		1	0	0	0	1
Total		29[1]	25	28	27	109

[1] *One equivocal drawing was excluded from analysis.*

The results showed that the use of the most sophisticated connection strategy was age-dependent, as there was a gradual increase in the use of the HLE device [$\chi^2(3) = 46.33$, $p<.001$] with significant shifts in response between successive age groups up to the age of 7 [4 to 5½: $\chi^2(1) = 4.69$, and 5½ to 7: $\chi^2(1) = 5.33$, $p<.05$ with continuity correction]. With respect to the specific solutions that children adopted, it seems that children were guided by two forces. They either drew a complete contour spatially distant from the pre-drawn house (segregation, 37%) or embedded an incomplete outline to the complete one (HLE, 58%). There was only a negligible number of children (5%) who attached the complete contour of the requisite house to its neighbouring one. This can be seen as a compromise solution between the former ones since it preserves the appearance of the array, attending to the structural closeness of forms, while maintaining the integrity of the drawn house. However due to the absence of an age-related trend in its use and to its rare occurrence, it does not seem to be an intermediate representational stage. There was also one 4 year old child who produced an enclosure configuration violating any array/view-specific information.

Since the overall distribution of the drawing strategies is essentially bimodal, the same developmental trends are revealed regardless of whether the data are analysed in terms of the form's integrity (HLE *vs*. segregation and attachment collapsed) or the form's proximity (HLE and attachment *vs*. segregation). Thus, the pattern of developmental change can be summarised as follows; the performance of 4 year olds is basically unimodal with the use of a complete-segregated contour being the dominant device. This pattern significantly changes at the age of 5½ where children oscillate between segregate and view-specific configurations. The bimodality is finally resolved at the age of 7 where the use of an incomplete-embedded contour becomes the preferred mode.

3.3 Discussion

In the present task there was no feature concealed from children's viewpoint. The omission of the joint wall from the graphic schema of the house does not induce a radical transformation of its structural integrity. Yet, the segregation bias among the younger children was not readily inhibited since the structural integration of the forms in the scene did not change the relative frequency of the alternative drawing strategies (unification vs. separation strategies). Young children are still reluctant to represent the point of contact between neighbouring parts, overlapping their boundaries, and continue to prefer assigning each form its own distinct space. These results suggest that spatial segregation might not render the internal structure of a scene, but rather signifies an all-purpose drawing convention, at least among the youngest children. However, the performance of the older children suggests that the properties of the scene not only foster view-specific representations in general, but the correct connection style in particular. The prevalence of partial occlusion over attachment configurations implies that, once children overcome the temptation of segregation, they can accomplish visual realism not by just preserving the appearance of the scene but by conveying its underlying spatial structure. And this reflects an intrinsic modification in the way children conceptualise the scene.

4. THE VASE TASK

The present task differs from the previous one. The spatial relationship between the occluding and the occluded form is one of partial enclosure and not adjacency, and children have to draw the requisite spatial relation from memory and not just copy it from a model. On the other hand, as in the Houses task, the portrayal of partial occlusion brings visual realism to the composition without omitting any defining features of the occluded form. Thus, the present task is of comparable difficulty to the previous task for those children who are compelled to preserve canonical features of forms. However the relation between the flower and vase (enclosure) is different from the relation between the terraced houses (adjacency). If children's alternative strategies are determined by the spatial relations particular to the array, then differences in the frequency of non-occlusion strategies should be obtained across the tasks

4.1 Apparatus

Children were provided with papers where a line drawing of a vase, 2.00 cm high and 1.50 cm wide (Figure 6.2), was pre-drawn. A flower with its stem missing was suspended over the rim of the vase, 1.30 mm distant form it. Children were asked to complete the display by drawing the stem of the flower. Drawings were scored for the solutions children adopted for depicting the non-visible part of the stem.

Figure 6.2 Schematic illustration of the stimulus material

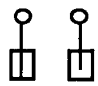
4.2 Results

4.2.1 Partial occlusion

A realistic portrayal of the scene requires children to draw only the part of the stem which protrudes from the vase, meeting its rim at the junction. Table 6.2 presents the number of children who produced partial occluded drawings with the implementation of the HLE device and the number of children who failed by using alternative strategies. The task elicited a high incidence of partial occlusion output in the overall sample, as ninety children (82%) succeeded in producing visually realistic drawings, whereas only twenty (18%) failed. With respect to age, the frequency of visually realistic drawings increased from the age of 4 to 5½

[$\chi^2(1) = 5.77$, $p = .02$, with continuity correction], while from that age onwards the majority of children reliably omitted the non-visible part of the stem.

Table 6.2 Frequency distribution of drawing strategies by age groups.

		Age groups				
Strategies		4	5 ½	7	8	Total
HLE (P. Occlusion)		18	23	23	26	90
Overlap		7	2	5	0	14
Segregation		5	0	0	1	6
Total		30	25	28	27	110

Among the non-visually realistic drawings, transparency (overlap) was the only alternative strategy used by all the children up to the age of 7, and its use fluctuated from 8% to 23% with no apparent age trend. The only group which performed differently to any other was the 4 year olds who also used segregation – by having the stem hanging above the rim – with comparable frequency to transparency. There was also a clear sub-group of transparent drawings where the stem, instead of running down to the bottom of the vase, was pendent within the vase's outline[2]. Those cases were mainly found among the 4 year old children.

4.2.2 Figural elaboration

The task required drawing from memory, so it also provided the opportunity to examine the graphic forms children used to draw the stem as well as the spontaneous depiction of additional features. In general, children's drawings improve by the substitution of two dimensional graphic devices (regions) for one dimensional ones (lines, dots) and by the subsequent inclusion of more detail. Thus, as children's graphic vocabulary expands, their drawings acquire volume and gain representational status.

In the present task, 64% of all children used a single line to delineate the stem, whereas 36% drew it as a region with a pair of lines. The choice of the graphic device was dependent on age [$\chi^2(3) = 15.41$, $p<.01$], as the use of a single line declined from 90% among the 4 year olds to 41% among the 8 year olds. In addition, 79% of children's drawings had no additional elements whereas 21% were embellished with leaves and flowers attached to the stem. Again, age was associated with the drawing's embellishment [$\chi^2(3) = 10.30$, $p<.05$] which was seen in one third of the drawings of older children (7 & 8 years old).

4.3 Discussion

The higher incidence of partial occlusion in the present task, compared to the previous one, suggests that children can readily use the HLE device even when they draw from memory. From the age of 5½, they not only produce mainly view-specific drawings, but also use exclusively an alternative strategy which encodes array-specific information (overlap). In addition the scene fosters the use of partial occlusion even among the 4 year old children and suppresses the frequency of separation mode which accounted for only the 5% of all drawings.

The fact that realistic portrayal of the scene does not necessitate the loss of any canonical features from the occluded form (blossom, stem) obviously contributes to the high incidence of partial occlusion over the total sample and within the younger children (4 & 5½). However, what makes this claim not totally satisfactory is the differential performance of children across the previous and the present task, since in both of them the form of the occluded item undergoes minimum transformation.

[2] Lenient criteria for junction formation were adopted in order not to penalise children for poor motor control. The stem was considered separate from or transparent within the vase if it was more than 3.00mm distant from or embedded in it respectively.

A number of researchers have maintained that the way children translate the scene to themselves has a substantial impact upon its conceptualisation and subsequently upon the planning and execution of the drawing (Cox 1991; Crook, 1984 & 1985; Van Sommers 1985). It is more likely that the structure of the scene in the present task is salient enough to elicit a mental description, where the spatial relation between the forms attains representational significance (flower-*inside*-vase). If in the previous task children conceptualised the scene as 'one house *next to* the other', this description is not adequate to elicit even array-specific representations (attachments). The shift in the use of alternative strategies across the two studies from segregation to overlap corroborates the latter argument and suggests that even if some children still have difficulty preserving view-specific properties, their drawings contain array-specific information which is conceptually driven (knowing that part of the stem is within the vase).

The fact that a small number of 4 year children still produce separate arrangements can be either interpreted on the basis of their 'synthetic incapacity' (Luquet, 1927/1977) or as an inability to understand the task requirements[3]. However, although instructions did not stress the spatial relation of the forms at the scene, the fact that the majority of children spatially integrated the two items argues against the latter assertion and suggests that spatial separation may constitute a drawing convention for very young children.

5. THE BOAT TASK

The present task is similar to the previous one in the sense that children have to draw from memory an occlusion scene where one form is partially enclosed within another. The only variation was the degree of figural transformation that the realistic portrayal of the spatial relation required. The purpose of the study was to investigate whether increasing the extent

[3] Specifically, requesting children to draw the missing stem might have led them to encode only the figurative scheme of the flower overlooking its spatial relationship to the vase.

and the nature of the hidden features that need to be omitted, suppresses the pictorial representation of partial occlusion.

5.1 Method

5.1.1 Apparatus

Children were presented with an outline of a boat pre-drawn on their paper, 8.00 cm long and 1.80 cm wide at its middle section, with no contextual features added. Children were told that they had to make their picture prettier by drawing a man fishing from the boat[4]. The boat's outline was drawn on the lower part of their paper to avoid constraining the younger of the children in terms of space availability.

5.1.2 Scoring Method

The scoring criteria and the drawing categories used were the same with the previous studies. However, due to the complexity of the human figure form compared to the occluded forms in the former tasks, it was expected that some 'immature' drawers could produce incomplete figures as a result of a less elaborate graphic schema, and not necessarily as a deliberate attempt to omit occluded features. As a consequence, for the accurate classification of the data into drawing categories, both the form's incompleteness and its correct placement, relatively to the boat's outline, were applied as criteria. Thus, view-centred representations were considered to be those where children eliminated *at least* the lower part of the body (from hips down) and firmly attached its visible top part to the edge of the boat. Some children succeeded by producing 'head only' drawings, whereas others made more elaborate drawings, representing a man down to the hips. Regardless of the scarcity of human features,

[4] It was assumed that instructing children to "draw a man in the boat" might have been interpreted by the younger children as if the man was hiding, totally occluded inside the boat, and it could have resulted in many enclosure and overlap representations. On the contrary, the notion of 'a man fishing from the boat' presupposes his placement inside boat's outline and implies that he is partially seen without making explicit reference to the appropriate spatial relation.

(a) any form spatially distant from the boat was classified as a *segregated* configuration, (b) totally embedded within the boat as an *enclosure* configuration, and (c) partially seen through the boat as an *overlap*. A drawing was scored as an *attachment* if children had drawn the lower part of the human figure (from the hip down) standing on the boat's boundary.

5.2 Results

5.2.1 Partial occlusion

The data are reported in Table 6.3. First, the distribution of visually accurate and inaccurate drawings by age revealed that the frequency of 'view-centred' drawings was age-dependent [$\chi^2(3) = 15.33$, $p<.01$], with a significant increase of partial occlusion output between the 7 and 8 years of age [$\chi^2(1) = 6.68$, $p<.01$, with continuity correction]. The pattern of the developmental progression is as follows; children of 4 years of age draw inaccurately significantly more often, whereas the reverse is true among the 8 years old, with the drawings of the two intermediate age groups being equally distributed between the two categories (Table 6.3, row with *p* values).

Table 6.3 Distribution of drawing strategies across age groups.

	Age groups				
Strategies	4	5½	7	8	Total
Correct					
HLE (P. Occlusion)	8	8	10	20	46
Incorrect					
Segregation	9	9	4	0	22
Attachment	3	6	7	6	22
Overlap	4	2	7	1	14
Enclosure	5	0	0	0	5
Total	21	17	18	7	63
Z	2.41	1.80	1.51	2.50	
P^1	<.05	*NS*[2]	*NS*	<0.05	
Total	29[3]	25	28	27	109

Note 1 The probability values given are for a two-tailed test.
Note 2 The difference is significant for an one-tailed test.
Note 3 One child produced an equivocal human from.

Turning to the types of alternative solution, spatial segregation and attachment were equally popular followed by overlap, whereas enclosure configurations were rather rare (rightmost column). Segregation, as the least structurally informative depiction, was mainly used up to the age of 5½, whereas enclosure was only used by the youngest children, and attachment was a constant temptation across all age groups. The frequency of overlap mode appears somewhat varied with no discernible peak representing a stage. Probably, its re-emergence at the age of 7 is the result of segregations being substituted for structurally integrated arrangements in general.

Despite the fact that enclosure does not seem to be equivalent to overlap as a graphic device, the fact that both preserve the internal relation of the forms within the array, allows us to consider them as array-specific modes among the alternative drawing strategies. Looking at the data in Table 6.3, it is rather interesting that, if we combine these strategies as a modified 'transparency' category, this array-specific strategy is as popular as any other non-structurally appropriate ones (separation 35%; attachment 35%; transparency 30%).

5.2.2 Figural elaboration

Further analysis of the data revealed that there was a lawful relationship between the type of figural 'incompleteness', the drawing strategy children used, and age. For example, "head only" drawings – as the most limited representation of the topic – were found basically among the 4 year olds (10 out of 12) whereas 'head-and-torso' drawings increased gradually with age from 24% among the youngest to 59% among the oldest children. In addition, there was an association between these two types of incomplete forms and the use of the visually-realistic drawing device [$\chi^2(1) = 23.08$, $p<.001$]. The 'head only' drawings were associated with the inappropriate strategy, since in the majority of them the head was segregated (8 out of 12), floating above the boat, and in one was enclosed within it. On the contrary, the 'head-and-torso' drawings conveyed an intentional attempt to omit the non-visible parts of the figure as only 5% (2/40) of those were non-occlusion drawings, presenting the man totally enclosed within the boat.

Considering the more elaborate figures (head-torso-hips-legs), the data does not suggest that the children who drew the lower part of the human figure had a depiction bias among the alternative strategies. Within the 52 perspective-inappropriate drawings, 14 depicted the whole figure spatially separate, 16 showed part of the figure overlapping the boat's boundary[5] and 22 presented the man as standing, anchored on the boat's boundary (attachment). Elaborate figures were scarcely drawn with partial occlusion (9%). Such drawings – portraying a man down to his hips with half of his legs occluded – were found exclusively among the oldest children. Obviously, these drawings are the most advanced in terms of planning and execution, as they require interruption of a long exercised sequential order (human form).

Finally, with age, children spontaneously added contextual elements to enhance the intelligibility of the scene such as a fishing rod, fish, sea-waves, and the human forms started to acquire more plasticity. Although in the majority of drawings the man was portrayed at frontal view (74%), profile views gradually increased with age from 8% at the age of 4 to 37% at the age of 8.

5.3 Discussion

As predicted, the present task suppressed partial occlusion output. The Boat task required children to modify their approach to the portrayal of a man both on a conceptual and a procedural level. They had to delete a distinctive part from its mental image and interrupt the sequential order of a highly practised graphic action. In this respect, it is not surprising that children before the age of 7 couldn't cope with the radical transformations the realistic depiction of the scene required, and produced more inaccurate drawings relatively to the Vase task. Nevertheless, what is rather vexing is the frequency of the alternative drawing solutions that children used to represent partial enclosure.

It is well documented that, when depicting structurally integrated scenes, transparency (overlap or enclosure) is a common solution to partial occlusion (Clark, 1897; Cox, 1992, p.109-116; Crook, 1984 & 1985). These types of transparency have been regarded as array-specific depictions on the basis of Luquet's notion of intellectual realism. According to this view, they exemplify a deliberate attempt to show what children know about the structure of a scene, and it is this informative attitude which interferes with the pictorial representation of partial occlusion. So, if the child's intention in drawing a transparency is to inform about what is been hidden, we should expect that increasing the complexity and the importance of the hidden form, should increase the frequency of transparency drawings. This was not the case. First, all the alternative drawing strategies (separation, attachment, transparency) had a similar frequency of occurrence in the overall sample. Second, transparency drawings increased from 13% in the Vase task to 17% in the Boat task; a rather negligible change. On the other hand, increasing the amount of concealed information promoted instead the tendency to produce separate arrangements which increased substantially from 5% in the Vase task to 20% in the present task.

These results make it clear that a transparency drawing has no privileged status as an array-specific solution. Perhaps our concern not to inadvertently promote transparency by structurally describing the scene in instructions could have generated a mental description devoid of the structural integration of forms. However, the social intelligibility of the task

[5] In the majority of cases, the features of the man that were transparent were part of his body with legs (14/16) and only two drawings had just the feet visible through the boat.

seems to convey such a relationship *per se*. Why then is transparency not primarily chosen to code the array-specific relation of 'inside'?

It is reasonable to argue that it denies the boat's solidity or perhaps violates a drawing convention according to which drawing within the periphery of a form is reserved for coding an 'in front of' relationship (Taylor & Bacharach, 1982). Either such ambiguity is not tolerated or can be prevented by children. It can be also said that the variation in the use of alternative strategies between the Vase and the Boat task can be explained on the grounds that children could be more willing to accept the vase as transparent, and thus make the stem visible through it; a case which is not applicable for a boat. Also, the Boat task yields more alternative unified strategies that the Vase task, since attachment is not pertinent in the latter case.

To summarise, increasing the information that needs to be concealed from the viewer reinforces the need to specify the unitary quality of the occluded object and to maintain its completeness. Whether this notional approach preserves the spatial relationships intrinsic to the array depends on children's age, the structure of scene and drawing conventions. Unlike the 'hatpin-through-an-apple' task (Clark, 1897; Crook, 1984), where we could not tell whether a transparency drawing informs us about the completeness of the object or about the actual organisation of the array, in the present task these two aspects could be disentangled.

The low frequency of transparencies in the present study hardly suggests a conceptually driven urge to draw an array-specific view which embodies the completeness of the occluded object. However, this does not necessarily suggest that the child lacks an informative attitude altogether, but rather that he is mainly concerned to represent adequately the occluded form (and not the inter-item relationship). The findings reveal that preserving the completeness of the occluded object is imperative for children up to the age of 7. However, object-centred views of the occluded object, totally detached from the spatial context to which it is embedded, are only found up to the age of 5½. This supports the results from the previous studies that for younger children segregation is more likely to be an all-purpose drawing convention, insensitive to the structure of the array. On the contrary, the use of attachments instead of transparencies from the age of 5½ perhaps indicates children's reluctance to violate a drawing convention, while at the same time they preserve the man's integrity and vaguely relate him to the boat.

CHAPTER VII

OCCLUSION: A COMPREHENSIVE LOOK

In the present chapter, a systematic comparison will be made between the results from the six partial occlusion tasks on structurally separate objects (Chapter V) with those three on the structurally integrated ones (Chapter VI). The aim is to examine the nature of fluctuations in the use of the HLE device across different scenes in general and different age groups in particular. This analysis will be also extended to the alternative devices children used to determine whether they are general-purpose drawing conventions or conceptually derived strategies that encode differently scene with different properties. As Freeman has stressed, the study of children's drawings

> *...should be based upon the systematic efforts which children make representing a drawing topic rather than their success; for surely the major part of development lies in the children making different strategic choices and changing their repertoires of rules, more than becoming "cleverer" (1980, p. 231).*

This interpretative stance finds an echo in the work reported here.

1. RESULTS

1.1 Partial occlusion

Table 7.1 presents the number of children drawing a partial occlusion for each task. The tasks have been ordered on the basis of children's performance from the least to the most effective in eliciting higher rates of partial occlusion output, within the two types of occlusion scenes. This pattern of task difficulty was present in the overall sample and at each age level, with moderate variations in the performance of 5½ year olds on the structurally separate scenes

Table 7.1 Number of children drawing a partial occlusion in each task by age level

Occluder — (black line) / Occluded — (red line)

		STRUCTURALLY SEPARATE SCENES (One object behind another)						STRUCTURALLY INTEGRATED SCENES (One object inside/next to another)		
		Wheel Wheel	Bottle Bottle	Apple Apple	Ball Ball	Hill Sun	Ball Ball	Boat Man	House House	Vase Flower
Age Groups	N									
4	30	0 (0%) C^{***}	0 (0%) C^{***}	2 (7%) C^{***}	6 (20%) C^{**}	10 (33%)	11 (37%)	8 (27%) C^{*}	4 (13%) C^{**}	18 (60%)
5½	25	4 (16%) C^{***}	1 (4%) C^{***}	11 (44%)	17 (68%) E^{*}	23 (92%) E^{***}	21 (84%) E^{***}	8 (32%) C^{*}	11 (44%)	23 (92%) E^{***}
7	28	6 (21%) C^{*}	11 (39%)	18 (64%)	24 (86%) E^{***}	27 (96%) E^{***}	28 (100%) E^{***}	10 (36%)	22 (79%) E^{**}	23 (82%) E^{**}
8	27	13 (48%)	18 (67%)	19 (70%) E^{*}	26 (96%) E^{***}	27 (100%) E^{***}	27 (100%) E^{***}	20 (74%) E^{*}	26 (96%) E^{***}	26 (96%) E^{***}
Total	110	23 (21%) C^{***}	30 (27%) C^{***}	50 (45%)	73 (66%) E^{***}	87 (79%) E^{***}	87 (79%) E^{***}	46 (42%)	63 (57%)	90 (82%) E^{***}

Notes *Based on the obtained rate of success, the tasks have been ordered from the most difficult to the easiest within the two types of occlusion scenes*
C =Complete forms, E =Eliminated forms.
The presence of a letter in a cell suggests device dominance and its direction, carrying out binomial tests. Asterisks show the level of significance from the test (*** $p<.001$, ** $p<.01$, * $p<.05$).

and in the performance of 4 year olds on the structurally integrated ones[1]. Also, using binomial tests, the table shows the absence or the presence of a depiction bias and its direction towards either completion (C) or elimination (E) of the hidden features in each cell[2]. As the table indicates, the pattern of developmental change varies across the tasks which display a different type of device dominance in the overall sample and within each age level. Cochran's Q tests yielded significant differences in the distribution of complete and partially eliminated forms across the nine tasks at each age level (the Q values for all age groups were significant at $p<.001$)[3].

With respect to the structurally separate scenes, there is a discernible depiction bias towards the use of complete forms up to the age of 7 for the Bus task, and up to the age of 5½ for the Bottles tasks, whereas in neither of these two conditions did the use of the HLE device attain the status of a preferred modality. On the other hand, performance in the Apples task shows a smoother developmental progression. Here, the youngest of the children continue to show a depiction preference in retaining the form's integrity, the two intermediate age groups oscillate between complete and incomplete forms, whereas this bimodality is finally resolved by the oldest of the children, who reliably employ the HLE device. As we progress to the more facilitating tasks, the acquisition of the visually realistic strategy levels off at a younger age. Thus, in the Balls tasks and the Sunset task, the ability to omit the concealed features of an object becomes well established from the age of 5½, whereas a considerable number of nursery children cease to rely as heavily on the use of complete contours in drawing the occluded forms.

---

[1] For the 5½ year olds, the Bus was easier than the Bottles task and the Balls with fine occlusion was easier than the Sunset task, while for the 4 year olds the Boat task was easier than the Houses task.

[2] To avoid discarding data from six nursery children who did not successfully completed the full set of tasks, their occasional equivocal responses were included in the present analysis. From those children, two produced no data and two other drew unrepresentative drawings in one out of nine tasks, whereas the remaining two made unscorable drawings in two tasks. The addition of those cases makes the use of the term 'complete' to describe failed responses of nursery children not strictly accurate.

Turning to performance on the structurally integrated scenes, the Boat task was the most difficult in promoting the use of the HLE device, the Houses task was of moderate complexity and finally the Vase task was the easiest. Specifically, the HLE strategy was reliably used only by the 8 year olds in the Boat task, whereas it was acquired from the age of 7 in the Houses task, and from the age of 5½ in the Vase task. The relative dominance of complete forms also varied. While in the Boat task, both 4 and 5½ year old children demonstrate a clear preference to preserve the integrity of the occluded form, this tendency is manifested only by the 4 year olds in the Houses task and is absent in the Vase task.

However, although it is evident that a bias towards a certain sort of depiction was task dependent, children of different ages demonstrate different amounts of performance variation. As the Table 7.1 shows, regardless of the nature of each scene, the performance of the youngest and the oldest of the children was relatively stable compared with that of the two intermediate age groups. Four year olds demonstrated a clear bias in drawing complete forms in six tasks, whereas no representational preference was observed in the remaining three tasks (the easiest of all). Similarly, 8 year olds consistently omitted the hidden part of the occluded form in seven tasks, while no depiction bias was obtained in the remaining two (the most difficult of all). On the contrary, children at transitional stages of development were sensitive to the properties of the scenes, and the rise and fall of the two drawing styles was evidently more variable. As can be seen from the summary table, 5½ year olds demonstrated no device dominance in two tasks, whereas they reliably used complete forms in three and incomplete contours in four tasks. A similar pattern in the prevalence of the pictorial devices was also obtained from the 7 year olds.

---

[3] Four year olds (n = 30): Q(8) = 57.16, 5½ year olds (n = 25): Q(8) = 89.59, 7 year olds (n = 28): Q(8) = 85.72, and 8 year olds (n = 27): Q(8) = 55.40.

1.1.1 Between-task differential: Paired comparisons

The extent of the differential performance ability across the tasks was more systematically assessed by comparing the matched frequencies of unsuccessful and successful representations across any pair of occlusion scenes. For this type of analysis, the McNemar test of change in the use of the two drawing devices (complete *vs.* partially eliminated forms) was employed, resulting in 36 comparisons for the total sample and within each age level. The results from this analysis are reported in Appendix C, Tables 1-5, and the general trends are schematically illustrated in Figure 7.1.

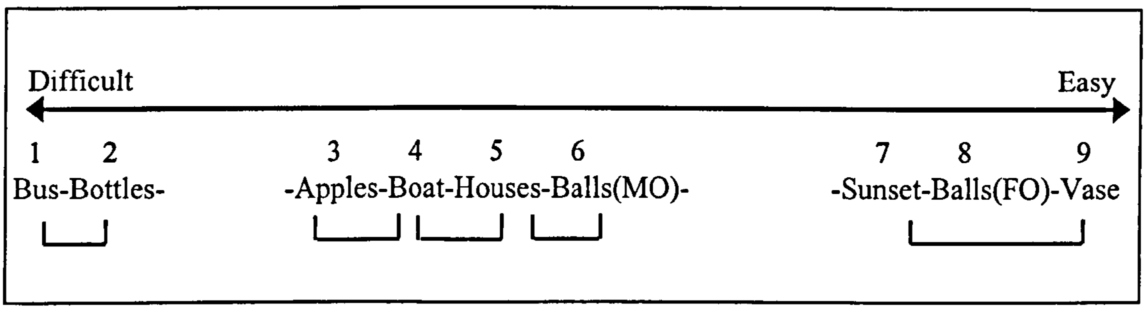


Figure 7.1 Schematic ordering of occlusion scenes by degree of difficulty and between-task differential. Tasks, jointed by inverted brackets, are of equivalent complexity.

First, the analysis of the drawings from the total sample revealed that the Bus and the Bottles tasks were the two relatively most difficult conditions. Children's performance was equally poor in both and significantly worse comparing each of them with any of the subsequent tasks. On the other hand, the Balls with fine occlusion, the Sunset and the Vase tasks were the most facilitating conditions. Children perform equally well between any two of these tasks, whereas each of them was significantly more likely to elicit better performance compared with any of the other tasks. The overall analysis also revealed that the Apples, the Boat, the Houses, and the Balls with medium occlusion were scenes of moderate complexity. As the Figure 7.1 shows, each of these scenes was easier than the preceding ones, of equal status to its adjacent, and more difficult to the proceeding scenes.

The analysis of paired performance across tasks by age, confirmed the general features of this hierarchical pattern (see Appendix C, Tables 2-5). First, the tasks at the extremes of the rank in Figure 7.1 retained their inhibitory and facilitating status. Second, although the magnitude of differences in difficulty between these scenes relative to the remaining tasks varied with age, what was consistent at each age level was that no task differential was obtained between the two most difficult conditions, and between the three most facilitating ones. Likewise, the lack of significant change in response between the successive pairs of tasks of moderate difficulty in Figure 7.1 was preserved at each age level overall[4].

1.2 Alternative strategies

1.2.1 Structural devices

Having charted the scenes on the basis of their relative tendency to evoke the use of the HLE device, we will now review the range of strategies used across the scenes. If there is a systematic transition from one strategy to the next that cuts across different contexts, then we can assume that children acquire the appropriate graphic device through a developmental accretion of drawing rules. On the other hand, if their graphic strategies are not general modes of approach, confined to a certain stage of representational ability, but their use is task-specific, then we would have some evidence to claim that children have a repertoire of strategies which they employ according to the structural features of the scenes at hand. First, Table 7.2 presents the number of strategies that each child used across all nine occlusion tasks by age level. The total number of strategies that could have been used in any task is five, thus maximum stereotypy is exhibited if children used consistently just one device for all the tasks and maximum variability if they used all five devices. From the table it becomes clear that only the two older age groups used one strategy, this being partial occlusion. On the other hand only one 4 year old used all five strategies. Yet the number of strategies that

[4] The only exception was the performance of 7 year olds who did significantly better at the Apples and the Houses tasks compared with the Boat scene.

the majority of children used within any age group range from 2 to 4. Children up to the age of 7 used mainly three and four strategies whereas 8 year olds used two and three devices. This clearly demonstrates that children from a very young age have at their disposal a repertoire of choices and that their use depends on what they consider as having representational precedence. Thus what changes with development is not the acquisition of newly emerging devices but children's shift of priorities.

Table 7.2 Number of strategies used at each age level.

	Number of strategies used				
Age groups	1	2	3	4	5
4	0	4	13	12	1
5½	0	5	12	8	0
7	1	6	11	10	0
8	6	8	12	1	0
Total	7	23	48	31	1

Further, Figure 7.2 presents the distribution of drawing devices by task. Since the Bus task elicited strategies that were not strictly comparable to those employed in the other tasks (i.e. paired compositions), it was excluded from this study. Similarly with previous findings (Chen & Holman, 1989; Freeman, 1980; Morra *et al.*, 1996), Figure 7.2 shows that the relative use of the alternative graphic strategies varies from task to task. However, regardless of the properties of a scene, the tendency to spatially separate the forms seems to be a steady temptation (see Ashton, 1997). Segregation might not always have a privileged status among the visually inappropriate strategies, but each time is a constant competitor to a different device across the tasks. It is clearly the strategy of choice for the Bottles and Houses tasks and to a lesser extent for the Balls with fine occlusion. On the other hand, it appears as frequently as overlap in the Apples, moderately occluded Balls, Sunset and Vase tasks, and as often as attachment in the Boat task. And although its use varies substantially when the level of task difficulty is high to moderate, is kept at an evenly low level among the most facilitating conditions.

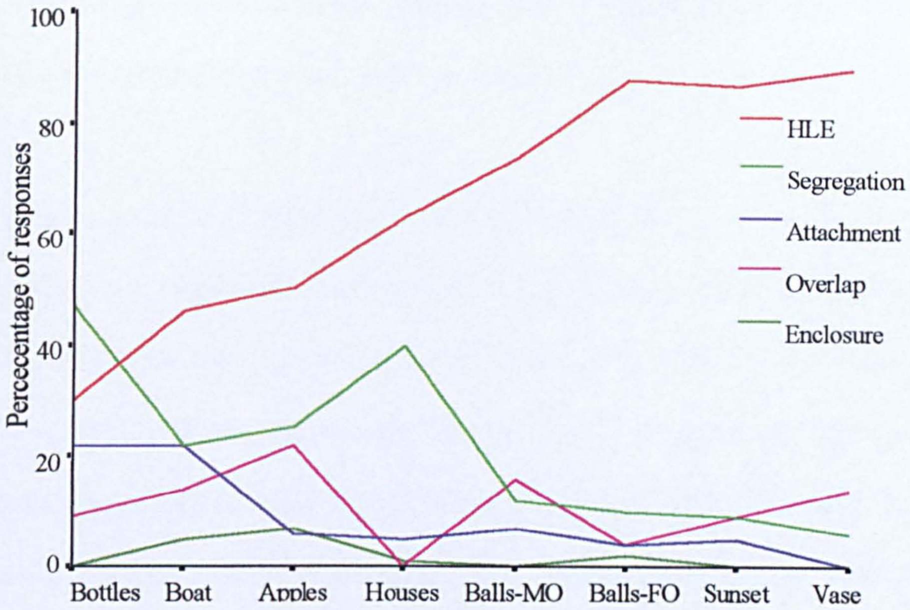


Figure 7.2 Distribution of drawing devices used across occlusion scenes.

Turning to the relative frequency of the unification strategies, the following observations can be made. First, enclosure configurations have the lowest rate of occurrence across all the scenes. With the exception of two tasks (Bottles, Boat), the strategy of attachment is the next least used with no fluctuations at its output across conditions. Overlap, on the other hand, is the most task-dependent strategy – as its uneven distribution suggests –, yet its use does not systematically vary according to whether the items of the array are integrated or separate.

At this point it will ease matters in interpreting this picture if we consider the developmental trends. The only strategy that revealed a reliable age-related trend in its use throughout the tasks is segregation. Of course, the pattern of its decline with age is largely dependent upon the properties of the scenes, as it was steeper and abrupt with 'easy' tasks and smoother and gradual with most difficult ones. This explains the fact that its distribution flattens towards the facilitating tasks in Figure 7.2, as a steady portion of 4 year olds still adhere to it. In a more limited sense, enclosures can be seen as associated with age since they are found only among the 4 year olds, and their occasional use by a few older children in the Apples task is probably an artifact of misinterpreting the task demands[5].

However, the use of attachment and overlap reveals a different picture. As in the case of segregation, moving towards tasks with higher levels of partial occlusion output, the use of attachment by the older children is scarce (as all the alternative strategies in general). However, with more challenging tasks different age groups preferred it for different scenes[6]. Conversely, overlap was used by children of different ages across the whole range of tasks[7]. What was also particularly interesting was that overlap was mainly used by 4 and 7 year olds in scenes where partial occlusion resulted from partial enclosure (Boat and Vase scenes). Considering how different these scenes are in all other respects, its consistent use by the same age groups suggests that the transparency mode could perhaps be age-related only when the forms are structurally integrated.

1.2.2 Spatial devices

There has been a general consensus in previous studies that the route of development progresses from separating the forms on the horizontal axis to segregating them on the vertical and finally to uniting them with the implementation of the HLE device. All these studies assessed children's graphic devices with mnemonic tasks or scenes of 3D models. The present series of studies with scenes of structurally separate objects mainly involved copying tasks from a 2D array. Thus, the use of the spatial axes is more indicative of

[5] A considerable number of children at each age level claimed to have drawn the front apple. As Freeman *et al.*, (study 2, 1977) showed, this situation increases the transparency mode.

[6] Children of all age levels used it in the Boat task, whereas only the two intermediate groups adopted it in the Bottles, and mainly the 4 and the 7 year olds in the Apples and Houses tasks.

[7] It was favoured by the oldest of the children in the Bottles tasks, it appeared with comparable frequency from the age of 5½ in the Apples task and it was mainly used by the 4 and the 7 year olds in the Boat and Vase tasks, whereas it was used only by the younger of the children in the remaining ones.

children's differential view-specific ability rather than of the graphic rules they employ to translate depth information from the visual field on the pictorial surface.

The results showed that children preserved the spatial alignment of the items of the array. When the further object was presented on an elevated plane relative to the front one (Apples, Balls, Sunset), the majority of children used the vertical (or vertical-oblique) axis whereas, when it was laterally displaced (Bottles), they aligned it along the horizontal axis. An age-related shift in the use of the view-specific alignment mode was observed only in the Apples and Bottles tasks, as children of 4 and 5½ years of age used both the vertical and the horizontal axes with comparable frequency.

1.3 Cognitive style and gender

In general, most of the published literature on partial occlusion deals with manipulation of experimental variables, and age is the only subject variable of research interest. To the knowledge of the author, the three studies which investigated the effect of gender found no reliable difference between boys and girls in the pictorial representation of partial occlusion (Morra *et al.*, 1996; Radkey & Enns, 1987; Teske *et al.*, 1992), and only Morra *et al.*, (1996) examined the possibility that cognitive style might exert an influence on the tendency towards hidden-feature inclusion. Their argument was that the role of attentional resources is central in situations which elicit a cognitive conflict, and offered empirical support for the association of FD/FI in the encoding of partial occlusion arrays of similar objects.

On the basis of this recent finding and the well-documented gender differences in cognitive style, we examined the association of gender and FD/FI[8] with successful performance in each

[8] Cognitive style was assessed with the Children's Embedded Figure Test (CEFT: Witkin *et al.*, 1971) for primary school children and with the Preschool Embedded Figure Test for 4 year olds (PEFT: Coates, 1972). Children's scores below the median of their age group were classified as FD and those above as FI.

task separately. Analysis of correct and incorrect responses by gender and cognitive style using chi-square tests did not yield significant results in any task[9]. Although the absence of a gender effect was credible, the lack of association between cognitive style and the use of the HLE was unexpected, particularly for similar-objects arrays (Balls, Apples). While the present results are inconsistent with those of Morra *et al.*, this difference could have been caused by the different method used for the assessment of cognitive style (CEFT & PEFT here, Block Design subtest of WISC in Morra's *at al.*, study), or the differences in the dimensionality of the materials used. Perhaps, more planning activity and thus more attentional processes are required when copying from a 3D array in Morra's *et al.*, study than drawing from a 2D model picture in the present one.

2. DISCUSSION

As in previous research on occlusion, the present series of studies showed that a bias towards certain drawing styles is related to age as well as to the properties of the scenes, and that the between-task variation in performance changes at different stages of development. In this section an attempt will be made to interpret children's performance across the tasks, identifying the various inhibitory and facilitating factors which affect adopting a view-specific approach. Next the course of developmental change will be delineated, discussing the fluctuation in performance within each age group.

2.1 Task demands

Some researchers state that asking children to draw a partially occluded scene from memory cannot reveal their true abilities (Cox, 1978 & 1981). A number of studies have shown that the degree of visual realism is reduced in children's drawings from memory compared with

[9] Analysis of variance was also performed on the total number of correct drawings (0-9) with gender and cognitive style as factors. Only the main effect of cognitive style was significant ($F_{1,105} = 4.78$, $p<.05$). This does not need to contradict the results from the chi-square tests since it is generally agreed that field articulation affects general performance in drawing tasks (Witkin & Goodenough, 1981).

copying from a model, especially if this is a line drawing (see Chen 1985; Chen & Cook, 1984). The reason is that working from memory puts children in the adverse position to have to access directly their mental descriptions of objects or arrays.

Although the provision of a model can be beneficial, the present results show that it is not a sufficient condition, and that drawing from memory is not always inhibitory. At all age levels, one of the two tasks which elicited the worse performance was a copying task (Bottles vs. Bus) whereas one of the three tasks which yielded the best results was a mnemonic one (Vase vs. Balls, Sunset). Thus, the distinction between copying and mnemonic tasks is not sufficient to explain the present findings.

In addition, other researchers have endorsed the view that when the concept of hiding is intrinsic in a scene, children can understand better the task requirements – often missed form the verbal instructions – and comprehend the pictorial relevance of omitting the non-visible part of the occluded form (e.g. Arrowsmith *et al.*, 1994; Cox, 1985, 1991). *Prima facie*, finding that the Sunset and the Vase tasks were invariably the easiest across all age levels seems to corroborate this view. Conceivably, it is not the idea of hiding as such, but rather the 'ecological validity' of partial occlusion inherent in a scene (for it is virtually in the main idea of a sunset –or a sunrise– that only part of the sun is seen). Why then cannot children do the Boat task which is of equivalent familiarity, and can do the Balls tasks – especially those of fine occlusion – which have much less validity in this respect?

A plausible interpretation is that familiarity is helpful only when the graphic form of the occluded object is a simple one *per se* which does not impose further difficulties on children. To return to the issue raised in the beginning of the discussion, it is only under these conditions that a copying task is of equivalent status to a mnemonic task. Reducing a sun from a circle to an arc is neither conceptually difficult – in the present socially intelligible context – nor graphically demanding. The same holds true in the Vase task, where children have just to interrupt a simple line standing for the stem of a flower. So, when the formal simplicity of the occluded form is preserved and the mental computations are reduced, copying from a picture array can be as easy as drawing from memory, even for four year olds.

Now consider children's performance in the Balls tasks with fine occlusion with the two easiest familiar tasks (Sunset and Vase). The absence of performance differential at all age levels provides the most compelling evidence in favour of the view that keeping the occluded form relatively simple renders familiarity redundant. Also, no matter how intrinsically salient is the partial occlusion in the scene (Boat task), if the general drawing skill factor is high, children will perform poorer compared with a less contextualised task where this factor is reduced (Balls). Thus, the value of 'valid' tasks in promoting partial occlusion is conditional on production difficulties.

Similarly, the ability to benefit from the provision of the appropriate solution at hand is subject to the amount of the graphic skill required for its application; for having access to an executive scheme does not ensure view-specificity. This is the case for the Bottles task which was reliably the most difficult of all the copying tasks. Even if children manage to conceptually delete the non-visible part of the occluded bottle, the computational complexity required for the contiguous placement of its resulting irregular contour next to the near bottle is beyond the graphic skills of the present age range[10]. Previous studies on occlusion have also demonstrated that children's performance is markedly suppressed when drawing

[10] Possibly, the younger children (4 & 5½) might have not seen the need for this deletion, since they could distinguish the near from the far bottle by preserving their orientation contrast. On the contrary, the older children might have difficulty executing the appropriate strategy.

graphically complex objects compared to simpler ones (pigs in Teske *et al.*, 1992; cars & piggy/singers in Morra *et al.*, 1996).

Empirical findings suggest that generally children have difficulty analysing an object into its component parts (Shepp & Schwartz, 1976; Smith & Kemler, 1977). Further support has been provided by occlusion studies which showed that it is more difficult to omit part of an object's shape rather than delete either a discrete feature (handle from a cup) or totally eliminate it (see Chen & Holman, 1989; Radkey & Enns, 1987). It is difficult to say with confidence what is the origin of this difficulty (conceptual, perceptual, graphic). Yet, as Morra *et al.* (1996) point out, the proper co-ordination of various units of information requires a lot of cognitive processing, and such demanding condition like the present one, activate over-learned but inappropriate schemes (complete bottle).

At this point it is relevant to draw attention to a particular graphic difficulty that the Bottles task entails. As previous research suggests, young children produce less advanced drawings and fewer alternative solutions when executive constraints are in conflict with the implementation of the appropriate strategy (see Cox, 1992; Freeman, 1980; Karmiloff-Smith 1990; Van Sommers, 1984). Usually, the preferred order in drawing a vertical bottle is top to bottom. The fact that in the present case the occluded edges of the further bottle are asymmetrical requires children to proceed in a broken sequence (left side-right side) to accurately align it with the near tilted bottle. The same applies to the Houses task, since houses are usually drawn from the leftmost edge in counter-clockwise direction. Presumably, when the implementation of the HLE device deviates from a well-practice drawing sequence, children are less likely to make a view-specific drawing.

Even if such graphic problems are minimal, a case which exemplifies the importance of the non-linguistic context in communicating the task requirements is the Apples task. Here we have a case where although the appropriate solution is provided and its application does not introduce production difficulties[11], the degree of visual realism in children's drawings is still limited. Children's performance in arrays with similar objects (Balls & Bottles) supports the view that in the Apples task children have difficulty understanding that they have to draw the occluded object. This problem seems to be attributed to the fact that the occluding apple was pre-drawn.

The inhibitory effect of this condition seems to depend on the fact that an array of identical items is used. Support for this contention is provided by children's far superior performance in the Sunset and Vase tasks where their graphic demands are as minimal, as in the Apples task, and the occluding form pre-drawn. This finding has two implications. First, the superior performance in the Balls compared to the Apples task disconfirms an assertion previously made (Cox, 1986; Morra *et al.* 1996) that, in a similar-object array, having children first draw the front object as complete activates an inappropriate graphic scheme for the form at the back. Second, with a similar-object array, having the occluding item pre-drawn on the children's papers impedes the correct interpretation of the task requirements. Such a misleading context undermines the full operation of the helpful qualities of the scene in fostering the mental and the graphic deletion of the non-visible part of an occluded form.

Finally the very poor performance in the Bus task (in comparison with any other) is not surprising, since children have to overcome both conceptual and graphic difficulties. In this situation, children have to rely totally on their conceptual knowledge in describing the scene

[11] We cannot plead production difficulties in the execution of the HLE device since children have to draw an arc to represent the visible part of the occluded form in the Apples as well as in the Balls and Sunset task. And although the 'social validity' of the Sunset task can foster its use and possibly

to themselves. However, neither the instruction nor the theme in itself guides children as to *how* they should conceptualise it. First, the instructions are neutral, as they do not stress the spatial differentiation of the near and far wheels. Second, the relevant information is neither intuitive in a model, nor an intrinsic property of the scene. Under such conditions, the conceptual deletion of the non-visible part of the occluded wheels is difficult to be initiated for the graphic omission to follow. This can explain the failure of the 8 year old children whose overall ability in using the HLE device suggests that their deficit is not at the level of graphic skill.

2.2 Developmental changes

Despite the different properties of the scenes, the prevalence of complete forms in the drawings of the four year olds supports the view that very young children are more vulnerable to their conceptual knowledge (Piaget) or their internal model (Luquet 1927/1977) which comprises the protopical or canonical (Freeman) attributes of forms. In general, we can claim that the four year olds are trapped in object-centred descriptions even when confronted with a drawn model. According to Piaget and Inhelder (1948/1967), this is the typical behaviour of preoperational children whose figurative schemes of objects are immune to their stationary visual appearance. Consequently, objects in an array are encoded as bounded and unitary entities with no spatial overlap. When it comes to pictorial representation, separation is the only viable solution which not only preserves the completeness of the figurative schemes – or the *Cute Gestalt* of the forms, as Kuttner and Reith (1995) express it – but conserves according to Goodnow (1977) the 'psychological space of objects'.

---

explains this task difference, the same cannot apply to the Balls tasks which are comparable in this respect to the Apples task.

This view has been embraced by Cox (1986), who claimed that, if young children's problem is just to preserve the conceptual completeness of forms, we would expect to find various structural solutions with some frequency (i.e. attachment, overlap). The present results fully support these contentions, as the separation bias was prevalent among the four year olds in the majority of tasks (except for the Sunset and Vase tasks), in contrast with older children who prefer other alternative strategies to separation. Moreover, the fact that this device is not confined to the drawing mode (picture selection: Cox, 1981, study 3) or to scenes involving the loss of a form's completeness (spatial adjacency: Reith, 1987; Piaget & Inhelder, 1966/1971, p. 54, Houses task in the present study), makes us more confident to claim that it is an all-purpose device, symptomatic of young children's 'synthetic incapacity' (Luquet, 1927/1977).

This is not to say that very young children are totally impervious to the spatial relation between forms, as we found that they could preserve some of the scene's view-specific properties in the spatial alignment of forms. However, this level of representational faithfulness does not necessitate the radical transformation of the figurative schemes of forms required for their spatial integration (see Dennis, 1992 and Morra, 1995). Likewise, the immunity of the four year olds to task variations is doubtless relative rather than absolute as there are some few occasions when they can be prompted to use the HLE device. Yet, as previous researchers have claimed (Barrett *et al.*, 1985; Light 1985), being cognitively and graphically restricted, their susceptibility to experimental manipulations is more limited than older children.

Turning to the performance of children at intermediate ages in the present age range, we obtained a rather variable drawing style which suggests that they have 'reserved ability' towards visual realism. As Piagetian theory predicts, with the emergence of operational thought children are able to transform their figurative static schemes and incorporate view-specific properties. However, since this ability is not well established, a number of factors can either lead or mislead children in recognising the need to modify their figurative schemes.

When provided with a model drawing to copy, where the graphic solution is easy to implement, both 5½ and 7 year olds perform with visual accuracy (Balls, Sunset tasks). In copying tasks with execution difficulties (Bottles, Houses), children cannot benefit from the accessibility of the appropriate device, whereas naturally the performance of 5½ year olds is notably more constrained compared to 7 year olds. In mnemonic tasks, both age groups regress to primitive solutions (complete forms) if the context and the content do not explicate the task requirements (Bus), and if the graphic equivalents are hard to deploy (Boat). On the contrary, when these demanding conditions are mitigated, both age groups demonstrate visual realism even in the absence of a model (Vase). Finally, the performance of the 8 year olds is less variable, since the use of the view-specific device prevailed in the majority of cases, while, even under obstructive conditions, complete forms were never dominant and segregating tendencies were readily suppressed. In this respect they seem to have gained more procedural knowledge and extended the use of the appropriate device to more situations (see Morra *et al.*, 1988 & Morra *et al.*, 1996).

Finally, Freeman's *et al.* (1977) model of drawing styles,which gradually approximate partial occlusion, does not show much empirical validity. The results provide no evidence for a lawful transition from one strategy to the next. The tendency to spatially separate the forms is a constant challenge with an age-related trend in its use. For younger children it seems to be an all-purpose graphic strategy which can be moderately suppressed only under certain propitious conditions. With age, its prevalence declines, and children use more adaptive and flexible depiction strategies. Enclosure, contrary to Freeman's findings, does not constitute an intermediate stage between separation and partial occlusion, whereas attachment and particularly overlap are scene-dependent strategies.

3. CONCLUSION

The present series of tasks have revealed a number of factors which exert either a facilitative or an inhibitory influence on the use of the view-specific drawing device. First, the social validity or "human sense" (cf. Donaldson, 1978), conferred by the material used, proved to be not as significant as Cox has claimed. The results lend support to the argument of Light and Foot (1986) that the familiarity of the materials *per se* is not a sufficient condition. As a number of studies have shown, it can even inadvertently promote canonical representations (Ingram & Butterworth, 1989; Krascum *et al.*, 1996; Morre, 1987; Reith, 1988). Above all, what needs to be meaningful is the concept of partial occlusion inherent in the scene.

To those who attribute children's failures to a "comprehension deficit", this manipulation promotes the understanding of what is about the array that the experimenter is alluding to (Davis, 1983, 1984 & 1985b; Davis & Bentley, 1984; Cox, 1986, chapter 2). For others, who are reluctant to credit children with such an intentional and communicative spirit, the validity of the occlusions in the arrays chosen works at the level of how children know a scene to be (Crook, 1984 &1985). So, if partial occlusion has functional salience (in the same manner as a form's completeness), it is more likely to be encoded in children's internal description of an array's structure, and ultimately to suppress the tendency towards hidden-feature inclusion.

Furthermore, in agreement with previous studies, the task demands can be also successfully communicated with the provision of a 2D model, which is neither familiar nor 'emdedded' in an explicit context (meaningful game, rigorous instructions). However, a number of children, who can reflect upon the view-specific properties of objects, will still have difficulty in drawing the requisite scene if they can not handle graphically the partially occluded shape. Thus, the formal simplicity of the visible boundaries of surfaces seems to be another contributing factor in representing a visually realistic intention, when this is present. In this sense, the results support the production-deficit claim that the lack of technical competence and organisational skills precludes a visually faithful drawing (Freeman 1980).

Overall, the present findings support the general view that the structural features of the array (separation, integration and continuity of the occluded), the physical properties of the items (like dimensionality, symmetry, canonicality) along with their representational salience affect the encoding and subsequently the depiction of spatial relationships. This is encapsulated in the argument of Teske *et al.*, (1992) that "the depiction of depth is not learned independently of such object characteristics" (p. 33). A number of factors differentiate the various tasks in terms of the accessibility/inaccessibility of a view-specific representation. Yet, there is no doubt that the developmental changes in children's drawings of partial occlusion take the form of an increasing sensitivity to these factors, reflected in the flexibility of their responses, rather than the presence or the absence of discrete abilities.

The studies reviewed in this chapter certainly show that the same children can eliminate the hidden feature from some occluded objects and not from others, and that, in agreement with previous research (Cox, 1981; Light, 1985; Smith & Brown, 1982), the use of partial occlusion with HLE is not in all cases confined to children over 8. In this respect Arrowsmith *et al.*, (1994) and Costall (1995) are right to insist that Luquet's (1913, 1927/1977) well-knowing saying "children draw what they know rather than what they see" should be interpreted not as a matter of fact but simply as a *tendency*. Similarly, the terms "intellectual realism" and "visual realism" should not be taken as absolute, but rather as generalised descriptions of clusters of phenomena which themselves need to be explicated in terms of task demands and processing factors. Costall (1995) proposes that we need to redefine development in terms of "the resourceful deployment of pre-existing abilities" (p. 21), a view shared by Light (1985) who claims that we should conceptualise age changes "as a reflection of shifting priorities within a hierarchy, rather than in terms of newly emerging capabilities" (p. 227). Their claims are fully supported by the finding that children from the age of 4 used a variety of strategies and in this sense they are not limited in their graphic means available. Thus, current research practices should be devoted to isolating the factors that impel the child towards one type of drawing or another.

In the present series of studies we identified some of the factors which relax children's conceptual obstinacy towards complete forms and foster visual realism. However, we cannot draw any conclusion as to the exact nature by which these factors interact with each other at different age levels. More rigorous experimental designs (see Lewis *et al.*, 1993; Morra *et al.*, 1996) are needed in the future to examine the independent and accumulative influence of these factors at different levels (most to least facilitating/inhibitory) cross-sectionally.

CHAPTER VIII
A DEVELOPMENTAL REVIEW ON THE ACQUISITION OF SPATIAL AXES

1. THE PIAGETIAN THEORY

Several seminal ideas about the development of representational space were presented by Piaget and Inhelder in *The Child's Conception of Space* (1948/1967). According to their theory, conceptual or representational space emerges around the age of two when the child starts building spatial concepts to understand and symbolise spatial relations. In their original formulation, Piaget and Inhelder claimed that children construct three systems of spatial concepts beginning with the mastery of topological concepts, proceeding later to the construction of projective and Euclidean concepts. Regardless of certain controversies[1] and subsequent reformulation of their original ideas, their developmental model highlights the difference between these three systems in conceptualising the properties of space.

In the beginning, very young children include in their representations only topological features of space which preserve relationships of proximity, continuity and inclusion. These concepts tap an understanding of "next to", "on" and "in" spatial relationships between objects but they do not provide any information about orientation, direction, distance and size. Later, they develop projective concepts, where the representation of spatial relationships incorporate changes in the point of view and reflect an understanding of left-right, in front-behind and up-down relationships. Children eventually develop a Euclidean conception of

[1] Piagetian theory implies that the three systems of spatial concepts develop sequentially. Empirical data, however, have challenged this view and initiated a lot of debate (see Beilin, 1989 for the sequential development of spatial concepts).

space, a system of integrated horizontal and vertical co-ordinates on a three dimensional grid with which they perceive and represent the physical world. The construction of this system creates a frame of reference that acts as a fixed "container" as objects or people move in reality or in imagination to different locations. It is comprehensive and generalised because:

> *...the essential character of a reference system does not reside in the choice of stationary reference objects, but in the possibility of co-ordinating positions and intervals without limit...* since *...the "container" differs from its "contents" in that these relations are not confined to the objects at a particular point...*
>
> (Piaget & Inhelder 1948/1967, p. 376-377, underline added)

Piaget and Inhelder traced the development of a spatial co-ordinate system, assessing children's ability to make reference to the natural axes. The acquisition of the concept of horizontal and vertical co-ordinates was studied by two methods. The standard test of horizontality assessment was the water-level task (WLT) where children were presented with two-dimensional line drawings of bottles tilted at various angles and were requested to imagine the level of the water in them, assuming that they were half-filled. In a predictive-production version of the task, children had then to draw the water level, whereas in a perceptual-recognition one, they had to select the picture which showed the water level as horizontal. Similarly, the use of the vertical axis was assessed by asking children to predict and draw vertical forms on oblique baselines; like trees, lampposts and houses on a mountainside (MST), chimneys on slanted roofs, and plumb-lines suspended from inclined posts or hooks attached on vehicles (PLT).

Piaget claimed that children's success in those tasks depends on their ability to relate the target stimuli to the spatial co-ordinates of an external frame of reference. The use of the term *external* implies that these reference points lie beyond the dimensions of the stimulus array itself. According to Piaget, the conceptual construction and the use of such reference system are advanced abilities symptomatic of the successful operation of deductive inference (geometric operations are involved in the acquisition of a Euclidean system). In the WLT, children, noting that the water-level always parallels an external horizontal cue like the table top, are led to the discovery of an invariant relationship, from which the accurate orientation of the water level follows. Likewise, finding a vertical cue external to the array, like the sides of paper, children discover a fixed spatial referent to judge and represent verticality.

The Piagetian theory claims that young children fail these tasks because their judgements are guided by local cues and their spatial concepts are determined by the particular configurations presented. As Bryant (1974, p.12-13) later expressed it, young children are 'relational coders' and older children are 'absolute' ones. In the WLT, the 'relational-coders' use the axes provided by the container while in the MST/PLT, they rely on the sloping edge of the mountain/post to judge and represent these natural horizontals and verticals in respect. On the contrary, absolute coders are not contextually dependent and rely on the invariant co-ordinates as the only absolute spatial reference in judging and representing orientation.

Piaget, based on the different patterns of errors that children made at different ages, concluded that the acquisition of a Euclidean system develops gradually in a stage-like fashion during early to late childhood. He proposed three primary stages in the development of the horizontal and vertical co-ordinate system with two sub-stages in each of the last two. Children at *Stage I* lack the idea of surfaces, lines, and planes. Consequently, in the WLT, they represent the water in a vessel by scribbles that often cross the walls of the container or - after acquiring better motor control - as a round ball with no indication of its planar surface. Accordingly, in the MST, the forms (houses/trees) are drawn either parallel to the slope or within the mountain's region, using it as a background. For Piaget, these cases exemplify topological representation since children are concerned in depicting the water as a mass inside a container, or houses and trees in proximity with the mountain's surface, with no regard of orientation and direction. *Stage I* lasts approximately until the age of 4 years.

At *Stage II,* children's representations are guided by figurative thought, a process which is determined by the structural components of the immediate field. Figural factors activate the schemes of space co-ordinates which subsequently derive from the axis provided by the stationary objects, such as the frame of the bottle or the mountainside. Children at *sub-stage IIA* draw the water-level always parallel to and at the base of the container, and trees perpendicular to the mountain's slope. So, with respect to the WLT, children are able to represent the water as a planar surface, but they cannot recognise its movement inside the container. Likewise, in the MST, children can anchor the forms on the mountainside, but they do not understand that they grow always vertical independent of the hill's slope.

At *Stage IIB*, they abandon the 'bottom-parallel rule' in WLT and recognise the change in the liquid's position relative to the tilt of the jar. However, they still fail to understand that it remains horizontal. According to Piaget, children's performance in this sub-stage is best described as a persistent trial and error attempt, which is preliminary to the subsequent construction of operational systems. This stage lasts from 5 to 7 years of age. Finally, children enter the concrete operational *Stage III* whose onset is between 7 to 8 years of age, "but not earlier" (Piaget & Inhelder 1948/1967, p. 379). At this stage children gradually discover the horizontal and vertical spatial co-ordinates. What distinguish *Sub-stage IIIA* from *IIIB* in the WLT is that, at the former stage, children are accurate in their predictions only when the vessel's position favours the axes (upright, inverted and resting on its side) whereas at the latter stage, they are accurate in all oblique positions. *Sub-stage IIIB* is said to occur about the age of 9 "with a few laggards at 12" (p. 408).

Piaget maintained that children who have not acquired a conceptual schema of the horizontal and the vertical co-ordinate do not know the physical principles pertinent to his tasks, and they cannot benefit from any demonstration and training. In his testing procedures, children were also asked to provide explanations of their drawings and were given feedback and/or training phases in which the physical facts were made evident to them. Children below the age of 9 revealed no knowledge of the principles and failed to discover them from visual inspection. Piaget regarded this 'perceptual' data to be the more compelling evidence for the absence of conceptual spatial schemes. This assertion implies that if the necessary cognitive structure - in this case an Euclidean conceptual system - has not been developed, the perceptual input cannot be accurately processed[2].

From the Piagetian theory certain predictions can be made. First, if the acquisition of mature spatial concepts culminates by middle childhood (12 years), a Euclidean reference system should be fully developed by late adolescence and virtually universal by adulthood in the normal population. Second, if the Piagetian spatial tasks depend upon a general, conceptual spatial process which develops in a stage-like fashion, then performance should be highly and significantly correlated across verticality and horizontality measures, and the same developmental progression should be evident in tasks with different performance requirements. Third, Piaget claimed that the conceptual construction of a reference system is a prerequisite for the mental co-ordination of spatial information and the subsequent discovery of the physical principles. As a consequence, the acquisition of spatial co-ordinates is considered as a necessary and sufficient condition for understanding the physical facts immersed in the Piagetian tasks.

[2] Piagetian theory viewed the cognitive and the perceptual factors as inextricably interwoven and claims that the former precede the latter in the process of acquisition of the physical principles. The operative schemes affect the assimilation of sensory input, and the visual information is misinterpreted if the necessary cognitive structures are missing. Simply, one should know to be able to recognise.

2. CHALLENGES TO PIAGET'S FORMULATIONS

Over the past 30 years, research on the Piagetian tasks has demonstrated that a large proportion of the adult population has considerable difficulty with those tasks. Empirical data suggests that approximately 40% of older subjects err on the Piagetian water-level tasks and fail to understand the principle (see reviews by Kalichman 1988; Liben 1991a; Pascual-Leone & Morra, 1991 and findings by Kalichman 1986; Liben 1991b; McAfee & Profit 1991; Rebelsky 1964; Signorella & Jamison 1978; Thomas *et al.,* 1973; Wittig & Allen 1984). Although there has been less research on verticality, the data leads to a similar conclusion (Liben 1978; Liben & Golbeck 1984; 1986; Meehan & Overtone 1986).

The first study, involving older samples by Rebelsky (1964), also revealed that females were less accurate than males on Piagetian tasks. Since then, research on gender differences has dominated the literature and Rebelsky's findings have been replicated by subsequent investigators who have reliably reported male superiority in the representation of the horizontals (WL) and the verticals (PL) of the physical world. Results across developmental periods suggest that gender differences seem to emerge at the time of concrete operations, at which time onwards, boys' performance improves relative faster than girls' (DeLisi, 1983; Liben, 1974, 1975; Liben & Golbeck, 1980; McGillicuddy-DeLisi *et al.,*1978).

Moreover, considerable variation has been observed in the correlational patterns of performance between the horizontal and the vertical concepts and within each measure. Empirical evidence suggests that performance depends strongly upon the context where the tasks are presented (Liben & Golbeck 1980, 1984, 1986; MacKay *et al.,*1972; Reith *et al.,*1994). These unpredictable findings have serious implications for the Piagetian theory. As Thomas & Lohaus (1993) assert:

> *" a theory that presumes that a spatial system is available by about age 9, and makes later cognitive development contingent on it, is certainly compromised by the frequency with which adults fail on the water-level task" (p. 12).*

All these empirical results indicate that these spatial problems are rather complex and cause difficulties to a wider developmental spectrum than was originally thought (Vasta & Liben, 1996). Like other Piagetian phenomena, the traditional assessment of vertical and horizontal concepts have initiated a lot of interest in the study of variables related to task performance, producing over 100 studies in the published literature. Researchers, trying to disentangle performance difficulties from operational deficiencies, have identified several components interwoven in the WLT and vertical tasks. Next, we will present these factors separately covering the principal studies which have systematically explored them. We will then conclude by reviewing the dominant theoretical models developed in the last decade, which attempt to integrate these components, and offer a comprehensive account of performance on the spatial Piagetian tasks.

3. THE KNOWLEDGE OF THE PHYSICAL PRINCIPLES

Some researchers have tried to investigate whether the failure on the Piagetian tasks by a substantial number of older children, adolescents and even adults reflects an underdeveloped abstract system of spatial co-ordinates or a lack of knowledge of physical behaviour of water and plumb-lines. Piaget and Inhelder were aware of the dual nature of the tasks that involve physical and geometric concepts:

> *On one hand, the concepts of vertical and horizontal are by nature physical ...Yet the elaboration of these concepts introduces a question independent of physics...and this is the development of a co-ordinate system as a simple tool of geometric orientation (Piaget & Inhelder, 1948/1967, p. 380).*

A number of studies have explored the relationship between verbally expressed understanding of the physical principles and performance on the spatial tasks with children, adolescents and adult population. Thomas and Lohaus (1993), testing 11 to 16 year olds, found that good performance on the WLT preceded verbally expressed knowledge but the reverse was true for the PLT. Collectively and despite the variation in the procedures used to measure performance (psychophysical apparatus: Thomas & Jamison, 1975; photographs in rapid succession: Howard, 1978; drawings: Liben & Golbeck 1984; Thomas & Lohaus 1993; Pulos 1997), there is general agreement that verbal expression of task principles is related to task performance as well as to other measures of spatial ability (Lohaus *et al.*, 1996).

However the relationship is not always strong. There have been cases where subjects (1) know the principles but perform poorly (Liben & Golbeck, 1984), or (2) do not know the principle but perform accurately[3]. Myer and Hensley (1984) found that 37% of their adult subjects who performed accurately did not know the water-level principle whereas 28% of those who correctly stated the rule were unable to apply this knowledge when performing the task. So, although knowledge of the principle of invariance was significantly related to performance, it was far from perfect as a predictor. These results show that there are cases where the knowledge of the principle is not directly accessible as well as conditions where accurate performance is not conditional on this knowledge.

If subjects' performance is particularly hindered by lack of knowledge of the physical behaviour of water and plumb-lines, it was hypothesised that their performance should be especially susceptible to methodological manipulations which eliminate or reduce this physical component. Two methods have been generally used to evaluate this possibility. The first approach included a training phase at the testing procedure and compared performance on the Piagetian tasks before and after training. The second method involved

[3] Howard (1978) suggested that his adult subjects could perform accurately on a forced-choice recognition task and a drawing task without demonstrating knowledge of the water-level principle (as assessed by their verbal reports). He claimed that a free choice recognition task can discriminate better between adults who know the principle compared to those who ignore it.

decontextualised tasks where horizontal and vertical lines were embedded or attached on titled frames.

3.1 Training procedures

Earlier studies have reported the positive effect of training on performance with young children but the magnitude of this effect appears to be limited. Improvement was either limited to post-test stimuli with shapes or orientations identical to those used during training (Beilin *et al.,*1966; Robertson & Youniss 1969; Sheppard, 1974) or to post-test stimuli with non-oblique orientations (Smedslund, 1963). Recent studies appear to confirm the method-dependent and task-specific nature of the training effect. McGillicuddy-DeLisi *et al.,*(1978), testing 6-10 year old children on the WLTs, report a significant pre-to-post inspection phase improvement only for the 6 year olds. However, this was confined to outlines in which one of bottle's axes was horizontal. Pennings (1991) succeeded in eliciting substantial improvement in performance of 7 to 8 year olds with tilted bottles, but this was not transferable to other spatial tasks, whereas Randall (1980) obtained a satisfactory generalizability of training with 7 year old children.

Comparable results were obtained with adult samples. Several studies demonstrated that subjects' performance is resistant to improvement even if they were given the chance to discover the physical principles through observation of the effects of the context's transformation. (Liben, 1978; Liben & Golbeck, 1980; Sholl & Liben study 2, 1995; Thomas *et al.,*1973). However, there were cases where training resulted in better performance (Barsky & Lachman, 1986; Liben & Golbeck, 1984). The inconsistent results of the training effect with children and older samples can be explained by the nature of the training procedures and the abilities of the subjects tested. Practice or visual demonstration seems to be completely ineffective (Liben, 1978; Liben & Golbeck, 1980; Sholl & Liben, 1995), or simply reduces the average size of errors without enhancing accuracy (Thomas *et al.*,1973). On the contrary, when an explicit verbal statement of the principles or corrective feedback follows demonstration, it results in a significant increase of accuracy (Liben, 1978; Penning, 1991; Randall, 1980). Also, the effect of intervention seems more beneficial for transitional children who could demonstrate some partial understanding on pre-test.

These findings have the following implication. The fact that young children and some adults can benefit from either direct instruction on the physical principles or perceptual feedback suggests that they might have the conceptual prerequisites for assimilating the relevant information and that certain instructional intervention may facilitate the application of an existing and potentially available accurate conceptual framework.

3.2 The "non-physical" tasks

An alternative way to investigate the effect that knowledge of the physical principles could have on demonstrating the acquisition of the spatial concepts was to administer tasks that remove this context altogether. Liben and Golbeck (1980) used a 'non-physical' version of the WLT and the PLT, where children of 8 to 16 years of age were asked to draw from memory lines that were "straight across" (horizontality assessment) or "straight up-and-down" (verticality assessment) inside tipped rectangles. The superiority of the non-physical versions in eliciting better performance relative to the traditional physical ones was statistical reliable and evident at every grade level. Perner (Perner *et al.*, 1984) empirically tested this finding with a younger age group (3-7) using a copying drawing task. He obtained similar results although the removal of the physical context had a facilitating effect only on the horizontality measure. The previous findings were also replicated with an adult population which performed near ceiling on the abstract versions of the spatial tasks but had difficulties with the standard Piagetian ones (Liben & Golbeck 1986). Thus, on the whole, the physical context of the Piagetian tasks seems to impose additional difficulty on performance even with adults.

The same issue was investigated by McGillicuddy-DeLisi who used the crossbar task as an alternative non-liquid procedure for the assessment of horizontality (McGillicuddy-DeLisi *et al.,*1978). The task involved a balanced crossbar attached by a pivot to a support bar, and children had to predict the orientation of the crossbar after the support bar was adjusted to oblique positions. She found that when performance was assessed on the mean number of correct responses, using a pass/fail criterion, the two tasks did not differ among children of 6, 8 and 10 years of age but that adult's performance was near ceiling on the crossbar task. However, when performance was assessed using mean angular deviation scores, the crossbar task elicited significantly smaller deviations than the WLT at each grade level while the task differential was not significant even by the age of 10 (DeLisi, 1983; DeLisi *et al.*, 1995).

Liben (1991b) suggested that when horizontality is assessed by the crossbar task, the less deviant scores observed among children, and the ceiling performance found in adults in DeLisi's studies need not to be attributed only to the removal of the physical component related to the behaviour of liquids. She pointed out that these tasks differ on a more crucial feature; that of the presence of an embedded context in the WLT and the absence of that context in the crossbar task. Her hypothesis was empirically confirmed, as adults' performance was not significantly better when they were tested with a crossbar embedded in an immediate oblique frame as opposed to the traditional WLT. However performance improved significantly with a disembedded crossbar compared with either of the former tasks. This last issue will be discussed in the following section.

4. VISUAL PERCEPTION

4.1 Illusory tilt hypothesis

Some researchers have argued that lines and edges embedded in tilted containers are not accurately encoded at the level of primary perception. Ford (1970) found that children made errors even when asked to match the WL in a 3D vessel selecting a picture from a set. According to one view, subjects are experiencing an illusory tilt which cause horizontal and vertical lines to be misperceived as tilted in a direction opposite to the tilt of the frame (Sholl & Liben, 1995).

On the basis of empirical findings, Sholl and Liben claimed that all their adult subjects experienced illusory tilt but good performers utilised Euclidean spatial concepts to overcome it. A number of studies investigated the directional errors made with perceptual tasks. Although these studies report that subjects' choices were biased to one direction, they disagree as to which direction this is. Howard (1978) and Amponsah (Amponsah & Krekling, 1994) found that, in a picture-selection task, adults were biased in choosing water-levels that climbed up towards the spout of the vessel consistent with the prediction from the illusory tilt model. However, McAfee and Proffitt (1991) found the opposite directional bias; water-lines tilted in the same direction as that of the vessel's tilt. It has to be noted that, depending on the differences in the testing procedures and instructions, some of these perceptual tasks might have encouraged more the use of conceptual strategies[4] rather than the use of purely perceptual processes.

4.2 Disembedding skills and cognitive style

Another approach recognising the role of visuo-spatial perception in Piagetian tasks focuses on disembedding skills. Research into this variable was started initially in order to explain the well-documented gender differences in the WLT. The basic argument was that, to succeed in the Piagetian tasks, subjects have to disregard the proximal visual field (bottle or inclined plane) and to locate the spatial axis in the presence of competing perceptual cues. Witkin (1978), in his analysis of cognitive style, contended that the ability to attend to a task or problem, independent of its context, characterised individuals who are field independent. Two methods have been adopted to examine the extent to which disembedding skills underlie performance on Piagetian spatial tasks.

4.2.1 Correlational studies

If Piagetian tasks require an ability to 'disembed' water lines or plumb lines from the immediate field, performance on the these tasks should be significantly related to field independence. Numerous studies have investigated the relationship between performance on Piagetian tasks and cognitive style using either the rod-and-frame test (RFT) or the embedded figures test (EFT) (reviews in Pascual-Leone & Morra, 1991; findings in Abravanel & Gingold, 1977; DeLisi, 1983; Liben, 1978; Signorella & Jamison, 1978). The correlational patterns are similar, although correlation estimates seem to vary widely depending on the method of assessing the spatial concepts (paper-and-pencil version, real 3-D test material, psychophysical apparatus) and cognitive style (RFT, EFT).

The magnitude of correlations in performance between paper-and-pencil versions of spatial tasks and EFT ranged from .57 to .70 for children of 7 to 12 years of age, supporting the contention that accuracy on verticality and horizontality tasks was associated with cognitive style status. In addition, factor-analytic research provides evidence that the EFT is structurally similar to the Piagetian task (Goodenough *et al.,*1987; Pascual-Leone, 1969; Witkin, 1978; Witkin & Goodenough, 1981).

[4] Researchers have identified two cognitive strategies related to the WLT: one is that the water level is parallel to the base and the other that it moves towards the spout of the vessel. These result in opposite

Liben (1978) investigated the possibility that the degree of overlap in performance on both tasks might be attributed to the general spatial abilities that the tests tap and not to disembedding skills in particular. By separately measuring general spatial ability (Guilford-Zimmerman test of spatial orientation), the correlations remained high and significant when spatial ability was partialled out. Abravanel (Abravanel & Gingold, 1977) provided further support for the reliable relationship between performance on Piagetian spatial tasks and field-dependence tasks, controlling for general intellectual maturity (IQ scores). However the results from both studies are limited to a restricted age group (17.6 and 10.2 year olds respectively).

When the pattern and magnitude of correlations are analysed in terms of age and gender, there is less consistency between studies. In general, the correlations reported for adults and children of 9 years and above are significant. However, with younger age groups results are inconclusive. DeLisi (1983) found highly significant correlations for 7 year olds whereas Van Esch (1978) obtained the opposite result. Pascual-Leone (Pascual-Leone & Morra, 1991) maintained that the correlations in younger children are less strong or even not significant. Few studies used a wide age range of children younger than 7 years of age (Fabian, 1982; Johnson, 1982). However, since the correlation estimates available are for the whole sample, no conclusion can be drawn about differences in the magnitude of correlations between age levels.

Turning to the gender variable, significant correlations were reported for both sexes with two exceptions; the studies of Signorella and Jamison (1978), and Otani and Leonard (1988). In the first study, performance on WLT was significantly correlated with EFT only for boys whereas the opposite result was reported in the second study which assessed cognitive style with Gottschaldt figures (similar to EFT). The results on the magnitude of correlations

---

directional errors.

between the sexes are also inconclusive. There were studies where correlations were higher for boys than for girls (EFT: Liben, 1978; RFT: Abravanel & Gingold, 1977, study 2) and others which demonstrate the reverse trend (RFT: Abravel & Gingold, 1977, study 1; DeLisi, 1983).

4.2.2 Contextual misleadingness

The previous results seem to indicate that the Piagetian tasks have a general misleading structure which could disrupt or hinder the conceptual representation of spatial co-ordinates because the proximal frame of reference (oblique straight sided frame) is in conflict with the true environmental co-ordinates. A number of studies manipulated the degree of conflict between the immediate and the distant spatial field in WLT by systematically varying the shape of the container (Abravanel & Gingold, 1977; DeLisi *et al.*,1995; Liben, 1978; Russell & Maxwell, 1980; Thomas & Jamison, 1975; Vasta *et al.*,1994; Willemsen & Reynolds, 1973).

In general, children's performance improved on WLT when the conflicting straight lines of the container were eliminated or totally reduced. Abravanel found that a triangular bottle was significantly more difficult than either a bell-shaped bottle or a cylinder with 9½ year olds, and DeLisi confirmed the facilitating effect of a spherical container in oblique position with children of 6:9 to 10:8 years of age. Russell observed that when the container was devoid of any straight lines (sphere) children of 4½ to 6½ years of age perform at ceiling ; however when the container had straight lines either at its base (cylinder, fish-bowl) or at its sides (tube) there was no difference in the degree of accuracy. Similar views are held by other researches who claim that the potential advantage of a non rectangular bottle can be attenuated by the presence of straight lines anywhere on its shape (neck, sides, base) (Liben, 1978; Vasta *et al.*,1994; Willemsen & Reynolds, 1973).

Nevertheless, some inconsistencies among the studies in the magnitude of improvement with a less misleading context were observed and can be explained on the basis of the method used to judge performance (criterion vs. absolute deviations). It seems that, providing a frame whose axes do not compete with the natural environmental co-ordinates reduces the magnitude of error without necessarily increasing accuracy, particularly among the younger children (DeLisi *et al.*,1995)[5]. Also, researchers observed similar developmental changes with less conflicting frames of references, while the differences of the mean angular deviations between successive age groups were considerably diminished (DeLisi *et al.*,1995; Thomas & Jamison, 1975).

5. THE ASPECT OF MOVEMENT

5.1 Mobile versus immobile reference systems

An aspect central to the Piagetian theoretical arguments was that the developmental stages for the acquisition of the horizontal concept parallel those for the acquisition of the vertical, since both are abstractions of the same spatial conceptual system[6]. This premise has been challenged by a number of researchers who, using different methodological approaches found a developmental discordance in the acquisition of the spatial concepts (drawing responses: Beard, 1964; Liben, 1978; Perner *et al.*,1984; verbally expressed understanding of the physical principles: Thomas & Lohaus, 1993, study 2, relevant discussion in p. 125).

Although Beard (1964) supports the idea that the understanding of the vertical as a natural dimension precedes that of the horizontal, studies with geometric tasks provide no evidence of a developmental lag in the development of the two co-ordinates axes as abstract dimensions (Liben & Golbeck, 1980, 1986; Perner *et al.*,1984). MacKay (MacKay *et*

[5] The gain in the percentage of accurate responses is minimal among 7 year olds (rectangular: 0% v s round: 4%) moderate among 8½ year olds (rectangular: 10% and round: 33%) and maximal with 10½ year olds (rectangular: 27% & round: 62%).

al.,1972) proposed that the differences between the spatial tasks may not reflect a developmental incongruity in the relevant concept formation, but rather the lack of content equivalence.

On the basis of empirical evidence, several authors (Freeman, 1980; Liben 1975, 1978; MacKay *et al.*,1972; Perner *et al.*,1984) suggest - contrary to Piaget- that the standard materials used for the assessment of the vertical and the horizontal are not always operatively analogous. According to their views, the WLT and the PLT involve a mobile reference system where the orientation of WLs and PLs should remain unaffected by the spatial displacement of their proximal frame (bottle, van). As Freeman (1980) claim, within such a context, the aspect of movement is conceptually salient and has representational precedence over the body of water or the flexible plumb-line. On the contrary, the mountainside and the chimney/roof tasks, used for the assessment of verticality, provide a stable reference system where children should have less difficulty to represent the main axis of solid forms on immobile stand-lines. On the basis of these views, variation in performance is expected between the two measures as well as within verticality measure, particularly when the pertinent concepts are assessed with predictive tasks, as the demands on mental imagery will be increased.

In fact, according to the between-measures performance, the invariance of verticality was just as difficult as the invariance of horizontality (Liben & Golbeck, 1984; MacKay *et al.*,1972), whereas within each measure the stable context elicited always better performance than the unstable one (Liben 1975, 1978; MacKay *et al.*,1972). However, results are less consistent when performance within and across the spatial measures is analysed at each age level and gender. In general, the correlational patterns seem to be higher for boys and men,

[6] Piaget had left some ground for assuming that horizontality might precede verticality (Piaget & Inhelder, 1948/1967, Chap.13)

although results on the magnitude of this gender difference are inconclusive across different developmental stages (Liben, 1975, 1978; Liben & Golbeck, 1980, 1984).

5.2 Contextual realism of mobile reference systems

Variation in performance was also observed in studies which systematically manipulated the degree of realism of the physical and mobile reference systems. It was hypothesised that performance on static and dynamic displays of the pertinent physical phenomena may be different and that movement provides additional information that could interfere with accuracy[7]. Several studies compared performance on the standard 'paper and pencil bottle outlines' task with others that used a more ecologically realistic water-level context (real tilted bottles, pouring scenes). Results were consistent in that realistic cues hinder the performance of children and adults (Kalichman 1988; Pascual-Leone & Morra, 1991).

Howard (1978), testing adults, utilised a dynamic display paradigm in a water-level study with a free-choice perceptual task. He found that performance with the dynamic display was less accurate than with the static one and he suggested that movement might add complexity to the task. Similarly, studies with children indicate that a paper-and-pencil (2D) version of the WLT was much easier than the real-bottle (3D) version, in that it reduces bottom-parallel responses (Pascual-Leone, 1969, 1989). It has also been claimed that pictures of tilted bottles presented unsupported are more likely to impede performance, being interpreted by young children as implicitly in motion. In view of these findings, it seems that, when movement is a contextually salient feature, it might activate task irrelevant schemes in subjects (especially children) who lack the knowledge of the invariant horizontality and verticality of WLs and PLs in respect.

[7] Opposing views had been also reported (Howard 1978, Kaiser *et al.*,1985; Pittenger, 1987) on the basis that more realistic scenes could facilitate the recognition of the correct orientation of WL and PL. However, these results had been obtained only with adults who are expected to have some grasp of the relevant physical principles.

5.3 Mental rotation and visualization

In the preceding sub-sections, we discovered that the aspect of implicit motion in some of the Piagetian spatial tasks and/or their presentation can impede performance. The principal reason for this result seems to originate from the cognitive demands they impose on visual imagery and mental rotation in particular. The component of visual imagery in Piagetian tasks has been discussed (DeLisi *et al.*,1976; Kalichman, 1988, Vasta & Liben, 1996), and empirical support for its association with the WLT has been also obtained (Signorella & Jamison 1978). Some authors have even argued that very young children may fail to use mental imagery altogether, inaccurately interpreting the stimulus (bottle or van) as a tilted picture of the stimulus in its original position and not as a picture of a tilted stimulus (Freeman, 1980; Howard, 1978).

Further support for the involvement of mental imagery in WLTs has been provided on the ground that the magnitude of deviant responses is analogous to the extent to which the proximal frame has been tilted for the subjects who perform poorly on the Piagetian tasks (Amponsah & Krekling, 1994; Lohaus *et al.*,1996; McAfee & Proffitt, 1991; Vasta *et al.*,1994). A cognitive process, which accounts for these findings, is claimed to be mental rotation. The argument of these researchers is that, subjects, encountering tilted containers, mentally rotate them from their 'canonical' upright position. The amount of the required mental rotation is positively related to the size of error. Partial support for this assertion is provided by an earlier study by Smedslund (1963) who found that children perform better on a picture selection compared to a drawing task only with bottles rotated the least from the vertical axis.

6. EXECUTIVE GRAPHIC SKILLS

6.1 Task's response modality

Variation in performance was also reported when the response modality was systematically manipulated (Perner *et al.*,1984; Reith *et al.*,1994). Perner compared performance on horizontality and verticality with a copying task and a picture selection task[8]. He found a significant main effect of response mode, with the perceptual task eliciting higher percentage of accurate responses in children of 4 to 7 years of age at each age level, and a significant task by response mode interaction. Specifically, children demonstrate significant improvement from drawing to recognition with the MST and the chimney task (especially in the incidence of perpendicular errors) although no such difference was observed for the WLT.

An earlier study by Smedslund (1963) found that performance in a WL drawing task and a multiple-choice perceptual task was comparable using bottles with three oblique positions (30°, 60° and 150°). However, the degree of equivalence between the tasks varied across the bottle's position, with the perceptual task being easier when the bottle had a smaller rotation from the vertical axis (± 30°). Within and between age variations were also observed in tasks with different demands on graphic skills. Reith *et al.,* (1994), testing 5 to 9 year olds on a WLT with a facilitating device such as a *Da Vinci window*[9], found that between age differences were sharply reduced and bottom-parallel responses were absent even among the youngest of the children. However, the age effect was more pronounced when the testing condition required the graphic translation of model's three-dimensional properties on a two-dimensional medium. In view of all these findings, we can conclude that when the production response requirements are removed or diminished, the performance of young children showed a general improvement.

[8] In Perner's study children had to choose between two orientations, one accurate and one perpendicular to the immediate context.

6.2 The geometric factor

Many researchers have tried to explain the finding that, when a task requires a manual response (producing a graphic trace in a drawing task, adjusting a target line with a psychophysical apparatus), children produce larger deviations from the vertical and horizontal axis orienting the forms towards the perpendicular. A number of authors have argued that geometric performance factors can interfere with the ability to use a reference spatial system external to the context. Ibbotson and Bryan (1976) provided evidence suggesting that the difficulty with the Piagetian tasks lies in angle reproduction rather than in a conceptually underdeveloped reference system. Using abstract tasks, with response requirements and perceptual features similar to the standard Piagetian ones, they found that children of 5 to 6.6 years made errors when copying a non-perpendicular line on an oblique baseline; their errors were reliably biased towards the perpendicular. The authors conclude that the perpendicular error is a basic geometrical tendency evident in various construction tasks[10]. Later studies confirmed these findings (Bremner, 1984; Bremner & Taylor, 1982; Freeman, 1980).

7. CONCEPTUAL FACTORS

On the basis that meaningful spatial tasks are more difficult than abstract tasks, some authors maintain that the perpendicular error is rather conceptually driven. Pascual-Leone (1969, with Goodman, 1979) explained that the Piagetian traditional tasks require a certain mental capacity to overcome figurative schemes that are frequently encountered. Bottles in children's environment are usually observed in an upright position, where the water is at the bottom and parallel to it. The WLT presents children with a situation which violates 'ecological frequency'. As a consequence, particularly children of a younger age might find

[9] Children could reproduce the projective shape of the model with contour lines on a plexiglass pane.
[10] Similar results were obtained when children had to position a wire on an oblique baseline.

it hard to mentally delete this canonical presentation because they are most susceptible in preserving the characteristic features of objects.

Pascual-Leone and Morra (1991) claim that people hold two types of inaccurate believes about WLs. One is that the water is always parallel to the base and perpendicular to the sides of the vessel and another that the liquid falls towards the lower region of the tilted frame and its level is higher at the pouring edge. Younger children are expected to be susceptible to the former whereas older subjects, having the ability to acknowledge transformation, will probably demonstrate the later cognitive bias. In fact, the verbal explanations, given by subject for their picture drawings or selections, supports the latter assumption (Howard, 1978). Thus, it seems that subject's developmental level, (spatial ability, cognitive style) will determine the operative strength of the cognitive biases.

Likewise, the vertical task is not immune from the consequences of conceptual constraints. Children see trees growing vertically from the ground and chimneys in front or behind roofs. As Arnheim (1969) has stressed, the perpendicular solution in the vertical tasks has to be seen as a local drawing convention for depicting 'uprightness', without violating this aspect of canonicality. Piaget (Piaget & Inhelder, 1948/1967) cited protocols from children, giving verbal explanations for their drawings. These reports reveal the nature of their conceptual schema which preserve the forms' canonical orientation. Predominantly, in the absence of a model, these figurative schemes, which are over-learned due to their ecological frequency, will be readily activated resulting in automatized responses (Pascual-Leone and Morra, 1991). In other words, when children have to make a predictive response, their "internal model" of the scene is more likely to guide their representational attempts (intellectual realism: Luquet, 1927/1977).

8. WHICH BIAS?

The view that children's perpendicular errors derive from conceptual constraints is not necessarily irreconcilable with the assertion that such errors originate in geometric constraints since both attribute children's failure to their reliance on a local cue and not on an invariant and distant one. However, in relation to the WLT, the former view interprets this reliance as a sign of children's conceptual schema which operates via the bottom of the vessels whereas the latter as a performance deficit which operates towards the sides of the container.

One study tried to disentangle the 'bottom-parallel' conceptual bias from the perpendicular motor-response bias. Sommerville (Sommerville & Cox 1988) noticed that the majority of water-level studies used stimuli where the sides of the container were perpendicular to its base making it impossible to identify the relative strength of these biases from the children's finished products. Using both a standard mug and a slanted one - with their sides oblique to its horizontal rim and base - she found that the mean angular deviations were significantly larger with the conventional mug than with the slanted mug. The same results were obtained analysing the direction of deviations from the horizontal axis

This finding points to the operation of both a conceptual and a motor bias. However, when the testing materials allow the congruent application of both biases, performance is impaired compared to cases where these two forces are acting in opposition. It has been already mentioned that the studies which tried to eliminate or reduce the linearity of the context elicited less deviant scores and more accurate orientations. It is likely that reducing contextually misleadingness not only helps children to disembed the target line at the level of primary perception but also fosters its accurate reproduction by obstructing the operation of the perpendicular bias.

9. OVERVIEW

In contrast to Piaget's original view, many who have studied extensively and systematically his tasks in relation to the spatial axes, now contend that a number of skills and abilities underlie accurate performance. Hitherto, three researchers attempted to offer a comprehensive view, accounting for all the factors involved. Although they were mainly concerned with the WLT, the same issues are involved in the vertical tasks (particularly in the PLT). Kalichman (1988) offered a six-component model, involving the skills of visual perception, disembedding, image generation, mental rotation, spatial co-ordination and information recall. Liben (1991a) affirmed that the tasks require competencies involving spatial systems, knowledge of the physical principles, perceptual skills and graphic skills. Finally Pascual-Leone and Morra (1991) presented a model developed from the premise that these tasks cause a cognitive conflict in which misleading (irrelevant) factors compete with correct (relevant) ones. These factors involve cognitive schemes, executive schemes, and field effects while the relative strength of these is determined by the child's mental attentional capacity.

When the testing procedures require the simultaneous use of most of these skills, the requisite spatial system is more likely to become inaccessible even if it is acquired. For instance, mnemonic tasks, requiring most of the aforementioned skills (visual imagery, rotation, executive and conceptual schemes) were found to be substantially more difficult than non mnemonic ones[11] (production tasks: adjusting a metal disc in Thomas & Jamison, 1975[12], drawing a line in Sommerville & Cox, 1988 and in Reith *et al.*, 1994, and perceptual task: Ford, 1970). These findings, along with many of those reviewed in the present chapter, indicate that a competence-deficit interpretation of young children's performance is not

[11] In mnemonic tasks children were tested without the target line present in the model, while in the non-mnemonic tasks, it was present.

always accurate; particularly for those at transitional stages. In the next chapters, we will examine the resources children of different age levels have at their disposal and the conditions under which they can effectively utilise them to solve horizontal and vertical tasks.

[12] Thomas used a psychophysical procedure where subjects had to adjust a disk's position inside a transparent half-filled cylinder, predicting the position that the level of real water had taken (in the first study) or matching that of the models' (in the second study).

CHAPTER IX

HORIZONTALITY ASSESSMENT

1. METHOD

1.1 Overview

Two WLTs were administered to assess the concept of horizontality. These tasks were similar to the traditional Piagetian one in that children had to draw the water level in containers with an oblique orientation. In the present study however, the water level was present in the 2D model and the children had to copy, rather than predict, the water level from the container's orientation. This variation was introduced to reduce the task's complexity and to eliminate the confounding effect of mental imagery involved in the performance on the original WLT (Kalichman, 1988; Signorella & Jamison, 1978; Vasta & Liben, 1996). A single orientation of the tilted container was examined and kept constant between the WLTs since previous studies found no difference in the performance of children across the various oblique positions of the bottles (Abravanel & Gingold, 1977; DeLisi *et al.*, 1995; Liben, 1974 & 1978; Thomas & Jamison, 1975). The shape of the containers was varied to observe the effect that the degree of conflict between the proximal frame of reference (container) and the true environmental co-ordinates would have on water level representation.

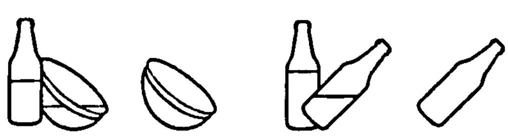
Performance was also examined in relation to the cognitive style variable of FD/FI (Witkin *et al.*, 1971). As noted elsewhere (Liben, 1978) the Piagetian WLT may be thought of as a disembedding task in which children have to ignore proximal cues (container's frame) and attend to more distal cues (edges of drawing paper, table) to succeed in the task.

1.2 Participants

The sample consisted of the same children who participated in occlusion task. For descriptions of their characteristics refer to Chapter V, section 2.1.

1.3 Materials

Horizontality Assessment: Children were presented with two WLTs; the Bowl and Bottle task. In the Bowl task, children were shown a picture outline of an array of two containers placed one next to the other in a straight line relative to the subject's viewpoint (Figure 9.1, a). The container to the left was a rectangular bottle, 3.00cm tall (2.00cm up to the neck) and 1.00cm wide at the bottom, in a vertical orientation. The container to the right was a round bottom bowl, 2.50 cm in diameter and 1.50cm tall (1.00 outer surface and .50cm inner surface), tilted 45° degree clockwise from the vertical. Neither its outline nor its rim contained any straight lines that would compete with the spatial co-ordinates of the environment. In the second task, the stimulus card depicted an array of two identical bottles, with the same dimensions of the upright bottle in the previous task. Likewise, the bottle to the left had a vertical orientation while the one to the right was tilted 45° from the vertical axis in a clockwise direction (Figure 9.1, b).



a_1 a_2 b_1 b_2

Figure 9.1 Schematic illustration of stimulus material.

The vessels' configurations were placed equidistant from the edges of the paper in order to keep the relative alignment of each container equally salient to the contextual frame. The water was drawn as a straight line across the vessels' walls. Two thirds of the vertical bottles and one third of the tilted containers were filled with water to make its level distinct in each vessel within the array. Neither the stimulus cards nor the children's paper had any extra proximal alignment cues that parallel a distal horizontal reference (horizontal axis of the paper)[1].

Cognitive style assessment: Field independence (FI)/field dependence (FD) was assessed on primary school children with the Children's Embedded Figure Test (CEFT: Witkin *et al.*, 1971) and on nursery children with the Preschool Embedded Figure Test (PEFT: Coates, 1972). For a description of the two versions of the Embedded Figure Test (EFT) refer to Chapter III, section 2.3.3.

1.4 Procedure

Children were tested individually, first with the CEFT or the PEFT, following standardised instructions (Coates, 1972; Witkin *et al.*, 1971 respectfully). Approximately after two weeks, the WLTs were administered in a single session. This order and the time between the two testing sessions were the same for each child. Since previous research seems to indicate that non-rectangular vessels facilitated performance, the experimental tasks were presented in an increasing order of difficulty, with the tilted bowl first followed by the tilted bottle. For each task, children were provided only with the outline of the tilted container drawn on their paper ("tilted vessel presentation" trial). After matching the pre-drawn tilted form with the equivalent one on the stimulus card, children were requested to look carefully to the model and draw the water line in the tilted container exactly as it looks in the model[2]. Children were not allowed to rotate their paper while drawing.

[1] In some previous WLTs (DeLisi, 1983; DeLisi *et al.*, 1995; Liben, 1975; Sommerville & Cox, exp.1, 1988) a horizontal reference line was drawn under the vessels to represent the table top. All these tasks were predictive in nature and this reference line served as a facilitator in making the appropriate axis directly perceptible and readily accessible in orienting the liquid surface. Since in the present tasks children had to copy the water level, the provision of an additional horizontal reference line was considered redundant.

[2] No explicit reference to the horizontality of the target line was made in the verbal instructions.

This procedure was followed with all children except with 4 year olds who were administered an additional task prior to each experiment proper. Those children, after the presentation of each picture, were provided with papers where the outlines of both vessels (vertical bottle and tilted container) were pre-drawn, and were requested to copy the water level in the tilted container (two-vessel presentation trial). In the Bowl task only, children were also asked to copy first the water line in the vertical bottle before that in the tilted bowl. This additional request served the need to ensure that the younger children could draw the water level as horizontal in an upright bottle within a 'different-shape' array of vessels where no conflicting proximal axes were present (tilted bottle). The 'two-vessel presentation' trials were conducted for three reasons; first to familiarise the younger children with the task requirements in general, second to provide children with additional spatial alignment cues, and third to direct their attention to the invariant horizontality of water level from an immediate comparison of vessels' orientation within each array.

1.5 Scoring Method

Previous investigators have used several methods to score and analyse performance on horizontality tasks. The various methods differ in three ways: (a) analysis of actual deviation scores (DeLisi, 1983; Perner *et al.*, 1984; Sommerville & Cox, 1987; Thomas & Jamison, 1975), (b) analysis of performance when scores are dichotomised as right and wrong using a fixed criterion of angular deviation from the horizontal axis[3] (Liben & Golbeck, 1980, 1984; MacKay *et al.*, 1972; Randall, 1980; Smedslund, 1963; Signorella & Jamison, 1978) and (c) classification of scores into categories of deviations that reflect different level of understanding and use of drawing rules[4] (DeLisi *et al.*, 1995; Liben, 1974, 1975; Pascual-Leone & Morra, 1991; Reith *et al.*, 1994).

[3] Usually the cut-off point is 5° for adolescents and adults, and 10° for children to whom a less accurate motor co-ordination is presumed.

[4] Angular deviations are frequently classified into four categories: (i) *accurate* responses, within 5°-10° from horizontal, (ii) *bottom-driven* responses, essentially parallel to the bottom of the vessel, (iii)

Each of the various scoring methods has its limitations. The use of actual deviation scores and measures of central tendency provide few and distorted views about the principles that govern performance. For example, a deviation of 30° is treated as twice as wrong as one of 15°, although it might be governed by the same principle. Also a correct/incorrect system of analysis conveys the wrong impression that within each category all responses are equivalent. Finally, frequency distribution tables capture the pattern of rules that govern performance but they might be statistically inadequate when certain classes of deviations have rare occurrences and variables with many levels are considered.

Due to the fact that any one method of scoring is a trade-off between theoretical and statistical adequacy, children's performance on horizontality assessment was scored and analysed in all three approaches to obtain a more comprehensive picture of their abilities. A best-fit straight line for the level of the top of the liquid was constructed and the angle between this line and the horizontal was measured in degrees using a hand-held protractor. Angular deviations were then measured as a discrete categorical variable – and treated with count frequency analyses–and as a continuous variable– and processed with conventional parametric statistical analyses. With the former approach, water level drawings were classified as correct and wrong setting the fixed critical value at 11°. As a consequence, if the water line was within the 0°-11° interval, it was defined as correct, otherwise it was wrong. To identify drawing strategies that children used, inaccurate responses were further classified into three angular categories. The quantitative definition of angular categories is as follows: (a) *compromise* category, with scores between 12° and 33°, comprises intermediate errors between the horizontal and the inclination of the container, (b) *context-driven* category, with scores between 34° to 50°, contains water lines congruent with the 45° tilt of the container,

---

excessive responses, beyond category (ii) and towards a vertical orientation, and (iv) *compromise* responses, intermediate to (i) and (iii). The angular range of each category varies across studies depending on the degree of inclination of the stimulus vessel and the task demands. Usually there is a and (c) *excessive* category, with scores more than 50°, accommodates water lines oriented a good deal more towards the vertical than the container is tilted.

Angular deviations as a continuous variable were assessed also on *(a) absolute deviation scores*, obtained by measuring deviations off the horizontal in either direction, and on *(b) directional deviation scores*, classified as positive and negative. The following Figure presents a schematic example of the directional criteria.

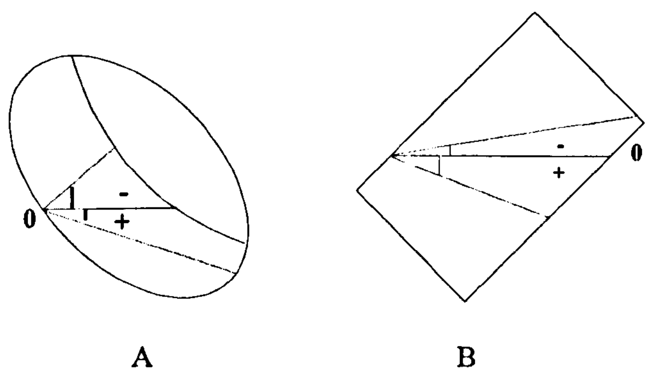


Figure 9.2 Directional errors scoring, related to the horizontal for Bowl (A) and Bottle (B).

As Figure 9.2 illustrates, a deviant score was considered as having a positive slope when the water level was drawn away from and below the horizontal axis, in a clockwise direction congruent with the tilt of the vessel (positive bias). On the contrary, it had a negative slope if the water line had a counter-clockwise inclination, in the opposite direction to that of the tilt of the vessel (negative bias). The 0° point of the protractor was always aligned with the edge of the water line on the non-pouring side of the container to obtain consistency in the scoring. These directional terms and definitions are consistent with those used by previous authors (Amponsah & Krekling, 1994; Howard, 1978; Lohaus *at al.*, 1996; McAfee & Proffitt, 1991; Sommerville & Cox, 1988). In the present study the terms *positive* and *negative* will be used descriptively without ascribing a sign on directional errors. Finally for the scoring of the two versions of EFT see Chapter III, section 2.3.3.

---

disparity of ± 10° for the bottom-driven and compromise category while there is more consistency for

2. RESULTS

At first, the results from the Bowl and Bottle tasks will be presented separately. For each WLT, the data of the nursery group will be analysed first from the initial 'two-vessel presentation' trial and compared with performance on the 'tilted-vessel presentation' trial. Analysis of actual deviation scores will be performed first, followed by directional scores and categorical data. In the last section, the results from the two WLTs will be considered together and correlational patterns between horizontality estimates and cognitive style will be examined.

2.1 The Bowl task

2.1.1 Nursery group: 'Two-vessel presentation' trial

Vertical bottle

The scoring of the water level in the vertical bottle showed that, from a total of 30 children tested, three produced unrepresentative drawings and four drew the water level as scribbles, leaving 23 drawings for further analysis. More than half of the children who represented the water as planar, could draw its level horizontal [14 (61%) vs. 9 (39%)] and the measures of central tendency and variability were low (M: 9.04, SD: 6.96) since all of the inaccurate drawings fell within the least deviant angular category (12° -33°).

Tilted bowl

The water level drawings of the tilted bowl, in both 'two-vessel' and 'tilted-vessel' trials, yielded the following results: 5 children produced unrepresentative drawings in both trials (4 scribbles, 1 equivocal) while 5 more children produced unrepresentative drawings in one or other of them, leaving 20 drawings where the water level had been drawn as a line in both of these trials. With the exception of three cases, children failed in both trials (85%) and did not benefit from the facilitating 'two-vessel' presentation. The consistency in children's

---

the quantitative definition of accurate and excessive responses.

performance across the trials was further confirmed from the absence of a significant difference in the mean absolute deviation scores (t test for paired samples t(19) = .59, p>.05).

2.1.2 Total sample: 'tilted-bowl presentation' trial

From a sample of 110 drawings, all children, except one from the nursery, produced recognisable representations of water in the tilted bowl. Seven 4-year old children failed to draw the planar surface of the water and instead they drew the water as scribbles in six drawings, and as an open circle in vessel's mouth in one drawing. These seven drawings were excluded from the analysis of results, leaving a total of 102 drawings where the water level was drawn with a line. Results from horizontality assessment will be first reported on deviation scores. A 4 (age) × 2 (gender) unweighted analysis of variance was first performed on absolute mean deviation scores. Means and standard deviations for this test are reported in Table 9.1 where higher scores reflect less accurate performance than lower scores.

Table 9.1 Means and standard deviations of absolute angular deviations off the horizontal by age level and gender for Bowl task.

			Boys		Girls		Total	
Age groups	N[a]		M	SD	M	SD	M	SD
4	22	B9;G13	33.22	19.87	41.62	20.36	37.42	20.11
5 ½	25	B10;G15	33.20	20.33	30.33	10.06	31.76	15.19
7	28	B8;G20	11.88	11.95	25.55	17.30	18.71	14.62
8	27	B10;G17	10.60	9.92	18.18	16.61	14.39	13.26
Total	102	B37;G65	22.22	15.52	28.92	16.08	25.57	15.80

a The number of children in each age group with number of boys (B) and girls (G) in each cell.

The main effect of age was significant with scores becoming less deviant as a function of children's age level ($F_{3,94}$ = 10.07, p<.001). Newman-Keuls post-hoc analysis for difference (p<.05) revealed that each of the two younger age groups produced significantly higher absolute mean deviation scores from each of the two older ones. Gender just failed to have a significant main effect ($F_{1,94}$ = 3.88, p = .052) and did not interact with age ($F_{3,94}$ = 1.05, p = .18). However, the trend is consistent with previous results, since, in the whole sample, boys were more accurate than girls, and this pattern was present in all age groups except with 5½ year olds where girls were slightly better than boys.

Angular deviations were classified further as positively and negatively biased depending on their direction from the horizontal axis. Perfect scores (0°), which were not biased to any direction, were excluded from this analysis. Results from the total number of 99 directional scores revealed that the majority of the water lines had a negative slope falling above the horizontal axis in an opposite direction to the tilt of the bowl (Negative: 80.8%, Positive: 19.2% see Table 9.2). Errors with a positive slope had also a significantly higher mean and a somewhat larger standard deviation compared to those with a negative slope (M = 34.80 and 24.81 respectively, t(97) = 2.18, $p<.05$[5]). Two separate 4 (age) × 2 (gender) unweighted analyses of variance were performed for each directional category since the distribution of positive biased errors were distinctly different from the distribution of negative biased ones. In both analyses, gender did not have a significant effect and did not interact with age[6]. On the basis of the consistent absence of any significant gender effect, results on directional scores are reported only by age level in Table 9.2.

Table 9.2 Means and standard deviations of directional scores by age within directional category.

		Scores with a negative slope			Scores with a positive slope		
Age groups	N	N	M	SD	N	M	SD
4	22	13	33.92	19.87	9	44.33	19.98
5 ½	25	21	32.19	15.17	4	27.75	13.17
7	27	25	22.04	16.30	2	27.50	29.00
8	25	21	15.19	10.46	4	24.00	30.01
Total	99	80	24.84	16.79	19	34.79	22.17

The main effect of age was significant only for scores with a negative slope (Negatively biased scores; $F_{3,72} = 6.75, p<.001$, Positively biased scores; $F_{3,12} = 0.73, p = 0.55$). Follow-up

[5] Levene's test for equality of variance: $F = .54, p = .47$.
[6] Negative scores; Gender: $F_{1,72} = 1.64, p = 0.21$, Gender × Age: $F_{3,72} = 0.63, p = 0.60$. Positive scores; Gender: $F_{1,12} = 0.55, p = 0.47$, Gender × Age: $F_{3,12} = 0.72, p = 0.51$.

study revealed that the 5½ year olds produced significant larger deviations in the opposite direction to the tilt of the bowl than any of their older counterparts and significant also was the difference in means of 4 and 8 year old children (post-hoc Student Newman-Keuls tests, $p<.05$). Considering the mean errors with a positive slope, 4 year old children appear to orient the water level a good deal away from the horizontal axis in comparison to all the other age groups. However, the absence of an age effect for positively biased scores can be attributed to the small number of observations in each cell and the difference in the within-age group variability. This variation in the magnitude of standard deviations between age groups suggests that positively biased children of an older age oscillate between very different response strategies whereas younger children are more consistent in their drawing solutions. Further, within age level, there was no difference in the mean scores across directional preference.

Similar results were obtained from the analysis of categorical data. Table 9.3 presents the frequency distribution of scores across angular categories by age. Collapsing all the non-horizontal categories, the distribution of correct (0°-11°) and wrong (≥12°) line drawings across age levels revealed that the number of drawings falling within the horizontal sector increased with age [$\chi^2(3) = 15.72$, $p<.01$]. Inspection of Table 9.3 indicates that there is an abrupt shift towards accuracy at the age of 7 (5½ with 7 year olds $\chi^2(1) = 5.10$, $p<.05$) with comparable performances within the two younger and the two older age groups. Further, the analysis of absolute mean deviation scores showed a moderate difference in the performance of the sexes ($F_{1,94} = 3.88$, $p = .052$) in favour of the boys. A similar analysis on the number of correct and incorrect water level drawings revealed a reliable difference in the performance of boys and girls. A chi square test showed that significantly more boys passed the WLT than girls [$\chi^2(1) = 5.73$, $p<.05$].

Breaking the non-horizontal responses into three angular categories, we observe comparable frequencies of accurate, compromise and context-driven drawings in the total sample (31%, 33% and 27% respectively). However, different age groups exhibit different response patterns since the distribution of drawings across all four angular categories by age was statistically different [$\chi^2(9) = 25.86$, $p<.01$]. As Table 9.3 shows, children evolve from a unimodal distribution at 4 years of age, with mode located between 34°-50° interval, to a bimodal one, located at the age of 5½ years within the middle two angular response categories and at the age of 7 within the first two angular categories.

Table 9.3 Distribution of angular deviations into drawing categories in the Bowl task by age.

		Drawing categories of angular deviations				
		Correct	Incorrect			
Age groups	N	Accurate 0°- 11°	Compromise 12°- 33°	Context-driven 34°-50°	Excessive >50°	Combined ≥12°
4	22	3	5	9	5	19
$5^{1/2}$	25	3	10	10	2	22
7	28	11	10	6	1	17
8	27	15	9	2	1	12
Total	102	32	34	27	9	70

Finally, by the age of 8, the frequency distribution of drawing responses gains unimodality with the mode clearly located within the accurate response category. This pattern of responses clearly indicates that there is a gradual age-related shift in the obtained frequencies towards the left hand-side of the Table 9.3 (accurate category). Excessive drawings towards the vertical are rare and mostly found among the 4 year olds, while context-driven responses are most frequent before the 7 years of age. By the age of 7 children's drawings are pooled towards the horizontal and demonstrate a higher degree of accuracy by the 8 years of age.

Pascual-Leone & Morra (1991) presents published and unpublished results from studies which report frequency distributions, categorising absolute deviations into angular classes. A problematic aspect of this approach is that not only the authors are ignoring the directional variable but they are also using the term 'bottom-driven' to define the 33° -50° angular class. This label conveys the false impression that all the drawings in this category have a positive slope. Since the water level drawings in the present task demonstrate a prevalent negative directional bias, the previous analysis was repeated separately for each directional class.

The distribution of negative biased scores across drawing categories was identical to those obtained on absolute scores both in terms of the age-related pattern and of the nature of modality within each age level. The only slight variation was the pattern of responses of 4 year olds whose drawings were evenly distributed across the four angular categories. A similar analysis was not feasible for scores with a positive slope due to their infrequent occurrence. However, collapsing the results from the two younger and the two older age groups, the modal response of the younger children was located within the bottom-driven category (7 out of 13 in 34°-50°) while that of the older group within the accurate category (4 out of 6 in 0°-11°). An examination of the total number of inaccurate drawings for each directional class indicates that the errors were not equally distributed across the three levels of angular deviations. The modal response for negatively biased errors was compromise (31 out of 56, 55%), and for positively biased errors was bottom-driven (8 out of 14, 57%). Although the results for the positively biased errors should be taken with caution due to their infrequent occurrence, categorical data also suggests that different drawing strategies underlay the two directional patterns.

Finally, children of all age levels demonstrate a clear negative bias which became more dominant with age [$\chi^2(3) = 9.44, p<.05$]. This directional preference was also independent of accuracy. Using the correct/incorrect classification, it was found that 24 out of 29 (83%) accurate drawings, and 56 out of 70 (80%) inaccurate ones had a negative slope. As a consequence, children who passed the Bowl task were as likely to be negatively biased as children who failed it. Dichotomising the drawings into pass and fail, a gender effect was found for scores with a negative slope. Boys demonstrating a negative bias were as likely to pass as they were to fail while negatively biased girls mainly failed the task [$\chi^2(1) = 4.06$, $p<.05$][7].

2.2 The Bottle task

2.2.1 Nursery group: 'Two-vessel presentation' trial

From a total of 30 nursery children tested, 3 represented the water level with scribbles in the tilted bottle, in both the 'two-vessel' and the 'tilted vessel' presentation trials, and 1 child in one of them. The analysis on the remaining 26 drawings showed that performance was remarkably consistent between the trials. Classifying the drawings into correct and incorrect, there was no evidence of any change in response since twenty-four (93%) children failed in both trials. The same pattern of results was obtained analysing the mean absolute deviation scores and the distribution of drawings across angular categories between the trials. There was no difference between the means [$t(25) = .81$, $p>.05$] nor was there a change in the drawing patterns between the tasks (Wilcoxon test, $z = .29$, $p>.05$). The results show that nursery children neither could use the additional horizontal lines of the vertical bottle to orient the water level in the tilted one nor could recognise the principle of invariance from a direct comparison in the 'two-vessel presentation'.

2.2.2 Total sample: 'tilted-bottle presentation' trial

The mean absolute deviations for each age level by gender are reported on Table 9.4. A 4 (age) × 2 (gender) unweighted analysis of variance showed that the main effect of age was significant ($F_{3,99} = 7.62$, $p<.001$) with the scores of 4 year olds significantly higher than any

[7] An association of gender with passing and failing the Bowl task had been also reported in the analysis of absolute scores.

of the other age groups (Newman-Keuls post-hoc analysis for difference, $p<.05$). Gender neither had a significant effect nor interacted with age. Boys and girls perform at similar levels in the total sample ($F_{1,99} = 0.73$, $p = 0.39$) and within each age group ($F_{3,99} = 1.64$, $p = 0.18$) with the only exception the 8 year olds where boys' performance surpassed girls'. However, the gender difference at this age is mainly attributed to a few discrepant scores in the girls' sample as reflected in the magnitude of difference in the standard deviation of boys' scores compared to that of girls'. Excluding these scores, the general results remained the same but the within group difference was reduced (8 year old girls, M: 20.33, SD: 19.65, sample's mean: 15.26).

Table 9.4 Means and standard deviations of absolute angular deviations off the horizontal by age level and gender for Bottle task.

Age groups		N[1]	Boys		Girls		Total	
			M	SD	M	SD	M	SD
4	27	B11;G16	40.45	14.61	41.44	12.47	40.95	13.54
5 ½	25	B10;G15	29.50	22.52	24.67	18.73	27.08	20.62
7	28	B8;G20	17.13	26.00	16.40	17.16	16.76	21.58
8	27	B10;G17	10.20	8.34	28.41	29.29	19.30	18.81
Total	107	B39;G68	24.32	17.87	27.73	19.41	26.02	18.64

[1] *The number of children in each age level with number of boys (B) and girls (G) in each cell.*

Absolute errors were further divided into directional ones. Five drawings which were not biased to any direction (0°) were excluded from the analysis. From the total sample of 102 directional scores, 88 (86%) drawings had a positive slope and 14 (14%) had a negative one. Comparing performance between the two directional categories, it appears that there was a significant difference in the means of positive and negative errors [t(100) = 4.19, $p<.001$][8]. As Table 9.5 shows, water lines tilted in the opposite direction to that of the bottles' (negative) were closely plotted near the horizontal axis (M: 4.14); the distribution of water lines with a slope congruent to that of the bottle's (positive) were more dispersed, with a mean well beyond the horizontal co-ordinate (M: 31.42). Two 4 (age) × 2 (gender)

[8] The Levene's test for equality of variance was significant $F = 28.74$, $p<.001$

unweighted analyses of variance were performed for each directional mode. In both analyses, gender did not have a significant effect and did not interact with age[9]. Due to the consistent absence of any significant gender effect, results on directional scores are reported only by age level in Table 9.5.

Table 9.5 Means and standard deviations of directional scores by age within directional category.

		Scores with a negative slope			Scores with a positive slope		
Age groups	N	N	M	SD	N	M	SD
4	27	1	15.00	-	26	42.04	12.28
5 ½	24	2	1.00	.00	22	30.14	18.68
7	26	5	2.20	1.30	21	21.62	20.25
8	25	6	5.00	4.65	19	29.21	26.51
Total	102	14	4.14	4.60	88	31.42	20.63

The main effect of age was significant for positive errors ($F_{3,84} = 4.45$, $p < .01$), with the line drawings of 4 year olds being significantly more deviant than those of 5½ and 7 year old children (Newman-Keuls post-hoc analysis for difference, $p<.05$). The moderate increase of the mean and standard deviation of positive errors at the age of 8 is due to the two extreme scores in the sample. Excluding again these two scores, the general pattern of the results remains the same[10] but a clearer picture of the developmental pattern was obtained with mean positive scores smoothly decreasing with age (8 years old: M = 22.17, SD = 17.06). The main effect of age for deviations with a negative slope was also significant, $F_{3,10} = 4.67$, $p < .05$; however due to a small number of entries in each cell, it should be interpreted with caution. From Table 9.5 it is clear that all, regardless of age, children demonstrating a negative bias, oriented the water level close to the horizontal axis with a single exception of one drawing of a 4 year old.

[9] Positive scores; Gender: $F_{1,80} = 2.82$, $p = .09$, Gender × Age: $F_{3,80} = 1.99$, $p = .12$, Negative scores; Gender: $F_{1,8} = 0.00$, $p = .95$, Gender × Age: $F_{3,8} = 0.12$, $p = .74$.

Performance on the Bottle task was further examined with the number of accurate (0° -11°) and inaccurate drawings (≥12°) along the different angular levels of divergence (Table 9.6). As expected, there was an age-related increase in the number of children who draw correct versus incorrect water levels [$\chi^2(3) = 24.0, p<.001$]. The shift towards accuracy took place at the age of 7 where nearly half of children produced accurate drawings, a rather infrequent occurrence before this age. Gender was not associated with the probability of passing or failing the Bottle task since 67% of boys and 70% girls failed the task in the overall sample.

Table 9.6 Frequency distribution of angular deviations into drawing categories in the Bottle task by age.

		Response categories of angular deviations				
		Correct	Incorrect			
Age groups	Total	Accurate 0°- 11°	Compromise 12°- 33°	Context-driven 34°-50°	Excessive >50°	Combined ≥12°
4	27	0	6	15	6	27
$5^{1/2}$	25	5	11	6	3	20
7	28	15	8	2	3	13
8	27	13	8	3	3	14
Total	107	33	33	26	15	74

Analysing the results across the four quantitative classes of deviations by age, the distribution of drawings was significantly different between children of different ages [$\chi^2(9) = 36.64, p<.001$]. The 4 year olds made the most of the excessive errors and the majority of their drawings were context driven without being able to produce any water level drawing with correct orientation. Five-and-a-half year olds reduced the number of context-driven responses while making mostly compromise errors and starting to draw few accurate water levels. The difference in the response patterns between the two younger groups was significant [$\chi^2(3) = 11.27, p<0.01$]. Considering the two older age groups, their modal response was clearly within the accurate angular category although they kept on making compromise drawings.

[10] In this analysis, post-hoc comparisons revealed that 4 year olds' positive errors were significantly more discrepant than any of the older counterparts.

An examination of the distribution of responses with respect to the directional preference reveals the following results. First, dichotomising the drawings into correct and incorrect, an association was found between passing the Bottle task and having a directional preference. Children having a positive bias were more likely to make an incorrect drawing (82%) compare to those with a negative bias (14%) [$\chi^2(1) = 24.37, p<.001$]. Second, considering the four angular categories, more than half of the positive biased scores were equally distributed between the compromise and bottom-driven category (35% and 30% respectively) while the remaining drawings were evenly divided between the accurate and excessive category. The distribution of positive scores within each age level revealed a similar pattern as that of absolute ones. The modal response of 4 year olds was a bottom-parallel one (58%) while that of 5½ year olds a compromise one (50%). The two older age groups seem to oscillate between accurate and compromise responses. Negative biased scores were mainly accurate (12 out of 14) with the two incorrect falling into the compromise angular category.

2.3 The EFT and WLTs

A 4 (age) × 2 (gender) unweighted analysis of variance on EFT scores[11] yielded only a significant main effect of age ($F_{3,94} = 19.31, p<.001$), although the main effect of gender was in the predicted direction (Boys: M = 14.35, Girls: M = 12.80, $F_{1,94} = 3.32$, $p = .07$). The results on age and gender differences are consistent with those reported in the manual (Witkin *et al.*, 1971) and in subsequent research where it is suggested that children are becoming more field independent with age[12] and that boys are performing better than girls. Since the main effect of age was significant, the median score of each age group was used to further classify children into two categories of cognitive style status. Children whose EFT

---

[11] Only EFT scores of children who drew the water as a line in both horizontality tasks were considered in this analysis (N = 102)

[12] The same analysis was repeated excluding the nursery group whose results were obtained on the pre-school version of EFT (PEFT) and whose mean was significantly different <u>only</u> from that of the older age group. The pattern of results remained generally the same. Age preserved its significant effect ($F_{2,74}$

scores were smaller or equal to the median of the relevant age group were classified as FD, and these children whose scores were above the median were classified as FI. Two separate tests of association between cognitive style status and age, and gender demonstrated that there was no difference in the number of FD/FI children across the age groups [$\chi^2(3) = 0.44$] and between boys and girls [$\chi^2(1) = 0.92$].

The results comparing performance on EFT with horizontality tasks will be first presented using the actual EFT scores and later the cognitive style status classification. The intercorrelation matrix for the horizontality tasks and the EFT scores are reported in Table 9.7 with children grouped by age and gender. Intertask (Bowl-Bottle) correlations were high and significant for the total sample and for both sexes. The correlational pattern across the age levels showed that, by the age of 7, children performed comparably in the Bowl and Bottle task whereas the performances of younger children were more variable across the tasks.

Table 9.7 Intercorrelation estimates of horizontality and EFT measures by age and gender of children.

	Subject Groups						
	Age Groups				Gender		
Measures	4	5 ½	7	8	Boys	Girls	Total
Bowl Bottle	.20 (21)	.15 (25)	.63[c] (28)	.46[a] (27)	.47[b] (37)	.45[c] (64)	.46[c] (101)
Bowl EFT	-.15 (21)	.08 (25)	-.06 (28)	-.43[a] (27)	-.35[a] (37)	-.20 (65)	-.27[b] (102)
Bottle EFT	-.01 (27)	-.49[a] (25)	-.45[a] (28)	-.46[a] (27)	-.60[c] (39)	-.16 (68)	-.34[c] (107)

Note 1 *The numbers in parenthesis show the number of children in each cell.*
Note 2 *The negative correlations between WLTs and EFT indicate that higher scores (larger deviation scores) on WLTs represent worse performance, whereas higher scores on EFT (number of items passed) represent better performance.*
Note 3 *a = p<.05, b = p<.01, c = p<.001.*

= 26.12, $p<.001$) and gender effect reached significance ($F_{1,74}$ = 5.00, $p<.05$) while the effect of interaction remained statistically non significant.

Intertask correlation were assessed by cognitive style status and significant results were obtained for both FD [r(55) = .50, p<.001] and FI children [r(46) = .36, p<.05]. The pattern of relationship between horizontality tasks was assessed when the cognitive style groups were further divided by age and gender. Within the FD group, performance on the Bowl and the Bottle tasks correlated for the 7 and 8 year olds [r(14) = .75, p<.01, r(15) = .59, p<.05 respectively] while for the FI group, the intertask correlations were significant for the 4 and the 7 year olds [r(9) = .70, and r(14) = .60 respectively, both p<.05]. Further, there were gender differences in intertask correlations within cognitive style groups. Significant correlations between the tasks were obtained within the FD group only for girls [r(37) = .59, p<.001] and within the FI group only for only boys [r(19) = .57, p<.05]. Differences in correlations for boys versus girls within cognitive style were not significant, but differences between cognitive style groups within gender were significant only for girls (FD .68 vs. FI .22, z = 1.78, p<.05 for one tailed comparison).

In addition, the correlations between performance on each horizontality task and performance on EFT were significant and of similar magnitude for the entire sample (-.27 vs. -.34, z = .025, p>.05), suggesting that performance on the two horizontality assessments was related to field independence. However, the magnitude and the pattern of correlations for Bowl-EFT and Bottle-EFT varied with age and gender. In general, the former correlations were uniformly smaller than were the latter ones as expected since the conflicting frame of the bottle required more disembedding abilities.

With respect to cognitive style, less deviant scores on the Bowl task were associated with a tendency towards field independence only for 8 year old children whereas, with respect to the Bottle task, this relationship was significant from the age of 5½. A test of differences between the correlated correlation coefficients (Steiger, 1980) of Bowl-EFT and Bottle-EFT within the age groups revealed that the difference was significant for the 5½ year olds, and the 7 year olds [t(22) = 2.34 and t(25) = 2.51, $p<.05$ for a two tailed comparison respectively]. In addition the two horizontality measures were significantly correlated with EFT for boys but not for girls, and the difference in the magnitude of these correlations between the sexes was significant for the Bottle-EFT (-.60 vs. -.16, z = 2.51, $p<.05$) but not for the Bowl-EFT (-.35 vs. -.20, z = 0.66, $p>.05$).

The results from Table 9.7 suggest that the Bottle and the Bowl task each assess horizontality representation and that the cognitive style of FD/FI is a factor which contributed to horizontality performance, especially when measured with the Bottle task. Performance of FD/FI children on the horizontality tasks was further examined with an unweighted 4 (age) × 2 (cognitive style) × 2 (horizontality task) ANOVA with a repeated measure on the last factor. Means for this analysis are reported in Table 9.8.

Table 9.8 Means and standard deviations of absolute angular deviations for the WLTs by cognitive style status and age.

	Field Dependent		Field Independent		Total	
Age	M	SD	M	SD	M	SD
Bowl Task						
4	39.42	23.16	33.11	13.27	36.26	18.21
5 ½	32.00	16.23	30.82	13.24	31.41	14.73
7	20.57	17.40	22.71	17.10	21.64	17.25
8	20.13	17.17	9.42	8.35	14.77	12.76
Total	28.03	18.49	24.01	12.99	26.02	15.74
Bottle Task						
4	41.75	15.58	43.33	7.10	42.54	11.34
5 ½	33.92	21.09	17.27	14.58	25.59	17.83
7	21.93	23.00	11.29	14.20	16.61	18.60
8	27.13	25.05	14.83	24.58	20.98	24.81
Total	31.18	21.18	21.68	15.11	26.43	18.14

Note: The total number of children is 101 with 21, 25, 28 and 27 children in each consecutive age level. The number of children classified as FD/FI at the respective grades were 12-9, 14-11, 14-14, 15-12.

The results revealed significant main effects for age ($F_{3,93}$ = 10.18, p<.001) and cognitive style ($F_{1,93}$ = 5.00, p<.05), suggesting that horizontality performance improves with age and that FI children were more accurate than FD children. However the absence of a main task effect indicates that children perform comparably in the two versions of WLTs. The ANOVA also revealed a significant interaction between age and task ($F_{3,93}$ = 2.73, p<.05) showing that the younger and the older age groups produce more deviant angular scores in the Bottle rather than in the Bowl task whereas the reverse was true for the two intermediate age groups. No significant interaction was obtained between cognitive style and task ($F_{1,93}$ = 1.80, p>.05) but the difference in performance between the cognitive style groups was bigger for the Bottle task compared to the Bowl.

Specific comparison between the tasks were further undertaken on the number of correct and incorrect drawings, and the distribution of drawings across categories of angular deviations. If we consider that the angular categories reflect different drawing strategies and level of understanding of horizontality invariance, we can explore whether the same stages of performance is evident on the two WLTs. Table 9.9 presents the combined frequency distribution of WL drawings across angular categories between the two tasks.

Table 9.9 Correlated performance of children in WLTs by angular categories.

	Bowl				
Bottle	Accurate	Compromise	Context-driven	Excessive	Total
Accurate	16	13	4	0	33
Compromise	10	12	8	1	31
Context-driven	5	8	7	3	24
Excessive	1	1	9	4	14
Total	32	34	28	8	102

Entries on the main diagonal represent the number of drawings where the same strategy has been used for both tasks. Additionally, entries below the diagonal represent drawings of children whose performance was worse on the Bottle compared to the Bowl task, as bottles' scores fall within a more deviant angular category to that of bowls' scores. On the contrary, cases above the diagonal show better performance on the Bottle compared to the Bowl task as bottles' scores fall within a less deviant angular category to that of bowls' scores.

Children tend to perform at similar levels across the tasks since 39 children consistently used the same strategy (cases on diagonal), while 29 choose a less advanced strategy on the Bottle compared to the Bowl task, and 34 used a more advanced one on the former than on the latter[13]. Among the latter cases, where a shift in the drawing strategy occurred, 51 children performed at adjacent levels (cases immediately above and below the diagonal) and only 12 used very discrepant strategies between the tasks, a response which was not age-related. Wilcoxon's signed-ranks test of pairs on scores of angular category and NcNemar test of change on pass/fail scores yield no significant difference between the tasks. These results suggest that the two tasks have similar frequency distributions and that they are of comparable level of difficulty since there was absolute no difference in the number of children who pass one task and failed the other in favour of either of the two (16 children Pass Bowl/Fail Bottle vs. 17 who Pass Bottle/Fail Bowl).

The only difference between the tasks seems to lie in the direction of deviant scores from the horizontal axis. Although the majority of water lines falling off the horizontal (>0°) in the Bottle task were oriented towards the perpendicular in the same direction as the tilt of the bottle, the water lines in the bowl had an orientation opposite to that of the tilt of the container. From the total of 93 paired comparisons, 66 demonstrate a positive bias on the Bottle and a negative on the Bowl task and only 2 showed the reverse directional pattern. A

[13] As the significant interaction of age × task showed from the previous ANOVA, the tasks' difficulty was different within each age level. While the 4 and 8 year olds did comparably worse on Bottle task, 5½ and 7 year olds did worse on Bowl task and the overall difference in the performance between the task although not statistically significant it is substantial only among the 4 year olds. In the latter sample, from a total of 18 out of 22 drawings where children change their response mode between the

test of difference in the shift of directional preference between the tasks was highly significant and of the same magnitude within each age group (McNemar test, $\chi^2(1) = 58.37$, $p<.001$).

3. DISCUSSION

3.1 Age differences

The results are in general agreement with the descriptive aspect of the Piagetian developmental account. Age-related stages were identified on each task as children progress from immature to more advanced drawing strategies. Children up to 5½ years of age have the most limited conception of the horizontality of water level since, even with the provision of a pictorial example, the incidence of accurate orientations is rare. Consistent with the Piagetian theory, the finding that these children are unable to appreciate the material facts made known to them, implies that they lack the conceptual system required for assimilating relevant perceptual experience.

Further support for a competence deficit interpretation of results from the younger children was provided by the equivalent performance of nursery children in the one- and, the intuitively easier, two-vessel presentation of horizontality tasks. Children could neither profit from the provision of congruous alignment cues (vertical bottle) on their drawing papers nor could reflect upon the invariance of water from the immediate contrast of vessel's orientation. There were also a considerable number of cases, especially in the Bowl task, where nursery children produced scribbles or equivocal drawings. Of perhaps greater interest is the finding that nursery children face similar difficulties drawing the water level in the vertical bottle. To our surprise, from the total number of children tested, only 47% produced horizontal levels, while 30% drew oblique levels –naturally of a lesser magnitude than those

WLTs, in 13 drawings, children used a more deviant strategy for the Bottle compared to the Bowl whereas only in 5 they did the opposite (4 year olds, Wilcoxon's signed-ranks test: $z = 1.70$, $p = .09$).

in tilted vessels– and even 23% scribbled or made ambiguous drawings. This finding clearly demonstrates that either due to poor motor co-ordination skills or to the representational salience of the mass of the liquid in a WLT, nursery children face difficulty even when the vessel is in its canonical orientation.

A shift towards accuracy took place, in both tasks, at the age of 7 where children substantially reduced the magnitude of angular distortions and mainly oscillated between accurate and compromise drawing solutions. This pattern of response persisted among the oldest children, indicating that the acquisition of the concept of horizontality continues beyond the age of 8. This finding is in agreement with the Piagetian theory which predicts the acquisition of Euclidean concepts around the age of 9 "with a few laggards at 12"(Piaget & Inhelder, 1948/1967, p. 408). Further, the finding that task variation did not have a significant effect and that the overall performance in the Bowl task correlated significantly with performance on the Bottle task also corroborate the Piagetian premise that a general, conceptual, spatial process underlies performance on WLTs. The most compelling evidence is obtained from the comparative analysis of drawing strategies. Not only do the same drawing strategies appear in the two alternative tasks with equivalent frequency, but also the same children produce drawings which reflect comparable levels of conceptual development across the tasks. Similar results were obtained from the study of DeLisi (DeLisi *et al.*, 1995) where children of 6 to 10 years of age performed at equivalent stage levels in a traditional and spherical water-bottle task.

On some other aspects, though, Piagetian theory is not fully supported. The pattern of results is consistent with theories which challenge the belief that children's drawings reflect a binary choice between two reference systems (Pascual-Leone & Morra, 1991). Forty-two and forty-five percent of all drawings in the Bowl and Bottle task respectfully, fell within the compromise and excessive category, suggesting that children do not always choose the axis between either the contextual and varied frame of the vessel (34°-50°) or the external and fixed frame of the co-ordinate system (0°-11°). In addition, the transition from one strategy to another across the age groups is not as all-or-none as Piagetian theory would suggest. All four response patterns exist at every age, so the developmental change takes the form of altering the strength of the various response patterns and progressing to the less deviant ones.

What comes from these two issues is that, as many authors maintain (Pascual-Leone & Morra, 1991; Kalichman 1988; Liben 1991a), there is a conflict between sets of latent processes (cognitive, perceptual, production) which the child tries to reconcile. As their relevant strength changes with development, the balance between the forces is overthrown, and the continuous effort to reinstate it, is what essentially constitutes the course of the developmental sequence of the water level task performance.

The rate of inaccuracy and the incidence of context-driven responses of the two older age groups is definitely lower than expected from the Piagetian theory and previously reported by other researchers (Pascual-Leone, 1969; Pascual-Leone & Goodman, 1979, Thomas and Lohaus 1993; Thomas & Turner 1991). This inconsistency can be attributed to the relative easiness of this task compared to the traditional Piagetian one, adopted in the previous studies, since it neither required a predictive response nor the graphic devices to translate 3D properties on a 2D surface. On the other hand, this result indicates that children of this age range do not lack the concept of horizontality but have difficulty applying this knowledge under certain conditions. Thus interpreting the developmental data of the present sample, it seems fair to attribute the performance of children up to 5 ½ years of age to competence-deficit whereas that of the older children to competence utilisation.

3.2 Gender differences

As for the gender differences, boys performed better than girls in both WLTs but the magnitude of difference is smaller than previously reported and significant only on pass/fail scores in the Bowl task. Although gender differences in WLT performance have been widely reported by numerous investigators using different scoring criteria and response modalities (Liben, 1974, 1975, 1978; Liben & Golbeck, 1980; Signorella & Jamison, 1978; Thomas & Jamison, 1975), developmental research seems to indicate that, although these differences are well established in adolescents and adults, they are inconsistent in studies of young children.

There is empirical evidence suggesting that sizeable gender differences appear approximately around the age of 10 (5^{th} to 6^{th} grade). Indeed, Liben (1975) failed to replicate the gender-differences in her nursery sample as well as Abravanel (Abravanel & Gingold, 1977) and DeLisi (1983), testing children from 7 to 11 years of age. McGillicuddy-DeLisi (McGillicuddy-DeLisi *et al.*, 1978) found a gender difference only with 10-year-olds which is also the age threshold where performance of boys and girls deviate substantially in Thomas's study (Thomas & Jamison, 1975).

Remarkable consistency with these results has been obtained from recent studies using a very different methodological framework (log linear analysis/binomial models) (Lohaus *et al.*, 1996; Thomas & Lohaus, 1993; Thomas & Turner, 1991). In an effort to explain individual differences in WLT, researchers approached the WL problem as a latent class problem and hypothesised that within each sex and generally at every age there are two classes of performers, poor and good. According to this conception, between-sex differences reflect within-sex differences as the proportion of good and poor performers vary within each sex, with larger proportions of good performers seen in the male group.

These researchers testing this hypothesis found the following results: Substantial individual differences seem to emerged around 9 years of age while at younger ages they are less noticeable, as one could argue that children become increasingly differentiated with development. From the age of 10 onwards there is clear and statistically significant evidence for two groups of performers, while the evidence is less convincing prior to this age. This results in the magnitude of the between-sex differences to be somewhat uncertain during the early primary years but well established by late childhood and early adolescence.

3.3 Cognitive style

Results from the correlational patterns and the analysis of deviation scores suggest that FD/FI is an important factor in WLT performance. In the overall sample, performance on EFTs correlated significantly with each of the horizontality tasks, and children who were better able to resist field effects produced significantly lower deviant scores. Consistent with our prediction, as the degree of perceptual misleadingness of the proximal frame decreases in the Bowl task, so does the strength of these relations. Consequently, the magnitude of correlations by age and gender varies across the tasks.

In particular, while from the age of 5½ the magnitude of correlations between performance on Bottle task and EFT was strikingly similar and significant across the age groups, it was significant only at the age of 8 for the Bowl task. On the contrary, the performance of 4 year olds in either task did not correlate with EFT which is consistent with the claim that cognitive style is not a predictable factor in WLTs in young children. Although the majority of studies lump together several age levels (review of studies in Pascual-Leone & Morra, 1991) to observe the power of this relation at each level, Foorman (1981), testing 4 year olds with the same measures as in the present investigation, obtained the same results (tilted bottles-PEFT; *rho* = -.02).

The significant correlations between horizontality measures and EFT for the boys are similar to the observations of previous researchers. However, the lack of correlations for girls is a gender difference that has not been found with older children, with an exception the study of Signorella (Signorella & Jamison 1978). This inconsistency can be explained on the basis of the instruments used to measure FD/FI. First, like in the present study, those which did not find a gender difference on cognitive style measured it with the EFT (Signorella & Jamison 1978; Liben, 1978) while those which reported significant differences in favour to the boys used the RFT (Abravanel & Gingold, 1977; DeLisi, 1983). The latter failed to find a gender difference in the relevant correlational patterns although the former obtained irreconcilable results.

Linn and Kyllonen (1981) suggest that the WLT is more contextually similar to the RFT than it is to the EFT, since it involves the orientation of lines in space which directly calls upon a conceptual spatial reference system. Thus, if the RFT measures spatial competence, of which disembedding ability is just one facet, between-sex differences in performance on cognitive style measures should be more consistently found using RFT than EFT, while the correlations between the former and WLT should be closer to unity for both sexes. Sufficient ground for this assumption has been given by Hyde (Hyde *et al.*, 1975) who obtained gender differences with the PRT but not with the EFT, and empirically proved that these differences were manifestations of gender differences in general spatial ability.

3.4 Contextual misleadingness: Task effect

With respect to the container's shape, the ovoid bowl –with its lack of a conflicting rectangular frame– was expected to be more conductive to the use of external referents. Substantial empirical support exists for the facilitating effect of non-rectangular vessels on performance in WLT (Abravanel & Gingol, 1977; DeLisi *et al.*, 1995; Thomas & Jamison,

1975) even in children as young as 5 and 6 years old (Russell & Maxwell, 1980). Only Liben (1975) found no significant difference between a round and straight bottle testing 17 year olds. Contrary to our prediction, no significant differences in the magnitude of angular distortions were obtained as a result of vessel's shape. There are three possible explanations for the absence of a task effect which will be presented and assessed in turn.

3.4.1 Order of presentation

First, the potential advantage of the Bowl relative to the Bottle task might have been attenuated by the fixed order of task presentation. If the absence of salient features (right-angle and parallelity cues) in the Bowl task inhibited task-irrelevant responses and activated task-relevant ones, children might have transferred the benefit gained from the Bowl to the Bottle task resulting in similar performance across the two testing conditions. This hypothesis presupposes the existence of the requisite conceptual system, which once activated, could be applied to a different context.

It seems, however, that there is not solid evidence in support of this claim. Practice or order effects have not been substantiated by studies using containers with different shapes and/or at various oblique positions (Abravanel & Gingold, 1977, study 3; DeLisi *et al.*, 1995; Thomas & Jamison, 1975). Two studies which found order effects, paradoxically enough, reported improvement in performance only from the least facilitating condition to the most facilitating one (bottle-cylinder: Abravanel & Gingold, 1977, study 2; straight jar-round jar: Randall, 1980). In addition, under more explicit testing conditions, where the effect of training had been investigated, children's benefit at the pre-test was not transferable to new bottle's shapes and orientations, and the effect was generally limited and task specific (Beilin *et al.*, 1966; Robertson & Youniss, 1969; Sheppard, 1974). All these findings seem to suggest that the transference of knowledge interpretation has limited explanatory power for the absence of task effect.

3.4.2 Cognitive style: A utilisation variable

The second explanation is connected with children's ability to overcome field effects. It is likely that FI operates by enhancing the ability to utilise knowledge of Euclidean co-ordinate rather than the acquisition of spatial operations on tasks like the water level problem. Although DeLisi (1983) considered both cases possible, a number of researches prefer to view cognitive style as a utilisation variable. Overton and Newman (1982) present studies which, investigating formal operational competence, found that latent and non-competent groups were not differentiated in terms of cognitive style status but competent ones were clearly field independent. A similar view is held by Shooll and Liben (1995) and Thomas and Lohaus (1993) who, on the basis of their results, maintain that both good and poor performers are influenced by the contextual misleadingness of WLT. However, those who are less susceptible to field effects, are able to override them by applying operational spatial schemes. Since children of the present age range seem to have either immature or incomplete spatial concepts, they are less likely to benefit from a less competing context. This seems to be supported by the absence of an interaction between task and cognitive style.

3.4.3 Stimulus structural features

Finally, the comparable performance across the tasks can be explained in terms of the bowl's structural features. Previous researchers have pointed out that round containers significantly improve WL performance only when all straight lines have been eliminated from their outline (Russell & Maxwell 1980; Willemse & Reynolds, 1973); a qualification which Liben (1975) referred to in explaining the absence of bottle shape effect in her study. However, the present results showed that this condition is necessary but not sufficient.

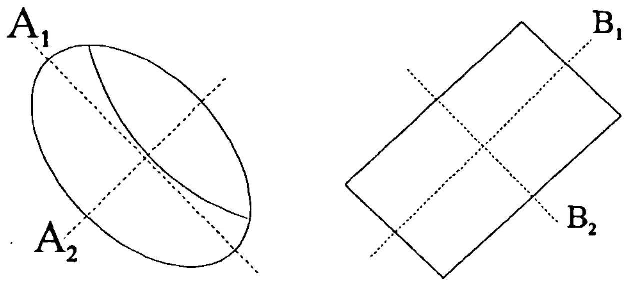


Figure 9.3 The structural axes of the Bowl (A) and the Bottle (B).

As Figure 9.3 shows, although the bowl is devoid of straight lines, it has a prominent structural axis and an explicit inclination, made apparent by the orientation of its visible spout. On the contrary, the circular containers used in the previous studies lack extendedness and have a less pronounced inclination. It is in fact these differences in the formal attributes between an elliptical and a circular container which can account for the irreconcilable results. Indeed, McAfee and Proffitt (1991) maintained that the orientation of the container needs to be indiscernible to enhance horizontality output in WLT. Thus, both the absence of non-straight lines on container's outline and the lack of explicit inclination and dimensionality of its shape need to be satisfied to facilitate performance relative to the traditional rectangular bottle.

Although there was no difference in the magnitude of angular distortions and the rate of success between the WLTs, there was a pronounced difference in the direction of deviation. The findings revealed that not only the containers elicited errors with an opposite directional bias from the horizontal axis, but also different magnitude of errors within directional category. Comparing the structural properties of the Bowl and the Bottle in Figure 9.3, it becomes apparent that both have (a) a clear inclination to the left, (b) a symmetric configuration along two intrinsic axes (A_1-A_2 and B_1-B_2), and (c) a distinctive dimensionality. However, they are incompatible in respect to their main structural axis since, when the vessels are at rest, the bowl's volume extends along the horizontal dimension (A_1) while the bottle's extends along the vertical (B_1). As a consequence, in the oblique position, the displacement of the bottle's main axis is congruent to the direction of its tilt whereas the displacement of the bowl's axis is incongruous to its tilt.

Only errors in the direction of the frame's tilt have been typically regarded as field effect errors, while empirical evidence suggests that errors in the opposite direction represent a very small portion of the total responses, and possibly comprise overcompensating attempts for the effects of the frame (Lohaus *et al.*, 1996). Surprisingly enough, although both containers yielded angular scores with a strong directional bias, only those in the Bottle task were in the expected direction. The fact that the Bowl also elicited a dominant directional preference does not leave any ground to deny its source in field influences.

To be able to understand why the field effects have an opposite result in the direction and magnitude of angular deviations across the tasks, we have to construe the various sources of field influences and weight their relative strength in each task. Four factors seem to operate:

(a) a geometric or production force will pull the responses towards the perpendicular to bottle's sides (Ibbotson & Bryant, 1976; Sommervile & Cox, 1988),

(b) the conceptual schema that the water is at the bottom and parallel to it will make its base salient and bias the angular deviations to the positive direction (Pacual-Leone & Morra, 1991),

(c) a perceptual force will make children draw the water line leaning in the same direction as the tilt of the vessel (McAfee & Proffitt, 1991; Vasta *et al.*, 1994; Sholl & Liben, 1995)

(d) and finally, a figurative factor will foster the operation of the Gestalt principles of 'good form' and/or symmetry (Bremner, 1985, Bremner & Taylor, 1982; Freeman 1980; Pacual-Leone & Morra, 1991).

The first three forces will shift deviation scores to the positive direction, and their relative strength will determine the magnitude of errors. Finally, the fact that both shapes are symmetrical, will set the operation of the gestaltist factor which in turn will prescribe the use of either of the vessels' intrinsic axes as alignment cues in orienting the level of the water. In the case of the Bottle, the child can draw the water level either parallel to the base or parallel to the walls, achieving symmetry which applies both to a local angular relationship (on the walls or bottom in respect), and to the overall figure. Although these alignment patterns produce errors with opposite direction (□to the walls: negative errors, □to the base: positive), the former strategy was never used possibly because it would have been inconsistent with the pictorial example provided. Since all the aforementioned factors elicit compatible directional preferences in the Bottle task, positive errors will be dominant across all classes of angular deviations and only an accurate conceptual strategy will reduce the potency of the positive bias. In fact, this is exactly what the data suggests. In all inaccurate angular categories (compromise to excessive), positive errors were prevalent, while within the accurate one, scores were equally divided into positive and negative.

The dynamic interplay of the field factors in the Bowl task is of a different nature. First the geometric force which will instigate perpendicular tendencies is attenuated by the lack of straight lines on bowls' contour whereas the conceptual and perceptual forces will still bias errors towards the positive direction. However, the gestaltist principles will potentially elicit two drawing strategies which are not directionally compatible. The child can either draw the water level along the bowl's principal axis – align it with the bottom of the bowl– or along its subordinate axis – align it with its 'walls'. Both solutions will 'bisect' its shape and contribute to figurative symmetry, but the former will produce errors with a positive slope and the latter errors with a negative one.

One would expect that, since the conceptual and perceptual constraints favour the positive direction and since the outline of bowl's rim offer an additional alignment cue, the most salient symmetrical solution would have been the one compatible with a positive inclination. Contrary to this prediction, children aligned the water line along with the bowl's walls to the opposite direction of its tilt. Possibly, this solution brings better symmetry to the overall figure as, with a pictorial example at hand, children are more likely to start drawing the water line from a point within the lower part of the base than the raised one. It might also be possible that children in general prefer to accomplish symmetry via the subordinate structural axis (A_2, B_2 in Figure 9.3). To be able to investigate these possibilities we need to monitor the starting location of strokes, use various containers with a different structure and include testing conditions with water lines present and absent from the pictorial example.

The dominance of the figurative forces over the cognitive and perceptual ones is expected to vary depending again on the cognitive strategy children adopt. In fact, the data suggests that when children have immature spatial schemata, there is a balance between the antagonistic forces. Angular deviations, which fell in the context-driven and excessive category, were as likely to be positive-driven as negative-driven. However, deviations within the accurate and compromise category are predominately negative in direction.

4. CONCLUSION

In summary, age related levels of performance were evident in both WLTs. Children's conceptual grasp of the horizontal co-ordinate is still not fully developed at the age of 8, consistent with Piagetian theory which expects these children to have reached sub-stage IIIA. However, the assessment of the concept of horizontality with a copying task improves the performance of the older children relative to previous studies. The fact that half of 7 and 8 year olds appreciated the relevance of the pictorial example points towards a competence-utilisation interpretation of the performance of the older children and a competence-deficit one of the younger children.

If the Bottle task is considered as the traditional way of testing the representation of horizontal invariance, these results provide evidence of the validity of the Bowl task as an alternative measure. They further corroborate the Piagetian theory in that a general, conceptual, spatial process that develops in a stage-like fashion cuts across tasks with different performance requirements. Although, the ability to overcome perceptual misleadingness is conducive to accuracy, in accordance with DeLisi's conclusion, "a conflicting frame of reference is not a necessary component of horizontality assessment in grade-school students" (DeLisi *et al.*, 1995, p. 240).

The WL problem introduces a conflict at many different levels and the provision of a non-rectangular container does not seem to reduce it. Performance on WLT is determined by a number of processes acting in opposition. As Pascual-Leone puts it "a dynamic conflict between sets of continuous contradictory processes is at the root of the water level task acquisition" (Pascual-Leone & Morra, 1991, p. 239). The pattern of directional preferences and the variation in the size of obtained frequencies across the drawing strategies by age indicates that the developmental progression is reflected in the prevalence of different factors at different age levels. In the next chapter, the acquisition of a Euclidean spatial system in children's pictorial representations will be assessed with respect to the concept of verticality. It will be of interest to examine whether the results in the WLTs will be comparable to those obtained in the verticality tasks.

CHAPTER X

VERTICALITY ASSESSMENT

1. GENERAL METHOD

1.1 Overview

Children's ability in depicting the vertical co-ordinate was assessed by three tasks with varied characteristics and requirements. In the first two, children had to copy the target stimulus which was anchored on a proximal frame of reference with different degrees of conflict to the spatial co-ordinates. In the first task children had to copy an upright tree on an round-shaped outline of a mountain whereas in the second they had to copy a chimney on a triangular-shaped roof. The third task differed from the other two in that children had to draw trees along the oblique edges of a road from memory.

It was hypothesised that the more the immediate frame competes with the environmental co-ordinates (roof, road), the higher the demands on children's disembedding abilities would be in extracting the appropriate axis (Liben, 1978; MacKay *et al.*, 1972). In addition, if children's failure on Piagetian tasks is related to a specific reproduction difficulty in constructing acute angles (Ibbotson & Bryant, 1976), the more linear the oblique proximal frame is, the greater the children's difficulty should be in overcoming their perpendicular bias. Further, if children's failure in representing the vertical co-ordinate stems from a general conceptual immaturity, the provision of the correct solution in the Mountain and Chimney task should have no effect on those children who lack the spatial concept of verticality (Liben, 1981). On the contrary, drawing from memory should be more difficult compared to copying tasks for children who possess the requisite co-ordinate axis system but have difficulty accessing and directly applying this knowledge. As a consequence, these alternative tasks were used to empirically investigate whether a conceptual spatial process underlies performance across tasks with different requirements, and to disentangle the factors hindering the demonstration of such conceptual abilities.

1.2 Participants

The sample consisted of the same children who participated in the previous experimental tasks. A description of their characteristics can be found in Chapter V, section 2.1. Due to the elimination of equivocal reproductions from the subsequent analysis of results, there is some moderate fluctuation in the sample size across the tasks.

1.3 General scoring criteria

The criteria adopted to assess the representation of the vertical spatial co-ordinate were the same with those used in the horizontal tasks. A best-fit straight line was constructed for the outline of the trees and the chimneys and the angle between this line and the vertical was measured in degrees with a hand-held protractor. When the trunk of the trees and the chimney had been drawn with two lines, an average estimate of orientation was taken. However, the critical feature for two-dimensional forms was that they had to extend to one dimension; thus circular-like shapes were excluded from analysis.

The deviation scores were then treated both as absolute and as directional, defining figural orientations away from the vertical and towards the perpendicular as positively biased and angular deviations way from the vertical and in the opposite direction to the perpendicular as negatively biased (see Figure 10.1) (Bremner & Taylor, 1982). A categorical scoring system was also used to identify the principles governing different error patterns from the magnitude of angular deviations, and to observe any age related difference in the strategies the children use to represent verticality. Verticality scores were clustered into four ranges of angular deviations in increasing order of severity: (a) *accurate*, (b) *compromise*, (c) *context-driven* and (d) *excessive.* The quantitative definition of these categories was the same as for the horizontality assessment (see Chapter IX, section 1.5).

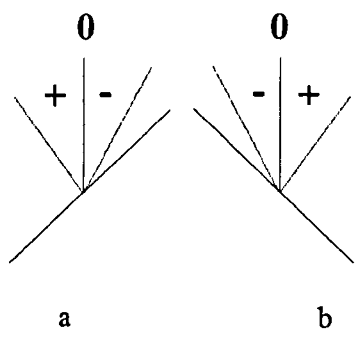


Figure 10.1 Directional errors scoring related to the vertical for left (a) and right slope (b)[1]

Since in the vertical tasks the target item was attached on a oblique baseline rather than embedded in a titled frame, as in the horizontality tasks, the use of the provided stand-lines and the quality of attachment were also considered in assessing the effect of the non-horizontal ground-line in figural orientation. Drawings where the requested forms were floating in undifferentiated space or were attached anywhere other than to the oblique outlines (imaginary horizontal ground-line, apex of mountain or roof), were excluded from any subsequent analysis. The reasons for eliminating forms lacking contact with the baseline were based on the belief that they either reflect the absence of any reference lines in guiding the spatial composition of the array or comprise an evasive solution to the construction of an acute angle. Empirical support for this contention comes from the study of Wilson and Wilson (1982) who found that floating figures above the outline of a hill were guided by the vertical factor while as they approached its surface they were pulled towards the perpendicular.

[1] For the Chimney and Road tasks two line drawings were obtained for each baseline orientation. Hereafter the left slope of the chimney/road will be referred alternatively as right oblique (RO) and the right slope of the chimney/road as left oblique (LO).

Lenient criteria were adopted in evaluating the quality of attachment of the requested form to the baseline accounting for the younger children's poor motor control. Therefore, lines failing to meet the target base by no more than a 3.00mm gap were considered as cases of incomplete attachment, guided by alignment principles, rather than as inadequate (immature) representations. These cases were included in the analysis of vertical orientation. Due to differences in the nature and characteristics of the drawing tasks, each study on verticality will be presented separately in the following section.

2. THE MOUNTAIN TASK

2.1 Apparatus and procedure

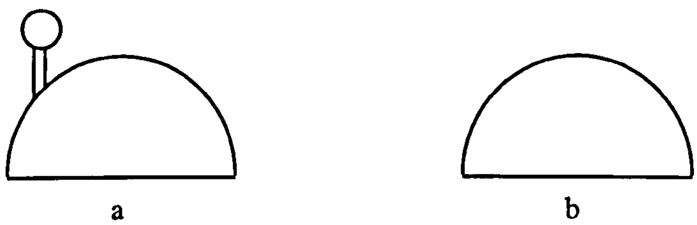


Figure 10.2 Schematic illustration of stimulus material.

Children were presented with a line drawing picture (Figure 10.2-a) depicting a mountain as a closed ellipse, 6.00 cm in diameter and 3.50 cm high at its tip. A tree, 3.00 cm tall, with an upright orientation was drawn to the left of the mountain's outline. The angle formed by the tangent passing from the point where the tree met the outline of the ellipse and the baseline of the ellipse was 50°. Children were provided with the outline of the mountain, pre-drawn on their paper (b). The outline drawing was centred on the stimulus card and on children's drawing papers. Children were requested to copy the tree that was missing from their paper exactly as it looks at the stimulus card. No explicit reference to the uprightness of the tree was made at the instructions. Children were not allowed to rotate their paper during drawing.

2.2 Results

Six drawings of 4 year olds were excluded from the analysis of results. In three drawings the tree was placed next to the horizontal base of the mountain, in one on its apex while in the remaining two the trunk was drawn with a circular shape. A 4 (age) × 2 (gender) × 2 (cognitive style[2]) unweighted ANOVA was conducted on the 104 mean absolute deviation scores. The means and standard deviations for this test are reported in Table 10.1. The analysis yield a significant main effect of age ($F_{3,88} = 29.43, p<.001$) with children producing less deviant scores off the vertical axis as they are getting older. As Table 10.1 indicates, this age difference is only attributed to the performance of the 4 year olds whose average absolute score lies well beyond the vertical (23.73) while the drawings of the remaining age groups were consistently close to the true vertical.

Table 10.1 Mean absolute deviation scores on Mountain task by age, gender and cognitive style.

	Boys				Girls				Total	
Age	N		FD	FI	N		FD	FI	N	M
4	8	6-2	22.17 (15.29)	40.50 (21.92)	16	8-8	17.62 (4.66)	14.62 (9.60)	24	23.73 (12.87)
5½	10	5-5	6.00 (5.96)	5.20 (5.76)	15	9-6	5.89 (4.54)	6.67 (10.25)	25	5.94 (6.63)
7	8	3-5	14.00 (10.44)	3.80 (4.44)	20	11-9	7.00 (5.10)	5.56 (4.10)	28	7.59 (6.20)
8	10	4-6	7.50 (5.07)	3.33 (2.16)	17	11-6	5.27 (4.56)	2.33 (3.39)	27	4.61 (3.80)
Total	36	18-18	12.42 (9.19)	13.21 (8.57)	68	39-29	8.94 (4.71)	7.29 (6.83)	104	10.46 (7.32)

Note 1 *The N is broken down by cognitive status style. The value to the left denotes the number of Field dependent children and the one to the right the number of Field independent ones.*
Note 2 *The numbers in parentheses are standard deviations.*

The ANOVA also revealed a significant main effect of gender ($F_{1,88} = 9.24, p<.01$) with boys performing worse than girls (12.81 vs. 8.12) and two significant interactions of age with gender ($F_{3,88}$ = 4.74, p<.01), and cognitive style ($F_{3,88}$ = 3.27, p<.05). According to the first interaction, girls were doing better than boys in all age levels except at the age of 5½ where boys slightly surpassed girls (smaller deviations). However, the performance of boys and girls within each age level is comparable except among the 4 year olds where the gender difference is the largest (Girls: 16.12 vs. Boys: 26.75). According to the age × cognitive style interaction, there was no performance difference between the two cognitive style groups up to the age of 5½, while from the age of 7, field dependent children produce more deviant scores than their field independent counterparts.

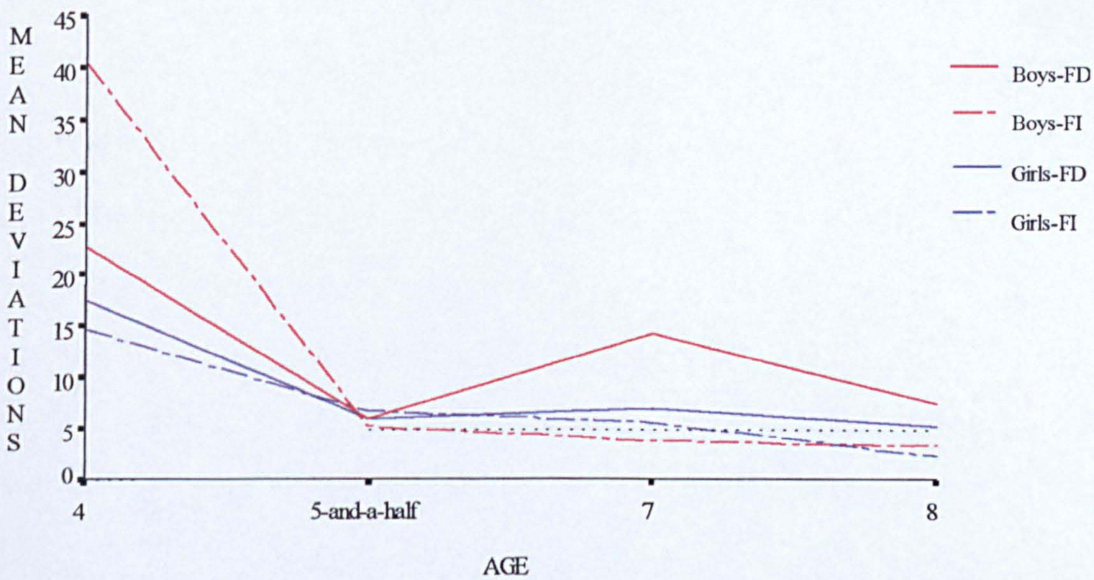


Figure 10.3 Differences of mean angular deviation scores across age levels as a function of gender and cognitive style, FD = field dependent, FI = field independent.

Finally the age × gender × cognitive status interaction was also significant and it is displayed in Figure 10.3. The three-way interaction indicates that within the field independent group, the gender difference in performance was minimal from the age of 5½ and onwards, whereas this difference was substantial at the age of four in favour of girls who outperform field

[2] For the classification of children into two cognitive style groups (FI/FD) refer to Chapter IX, section 2.3.

independent boys (smaller mean score). However, this result should be interpreted with caution since there were only two four year old boys in the FI group. Finally, with respect to the field dependent group, the performance of boys fluctuates widely across the age levels. In contrast, FD girls demonstrate marked improvement from the age of four to the age of 5-and-a-half, with performance stabilising afterwards. This resulted in gender differences for the field dependent children being the largest at the age of 7 and the smallest at the age of 5½.

Further analysis was conducted on directional data excluding 11 drawings where trees had 0° orientation with respect to the vertical axis. In general, children demonstrate a dominant directional bias in the predicted direction; in 58 (62%) drawings, the trees had a positive slope, while in 35 (38%) the trees were tipped in the opposite direction towards the mountainside. An analysis on the total angular deviations with direction as the independent factor showed that although there was no significant difference in the means between slanted trees with a positive slope compared to those with a negative one (11.07 vs. 8.26 respectively), the standard deviation of the former was double that of the latter (11.24 vs. 5.65, Levene's test of equality of variance, $F = 8.29, p<.01$). The present results indicate that, in general terms, both directional tendencies yield average scores within the vertical sector (0°-11°) but a positive bias cause greater variability in the distribution of errors. On the basis of the different variability in the distribution of positive and negative biased errors, two separate analyses were conducted for each direction.

The 4 (age) × 2 (gender) × 2 (cognitive style) unweighted ANOVA on deviation scores for each directional category[3] yielded only a significant main effect of age (positive biased score:

[3] A 4 (age) × 2 (gender) × 2 (cognitive style) unweighted ANOVA on signed errors (0° scores excluded) was also computed and the same results were obtained with the equivalent analysis on absolute errors for each direction. The main effect of age and gender, and their interaction remained significant while the interaction of age with cognitive status felt short of significance ($F_{1,77} = 2.41$, p=.07).

$F_{3,42}$ = 13.52, $p<.001$, negative biased scores: $F_{3,23}$ = 3.29, $p<.05$). Follow-up study revealed the same pattern of age differences with that on total scores; regardless of directional bias, 4 year olds perform significantly worse than any of the older age groups whose mean directional errors fell within the vertical sector (0°-11°) (Newman-Keuls test, $p<.05$). In addition, comparison of errors across directional category within age group – whilst it didn't yield significant differences[4] – revealed that only for the younger group was the difference relative bigger, with the mean and standard deviation of trees with a positive slope both larger than for those with a negative slope (M: 22.47, SD: 14.65 vs. M: 15.00, SD: 5.79 respectively). This suggests that the aforementioned larger variation in the distribution of positively biased errors is mainly attributed to the performance of the younger group which was more susceptible to perpendicular forces.

Two more results from the ANOVAs are worth mentioning. The analysis on positive biased errors revealed a significant interaction of age × gender ($F_{3,42}$ = 3.41, $p<.05$) according which boys performed worse than girls only in the younger age group whereas at the age of 7 there was no gender difference in performance. The analysis of negative biased errors showed that although the main effect of cognitive style status was short of significance ($F_{1,23}$ = 3.83, p = .063), FD children tipped their trees more towards the mountainside than did their FI counterparts (Means: 11.12 vs. 5.79 respectively).

To obtain a better understanding of the principles guiding performance on the Mountain task, the data was analysed on the frequency of vertical and non-vertical drawings with respect to the angular criteria of magnitude of deviation. The frequency data in Table 10.2 outline the same developmental pattern as the one obtained from the analysis of variance on mean scores. Children older than four years performed at ceiling level, depicting the tree with

[4] Analysis of difference of means for two independent samples yielded t values from 1.45 to .81.

accurate orientation (0-11°) while an abrupt shift from compromise to accurate drawings took place from the age of four to the age of 5-and-a-half [$\chi^2(1) = 15.19$ with continuity correction, $p<.001$]. In general, the task elicited a high incidence of accurate orientations and a narrow range of deviant responses illustrated by the occurrence of mainly compromise errors.

Table 10.2 Distribution of angular deviations across drawing categories by age.

		Drawing categories of angular deviations				
		Correct	Incorrect			
Age groups	N	Accurate 0°-11°	Compromise 12°-33°	Context-driven 34°-50°	Excessive >50°	Combined ≥12°
4	24	7	15	1	1	17
5½	25	22	3	0	0	3
7	28	24	4	0	0	4
8	27	25	3	0	0	3
Total	104	78	25	1	1	27

According to the directional preference, there is no evidence that young children are biased towards the direction of the perpendicular as only two children of the nursery group made context-driven and excessive drawings with a positive slope. Further, the directional preference appears to be independent of children's age, success in the task, and cognitive style status since within each directional category there were comparable numbers of accurate and compromise drawings, and FD and FI children. However, the higher incidence of trees drawn with a positive slope in the overall sample and within the fail category (73% vs. 58%) and the FD group (69% vs. 53%) suggests that a directional preference exists and its source affects the results to a moderate extent in a predictable way. Finally, the performance on Mountain task does not seem to correlate with performance on EFT. All correlation coefficients for the total sample and for each age level and gender failed to reach significance with the sole exception of 7 year old children [$r(28) = -.41, p = .03$].

2.3 Discussion

Piagetian theory predicts that children younger than 9 years of age will not have developed a Euclidean conception of space and that their operatively immature concepts will affect the way they will perceive and represent reality. This theoretical premise predicts that children's drawings, when they are copying from an operatively advanced stimulus, will contain distortions because the cognitive structures necessary to assimilate perceptual reality are lacking. However, children's overall performance on the Mountain task does not seem to conform to this prediction. The general pattern of results indicates that children have no difficulty reproducing the vertical co-ordinate on a sloping baseline copying form a model. The ceiling level performance of children from the age of 5½, both in terms of success rates and magnitude of angular distortions, suggests that they should have developed the relevant spatial concept to be able to recognise the invariant verticality of trees and benefit from the provision of a pictorial example.

The comparable level of performance between field dependent and field independent children and the overall absence of a correlational pattern between the Mountain task with measures of cognitive style clearly demonstrate that disembedding abilities are not engaged in the task. This finding can be explained by the geometry of the task which reduces the degree of conflict from the immediate visual field. The fact that the target line is not fully embedded in a slanted context but rather attached on it, along with the elliptical shape of the mountain which provides a non-perpendicular referent without competing with the true vertical spatial co-ordinate, makes the recognition of vertical invariance independent of cognitive style status.

These testing conditions dramatically increase success rate and accuracy. MacKay (MacKay *et al.*, 1972) assessed the concept of verticality with a mnemonic task where children had to draw a tree on an isosceles mountain. Only 30% of 7 and 8 year old children passed the task compared to 89% in the present study. Also, the ability of children as young as 4 years of age to benefit from the facilitating nature of a copying task was supported by Perner (Perner *et al.*, 1984) who found a 22.5% of success rate in the sample of 4 year olds compared to 29% in this study. Although Perner used a more lenient scoring criterion of success (≤ 22.5°) and reports cross-sectional rates of success totalling across 10 vertical and horizontal variants to allow direct comparison, his conclusions are in line with present results.

Interestingly, the well-documented predisposition of younger children to orient trees perpendicular to hillsides has not been found in this study. Although a dominant directional bias had been displayed by the 'out-of sector' deviations (positive:73%, negative: 35%), positively-driven children did not err towards the perpendicular. The performance of 4 year olds provides the most compelling evidence against the notion that children's errors systematically favour right-angle relations either due to reproduction difficulties or to representational concepts when tested with a socially intelligible task[5]. Further, although different forces operate within each region at either side of the vertical, the attributes of the task render these forces comparable in strength since the performance of positively and negatively biased children do not differ in a dramatic way.

Analysis of directional data with respect to cognitive style showed that positive driven children were more likely to be FD than FI but both perform at equivalent levels. On the contrary, negatively biased children were equally likely to be FD or FI but the mean score of the former was shifted more towards the mountainside. A possible explanation for this differential performance is that either a structural (more space to the left than to the right of

[5] It has been claimed that perpendicular errors in children's drawings of trees and chimneys stem from their conceptual knowledge and perceptual experience that vertical figures grow perpendicular to the ground-line.

the vertical), geometric (perpendicular) or perceptual factor operate shifting the majority of trees' orientation to the positive direction. Field dependent children seems to be more susceptible to these forces, as it will be expected but, due to the propitious nature of the task, they could readily apply their Euclidean concepts and perform equally well to their FI counterparts. On the contrary, children immune to these factors are negative-biased with a non-dominant cognitive style status. However only FD children, possibly more influenced by the slope of the mountain as an alignment cue, overcompensate by pulling their drawings more towards it.

Finally, the difference in performance between the sexes in favour to girls was rather unexpected. However, the significant interaction between age and gender showed that this variation in the overall sample was mainly attributed to the performance of 4 year olds girls who produced substantially smaller deviations compared to boys, whereas from the age of 5½ both sexes perform at comparable levels.

Concluding, while the present experimental procedure deviates from the one used by Piaget and Inhelder, the results raise some questions to the authors' theoretical claim that pre-operational children, lacking a conceptual spatial system, are unable to profit from observation. The finding that there was absolute no age trend apparent on this task from the age of 5½ suggests that young children may indeed possess the requisite co-ordinate axis system but have difficulty accessing their concepts and applying them under certain conditions. The next study was designed to investigate if the same children tested in the presence of distracting perceptual alignment cues will demonstrate differential performance.

3. THE CHIMNEY TASK

3.1 Apparatus

Children were presented with a line drawing of two adjacent houses centred on the stimulus card. Each house was 3.00cm in total height (2.00cm facade, 1.00 cm roof) and had a chimney, 0.50cm tall, drawn on the mid point of the right slope of its roof (Figure 10.4-a). The roof was an isosceles triangle with base angles of 45°. All children were provided only with the left house of the array, pre-drawn on their paper (b).

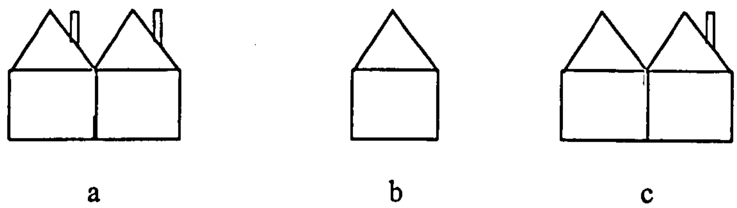


Figure 10.4 Schematic illustrations of layout (a) and pre-drawn forms (b, c).

3.2 Procedure

After matching the pre-drawn house with the corresponding one on the array, children were requested to draw its missing chimney exactly as it looks on the stimulus card. Nursery children were given an intuitively easier version of the task (Figure 10.4-c), prior to the standard version, where the whole array was reproduced on their drawing papers and the only feature missing was the chimney of the left house (two-house presentation trial).

Based on the variety of placement solutions that nursery children adopted in the Mountain task, the immediate comparison between the houses in the 'two-house presentation' trial should have directed children's attention to the appropriate placement and attachment of the requested chimney and should have emphasised further the task requirements. Besides, the provision of the whole array of houses along with the chimney of the right house should have conveyed the immediate context with additional vertical lines and should have subsequently facilitated chimney's accurate orientation. No further guidance was given and it was ensured that children were not rotating their papers while drawing.

3.3 Scoring criteria

Drawings were scored on the basis of the chimney's orientation to the roof's slope adopting the same criteria as in the Mountain task. A preliminary inspection of the drawings revealed that there were cases were the chimney had been placed inaccurately on the left slope of the roof. To account for the possibility that the orientation of the baseline (left and right slope of the roof) could have influenced differently the representation of vertical orientation, the drawings of chimneys were also scored for their placement with respect to the roof's slope, and the analysis of orientation was performed at first for each slope separately. Since no differences in performance was found as a function of baseline orientation, deviation scores from both slopes were collapsed and analysed together.

3.4 Results

The results from the nursery group will be presented first comparing performance between the 'two-house' and 'single-house presentation' trials of the Chimney task. Subsequently, the findings from the total sample will be reported from the 'single-house presentation' trial.

3.4.1 Nursery group

Table 10.3 summarises children's performance on each trial separately (left and middle column) and on both trials comparatively (right column). In the first trial, from a total of 30 children tested, five produced baseline-free drawings and more than half inaccurately placed the chimney on the left slope of the roof. The same children tested on the standard task produced four baseline-free drawings and, contrary to the first trial, they were accurate in placing the chimney on the right slope of the roof in 21 out of the 26 total drawings.

Considering chimney's orientation, an overall inspection of Table 10.3 suggests that the chimney's orientations on the right slope (left oblique) were less deviant than those on the left slope (right oblique) and that children produce less discrepant scores with the 'two-house presentation' trial than with the standard one.

Table 10.3 Means and standard deviations of chimneys' absolute deviations from the vertical by placement pattern and trial.

	Trials							
	2-house Trial		1-house Trial		Paired Performance			
Placement	N	M	N	M	Placement	N	2-house	1-house
Left slope	11	35.36 (17.35)	5	42.20 (24.42)	Both Left	4	36.00 (24.47)	52.25 (11.03)
Right slope	14	27.07 (15.32)	21	33.48 (11.30)	Both Right	12	26.17 (13.06)	34.25 (11.92)
					Different	7	30.14 (16.62)	28.43 (13.35)
Total	25	31.21 (16.33)	26	37.84 (17.89)	Total	23	29.09 (16.00)	35.61 (14.28)

Note 1 The figures in parentheses are standard deviations
Note 2 Free-floating forms were excluded from the analysis of angular deviations.

These differences were further investigated with statistical tests which yield the following results: A test of difference in the mean scores for the right and left oblique, for each trial separately, showed that chimneys' deviation scores for the left slope were not significant different to those for the right slope[6] (independent two sample *t* test); thus the orientation of the baseline had no significant effect on the representation of verticality. In addition analysis of the 23 drawings, where children produce representative drawings in both tasks, revealed that, although the provision of additional vertical axes and an accurate solution on children's drawing papers reduces the magnitude of deviation, children perform comparably in both tasks [$t(22) = 1.79$, $p = 0.08$]. However, due to some sparse indication that left slope was more difficult to the right, the analysis of paired differences was repeated excluding the seven cases where children used different placement patterns between the task. The analysis supported the superiority of the 'two-house presentation' trial in eliciting less deviant scores in cases where the same slope was chosen for chimney's anchorage for both trials [t(15) = 2.15, $p = .049$].

3.4.2 Total sample

A total of 106 drawings were analysed excluding only the four baseline-free drawings of the nursery group. The frequency of drawings with chimneys placed on the wrong slope was rather rare and mainly found among the nursery group (5 out of 7). On the basis of the negligible proportion of chimneys on left slope in the total sample, and of the absence of a significant baseline orientation effect in the preceding analysis, the results will be presented on the scores from both slopes collapsed. A 4 (age) × 2 (gender) × 2 (cognitive style) unweighted ANOVA was conducted on mean absolute deviation scores. The means and standard deviations for this test are reported in Table 10.4. The analysis yielded a significant main effect of age ($F_{3,90} = 34.61, p<.001$) with children producing scores closer to the vertical with age. Post-hoc Newman-Keuls tests ($p<.05$) indicated that 4 and 5½ year olds draw significantly more deviant scores than any of their older counterparts. Boys performed slightly better than girls (M: 14.82 vs. 16.15 respectively) and FI children were more accurate than FD children (M: 13.55 vs. 17.42) but neither of these main effects was significant nor any interaction. When the ANOVA was repeated only on deviation scores for right slope to reduce possible variation, we obtained the same results with the exception of the main effect of cognitive status style which just reached significance[7] (total scores: $F_{1,90} = 2.95, p = .09$; right slope scores: $F_{1,83} = 3.80, p = .055$).

[6] Two-house trial: $t_{23} = 1.27, p>.05$, Levene's Test for Equality of Variances $F = .015, p = .902$; one-house trial: $t_{24} = 1.22, p>.05$, Levene's Test for Equality of Variances $F= 3.502, p = .074$.

Table 10.4 Mean absolute deviation scores on Chimney task by age, gender and cognitive style status.

	Boys				Girls				Total	
Age	N		FD	FI	N		FD	FI	N	M
4	10	7-3	30.14 (15.29)	32.33 (10.21)	16	10-6	39.20 (15.97)	35.67 (13.94)	26	34.34 (13.85)
5	10	5-5	14.60 (10.67)	15.00 (8.12)	15	9-6	23.78 (12.19)	6.17 (7.39)	25	14.89 (9.59)
7	8	3-5	13.33 (9.89)	8.40 (11.61)	20	11-9	8.82 (8.52)	7.00 (10.59)	28	9.39 (10.15)
8	10	4-6	3.25 (2.63)	1.50 (1.52)	17	11-6	6.27 (9.48)	2.33 (2.06)	27	3.34 (3.92)
Total	38	19-19	15.33 (9.62)	14.31 (7.86)	68	41-27	19.52 (11.54)	12.79 (8.49)	106	15.49 (9.38)

Note 1 *The N is broken down by cognitive status style. The value to the left denote the number of Field Dependent children and the one to the right the number of Field independent ones.*
Note 2 *The numbers in parentheses are standard deviations.*

The drawings were further examined with respect to the direction of deviation from the vertical axis. Excluding the 16 drawings where children attained perfect scores (0°), the remaining 90 drawings demonstrate a dominant positive bias since in 73 (81%) of them the chimney was drawn with a tilt away from the vertical in the direction of the perpendicular. A test of difference in the means of negative and positive biased scores was significant [$t(88) = 4.05$, $p<.001$] and showed that the former were mostly near the vertical (M: 5.47) whereas the latter were moderately affected by perpendicular forces (M: 21.47)[8].

Since the distributions of positive biased and negative biased scores were distinctively different, two separate analyses of variance were performed for each direction of deviation[9]. A 4 (age) × 2 (gender) × 2 (cognitive style) unweighted ANOVA on positive absolute scores

[7] The means for the FD/FI groups were essentially the same, so it was the reduced number of the sample in this analysis which induced the statistical adequacy.
[8] Also Levene's Test for Equality of Variances was significant $F = 9.10$, $p = .003$ suggesting greater variability within positive biased scores.

revealed only a main effect of age ($F_{3,57}$ = 18.24, p<.001) with a gradual decrease of mean errors from 34.89 among the 4 year olds to 5.76 among the 8 year olds[10]. Due to the small number of negative biased errors only age as a factor entered into this analysis which just approached significance ($F_{3,13}$ = 3.33, p = .053). However, inspection of the 16 negatively biased scores showed that all children drew chimneys within the vertical sector with the exception of one of the two 4 year olds in the sample who produced a very extreme score (48.00). So, the age difference was only caused by a single score while the range of means for the children older than 4 years of age fell between 2.5 to 3.00 from the vertical.

Drawings were further classified into categories of responses on the basis of magnitude of deviation. Using first the pass/fail criterion, a test of association between age and success in the Chimney task was highly significant [$\chi^2(3)$ = 38.2, p<.001]. Comparisons between successive age groups revealed that there was an abrupt shift towards accuracy from the age of 4 to 5½ [$\chi^2(1)$ = 5.71, p<.05] with a smooth increase in the number of correct orientations after this age. The frequency data in Table 10.5 clearly shows that the 4 year olds drew mainly incorrect representation, the 5½ and 7 year olds drew comparable numbers of correct and incorrect orientations whereas 8 year olds children were mainly successful in accurately orienting the chimney with respect to the vertical axis.

Different age groups demonstrate also different drawing patters. The rather rare excessive drawings and nearly all of the context-driven ones in the overall sample were found among the 4 year olds whose modal response was context-driven. On the contrary, the frequency distribution of the drawings of 5½ and 7 year olds were rather bimodal with modes located in

[9] The ANOVA computed on signed errors (0° scores were excluded) revealed again a significant main effect of age ($F_{3,74}$ = 12.29, p<.001) and also a main effect of cognitive style ($F_{1,74}$ = 4.19, p<.05) with the mean error of FD children larger than that of the FI children (17.01 vs. 10.29 respectively).

the accurate and compromise category whereas by the age of 8 the distribution achieves unimodality clearly located within the accurate category.

Table 10.5 Distribution of angular deviations across drawing categories by age.

	Response categories of angular deviations					
	Correct	Incorrect				
Age groups	Accurate 0°- 11°	Compromise 12°- 33°	Context-driven 34°-50°	Excessive >50°	Combined ≥12°	Total
4	2	8	13	3	24	26
$5^{1/2}$	10	13	2	0	15	25
7	18	10	0	0	10	28
8	24	3	0	0	3	27
Total	54	34	15	3	52	106

The developmental pattern presented in the table above was essentially the same when the drawing patterns were considered according to children's directional preference. All except one of the incorrect drawings were positively biased suggesting that all out-of-vertical-sector orientations were to a lesser (compromise) or greater degree (context-driven) affected by the perpendicular factor, and the younger children were the most susceptible to it. This finding has two additional implications. First, there is an association between directional preference and passing or failing the Chimney task since nearly all of the children who failed the task were positively biased whereas no directional preference was seen among those who passed it (16-22), [$\chi^2(1)$ = 20.59, p<.001, with continuity correction]. Also, the directional preference is associated with age [$\chi^2(3)$ = 7.61, p = .055]. Children up to 5½ years of age were primarily positively biased (91.6%) as they were influenced more by the perpendicular forces, whereas, after this age, with an increase of accurate orientation, one third of their drawings demonstrate a negative bias [$\chi^2(1)$ = 6.08, p<.05, with continuity correction].Finally

[10] Follow-up study revealed a similar pattern of age differences with that on absolute scores; 4 year olds produce significantly more deviant scores than any of their older children and 5½ year olds perform significantly worse than the 8 year olds.

directional preference seems to be associated with cognitive style status. Whereas the majority of FD children (89%) tilted the chimney in the positive direction, FI children demonstrate no preference in the frequency distribution of errors at each side from the vertical [χ^2 (1) = 4.14, p<.05, with continuity correction].

Finally, the overall performance on Chimney task correlated with performance on EFT measures [r(106) = -.27, p<.01] while the correlation estimates by age group and gender showed that it was only at the age of 5½ and among the boys that performance on the task correlated with cognitive style measures [5½: r(25) = -.49, p = .012, boys: r(38) = -.42, p<.01]. A test of difference between the correlation coefficients between the sexes indicated that this difference was not statistically significant (z = 1.22, p>.05).

3.5 Discussion

3.5.1 Nursery Group

The performance of 4 years old children in the 'two-house' and 'single-house presentation' trials showed that they can benefit moderately from the facilitating version of Chimney task with ample parallelity cues (vertical reference axes). It seems that such a task reduces the magnitude of errors, especially for children who used the same slope in both trials to anchor the chimney, but does not increase the rate of success[11] (correct orientation).

This result supports previous findings from abstract and contextualised tasks which indicate that young children cannot radically benefit from the relationship of parallelism between the target line and contextual features whether within the array or in the frame (Bayraktar, experiments I & Ia, 1985; Freeman & Kelham, 1981; Perner *et al.*, 1984). It further supports

[11] A McNemar test of change on pass/fail scores was not statistical significant since only 4 children passed the 'two-house presentation' trial and failed the standard one compared to none who display the opposite pattern of inconsistency.

the premise from Piagetian theory that, in the absence of an integrated co-ordinate system, young children, liable to 'centration' effects, are guided only by the proximal reference line (baseline) without being able to recognise the relevance of other appropriate spatial referents. As Mitchelmore (1985) suggested "the salience of a proximal spatial relation is always likely to be greater than that of a distal relation" (p.307).

The 'two-house presentation' trial did not increase 'attached-to-the-baseline' responses and even had a negative effect on placement accuracy. With respect to the former, 4 year old children are either poor marksmen or are applying a loose criterion of proximity to represent spatial relationships. The difference in placement accuracy between the two trials, albeit unexpected, could be explained in the light of children's sensitivity to symmetry. It is well documented that young children are guided in their graphic representations by symmetrical forces often at the expense of visual realism. Possibly, the placement of the chimney on the right slope of the roof for both the left and right house in the stimulus card introduces a compositional imbalance. The younger children are given the opportunity to 'amend' this imbalance when presented with the whole array (two-house presentation) by inaccurately placing the requested chimney of the left house on the left side of its roof. This explains the result of obtaining, contrary to expectations, chimney placement errors in the 'two-house' compared to the 'single-house presentation' trial.

Finally, from those inaccurate placements, it seems that a chimney was somewhat easier to draw on an oblique baseline with an orientation to the left (right roof's slope) than on one with an orientation to the right (left slope). Bremner (1984) testing 4 year old children with an abstract task also found a marginal superiority of left oblique in aiding performance. However, the absence of a reliable orientation effect with oblique baselines is in agreement with results from abstract tasks on young children (Bayraktar, experiment Id, 1985; Bremner

& Taylor, 1982; Ibbotson & Bryant, 1976) where no significant difference was found between the two types of oblique presentations.

3.5.2 Total sample

Results from the Chimney task demonstrate a gradual incremental improvement in accuracy with age, indicated by a reduction in the magnitude of deviant scores, an increase in the proportion of correct responses, and a progressive use of more advanced drawing strategies. Therefore, the present testing procedure elicited cross-sectional data that fit to a stage-like model of conceptual spatial-processing ability. The smooth pattern of variation in the frequencies of different drawing strategies and modal responses across age levels elucidate further this developmental progression which takes the form of adjusting the strength of different guiding factors.

Although these findings are in general agreement with Piaget's observations about the drawing strategies and the gradual use of developmentally advanced drawing alternatives, there are certain inconsistencies with his model. Children at all age levels performed much better than the theory predicted. In particular, the perpendicular solution was abandoned from the age of 5½ while accurate orientations were the dominant response mode by the 8 years of age. These results are in line with Perner's study (Perner *et al.*, 1984) who adopted the same methodological approach (copying tasks), and support further the claim that this inconsistency results from procedural variations and that the copying tasks are not equivalent to the mnemonic tasks in the assessment of spatial concepts.

From Piaget's descriptive account of children's stages in the representation of verticality, we obtain the impression that the transition from perpendicular to accurate orientations is an all-or-none. There is a substantial body of empirical evidence (Thomas & Lohaus, 1993; Wilson & Wilson, 1982) against the notion that children's drawings reflect a clear binary choice between two reference systems; proximate and varied, distal and invariant. The large proportion of compromise responses in the total sample suggests that children often make an adjustment between the two incompatible spatial referents which appears to be the preliminary step of accuracy. Many theorists claim that these errors illustrate the cue-conflict situation that the Piagetian spatial tasks entail and demonstrate that children are not totally insensitive to spatial co-ordinates but are caught in a conflict between vertical forces and various other acting in opposition (Bremner, 1985; Freeman, 1980).

A number of findings from the present study illustrate the presence of different forces, the interplay of which is accentuated by the geometry of this task. First, the difference in the means and distribution of directional scores showed that although inclinations towards the edge of the roof (negative) are influenced by the vertical factor, inclinations to the opposite direction (positive) lay in a tension area (compromise), midway between the vertical and perpendicular. In addition, the overall correlation between angular scores and performance on EFT, along with the moderate difference (significant for chimneys on the right-slope) in performance between FD and FI children suggests that success in the Chimney task requires an ability to overcome conflicting cues. The significant association between directional preference with cognitive style status and rate of success further suggests that children producing positive inclinations are more likely to be context-driven and field-dependent.

Children of different ages, varying in the abilities having on their disposal, display a different degree of susceptibility to task relevant and task irrelevant cues. The obliquity of the baseline accentuates the operation of perpendicular forces and shifts the majority of scores to the positive direction in the overall sample. Four year old children, lacking an internalised system of spatial referents were readily influenced by the immediate spatial context and produced perpendicular errors. Children at intermediate stages of spatial ability were balancing antagonistic forces and produced compromise errors whereas older children were overcoming distracting forces having a more developed system of spatial co-ordinates and better disembedding abilities.

With respect to gender differences the issue is not settled. Both sexes perform comparably in the Chimney task whereas only the performance of boys seems to correlate with cognitive style measures. These results do not corroborate the claim that the gender differences in cognitive style produce the gender differences in the representation of vertical concept. However, the fact that the magnitude of difference between the correlation coefficients for boys and girls is not statistical significant may have contributed to the moderate superiority of boys' drawings compared to those of girls.

3.5.3 Between-tasks comparison.

Children in both studies were requested to copy from a model a vertical feature resting on a non-horizontal baseline, while the degree of conflict from the proximal alignment cue was increased in the Chimney task. The Mountain task was easier than the Chimney task for all children. This difference cannot be attributed to conceptual differences between the tasks since Piaget treated both as operatively equivalent in the assessment of the concept of verticality.

There is also a remote possibility that the figural differences between the tasks could have contributed to differential performance. Perner (Perner *et al.*, 1984) considered the possibility that the modest superiority of the tree as opposed to the Chimney task in eliciting better performance could have resulted from differences in size between the target objects, and between the proximal alignment referents (experiment 1). He found no evidence to support this claim, as children's performance did not change substantially as a result of using stimuli matched in size (experiment 2).

Further, based on the weak 'baseline orientation effect' obtained from the analysis of data from the nursery, the fact that the tree was placed on the left side of the mountain whereas the chimney on the right side of the roof could also have contributed to differences between the tasks. However the direction of baseline effect is not in line with the direction of tasks' difference. Since a left oblique appears to be easier than a right oblique, the Chimney task should have been easier compared to the Mountain task.

Finally, the fact that perpendicular errors were found in the chimney but not in the Mountain task demonstrates that this error cannot be conceptual – as has been claimed elsewhere – (eg. Arnheim, 1969). If children's perpendicular drawings result from their conceptual knowledge and perceptual experience that vertical figures grow perpendicular to the ground-line, perpendicular drawings should have been of equal frequency in both tasks.

Ruling out conceptual and figural differences as possible sources of between-task variation, the only obvious factor of differential performance is the degree of conflict between the invariant vertical and the proximal oblique referent. The differential effect of cognitive style status on performance across the tasks supports this assertion. In the last study of this chapter, we will investigate the pictorial representation of the vertical concept with a more demanding task. The critical issue is whether children, assessed with a predictive-mnemonic task, demonstrate equivalent conceptual levels of spatial ability.

4. THE ROAD TASK

4.1 Apparatus and procedure

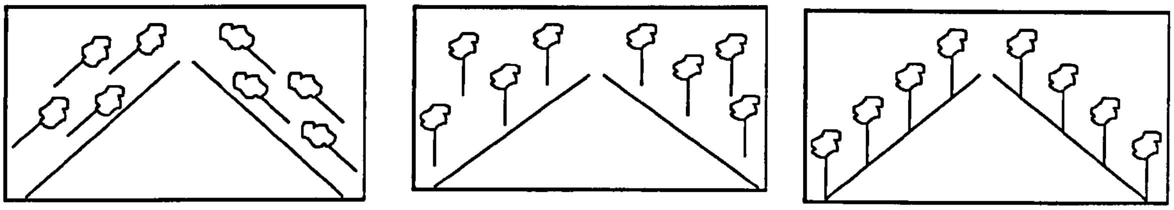
Children were presented with a configuration of two converging oblique lines, pre-drawn in the centre of their paper and oriented ±55° from the vertical axis. The lines were 7.00 cm long with a minimum distance of 3.00cm and a maximum distance of 11.00 cm. Children were told that the picture shows the edges of the road which is receding into the distance and were asked to draw four trees on each side to make the picture prettier. For each drawing, the production sequence was recorded by writing down the temporal order in which children drew the 8 tress. No further instructions were given except that children were emphatically told that they should drawn the trees *on* the lines and that they should not rotate their paper during drawing.

4.2 Scoring criteria

The same conditions were applied to decide upon the drawings eligible for analysis, and the same scoring criteria were used to examine the orientation of the trees[12]. To avoid the effects of perseveration and confusion with previously drawn orientations, the analysis was confined only to the first-drawn tree on each side of the road's edge. In the previous tasks children had to demonstrate a reasonable use of the provided baseline in attaching the requested forms, and free floating figures were excluded from the analysis of orientation. In the present task a more refined criterion was used since, due to the absence of a model, children adopted a number of various solution to the spatial composition of the whole scene.

A closer examination of the drawings revealed that the lack of attachment was not always indicative of a failure in the use of a reference axis. There were cases where children had not explicitly used the provided baselines for local anchorage but they had distributed the forms employing alignment cues either internal or external to the array (a & b in Figure. 10.5).

[12] The conditions of a form's linearity and its firm anchorage on the oblique baselines had to be satisfied for a drawing to be included in subsequent analysis which consisted of (1) measuring the degrees by which the tree's orientation departed from the true vertical (2) indicating the direction of tilt in terms of positive and negative slope and (3) assigning the drawings into deviant classes to identify drawing patterns. (For a detailed description of the general scoring principles and conditions refer to section 1.3 of the present chapter).



a b c

Figure 10.5 Examples of drawings without baseline anchorage with internal (a) or external spatial alignment (b), and with baseline anchorage with external spatial alignment (c).

As a consequence, drawings were totally excluded from analysis only when the forms were lacking solidity and alignment to any spatial frame. Two separate analyses were performed on the remaining drawings; first results were computed on the total sample of representative drawings and second only on those with baseline attachment. This was done to account for the possibility that the lack of attachment may either reflect children's inadequacy in constructing an acute angle or an attempt to reduce the extend of conflict from the surface's contact (this qualitative analysis would particularly differentiate drawings with correct orientations).

4.3 Results

From a total of 110 children tested, nursery children produced 6 drawings of trees for the left slope and 7 drawings for the right slope which were eliminated from the analysis, leaving a sample of 103 drawings with representative trees for both slopes. A 4 (age) × 2 (gender) × 2 (cognitive style) × 2 (slope's orientation) unweighted ANOVA with a repeated measure on the last factor was performed on the absolute deviation scores (Table 10.6). The results of the between subjects effects revealed only a significant main effect of age ($F_{3,87} = 7.14, p<.001$) while the within subject effect fell short of significance ($F_{1,87} = 2.25, p = .14$).

Table 10.6 Mean absolute deviation scores for each side of the road by age, gender and cognitive style status.

		Baseline Orientation	
			
Factors	N	Left slope (RO)	Right slope (LO)
A. Age groups			
4	23	50.92 (17.14)	50.81 (16.67)
5½	25	40.62 (18.01)	44.88 (20.42)
7	28	39.46 (19.93)	41.73 (15.88)
8	27	25.01 (18.99)	30.17 (20.44)
B. Gender			
Boys	39	39.61 (21.74)	40.91 (19.57)
Girls	64	38.40 (19.50)	42.88 (20.00)
C. Cognitive style			
FD	58	40.85 (18.61)	43.30 (19.40)
FI	45	37.16 (22.17)	40.49 (20.30)

Note 1 RO = Right oblique, LO = Left oblique.
Note 2 The numbers in parentheses are standard deviations.
Note 3 FD = Field Dependent, FI = Field Independent.

To locate the source of variance, posteriori comparisons for each baseline orientation revealed that only the mean absolute scores of 8 year olds were significantly different to any of their younger counterparts (Newman-Keuls, $p<.05$). These results simply suggest that verticality performance in the Road task improves smoothly up to the age of 7 while at the age of 8 there is a marked reduction in the magnitude of deviations from the vertical axis. Further, there is essentially no difference in performance between boys and girls, and a moderate difference between FD and FI children in favour of the latter. The parallelism in the pattern of the result between the two slopes is demonstrated by the absence of baseline orientation effect, although for all factors and within each level, the mean absolute errors for the left slope were smaller than those for the right.

Repeating the ANOVA only on drawings where the trees made contact with the surface of road's edge (N = 90), the pattern of the between-subject effects remained the same; the main effect of age retained significance ($F_{3,74} = 5.31$, $p = .002$) while all other main effects and interactions were non-significant. However, the superiority of the left slope in aiding performance was more pronounced since the difference in the mean absolute scores between the left and the right slope approached significance ($F_{1,74}$ = 3.66, p = .06). This analysis also yielded a significant 3-way interaction between age, cognitive style status and baseline orientation ($F_{1,74} = 3.03, p = .035$) which is presented in Figure 10.6.

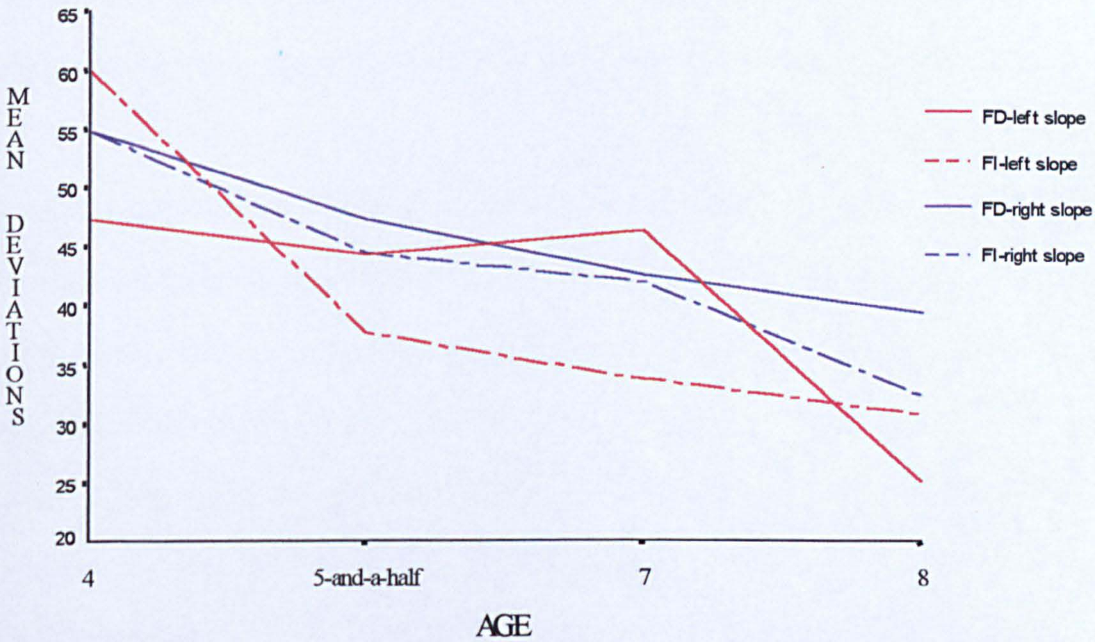


Figure 10.6 Differences of mean angular deviation scores across age levels as a function of slope's orientation and cognitive style[13].

As the figure shows, for the right slope, FI children outperformed (smaller deviations) their FD counterparts from the age of 5½, while for the left slope we obtain a wide variation in the performance of the cognitive style groups across the age levels. The difference in performance of cognitive style groups with respect to age and baseline's orientation had the following result; at the age of 5½ and 8 the right slope was more difficult than the left regardless of cognitive style membership. However, at the age of 4, the performance of both FD/FI children on the right slope was surpassed (smaller deviations) only by the performance of FD children on the left slope. The opposite is true at the age of 7 where the performance on the right slope was worse than performance on the left slope by FI children.

With respect to directional preference, children exhibited a positive bias in 79% of the drawings[14] where they consistently tilted the trees in the direction of the perpendicular for both edges of the road, with no significant change in directional preference across baseline orientations (McNemar test of change, $z = .09$). An independent t test was used to compare positive and negative biased scores within baseline orientation (bottom row of Table 10.7).

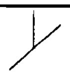
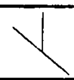
Considering the left side of the road, the absence of a significance difference in the mean positive and negative biased scores [$t(102) = 1.51$, $p > .05$, Levene's test for equality of variances $F = 2.47$, *ns*], and their magnitude suggest that the same forces (context-driven) operate for both directional categories. More specifically, errors with a positive inclination tended to be perpendicular to the left slope of the road and errors with a negative one tended to be parallel to it. On the contrary, the directional mean errors for the right side of the road were significantly different, since those having a positive tilt were more deviant than those having a negative one [$t(101) = 4.92$, $p < .001$, Levene's test for equality of variance $F = 2.15$, *ns*].

Each class of directional scores was further examined with factorial analysis for each slope of the road. Positive biased errors were analysed with age (4) × gender (2) × cognitive style (2) as factors. For negative biased scores, only age as a factor could enter into the ANOVA due to the small overall number of drawings with a negative inclination. The analyses revealed that age appears to be the main determining factor to account for the variability in

[13] The distribution of FD/FI children within successive age levels was 12/7, 11/13, 13/14 and 10/10.
[14] Excluding the 'baseline-free' drawings, children displayed directional consistency towards the perpendicular in 88% of the drawings.

verticality performance on Road task[15]. Thus for a more direct comparison between directional tendencies, mean positive and negative scores had been broken down only by age in Table 10.7.

Table 10.7 Mean directional scores by baseline orientation and age.

	Baseline orientation							
								
			←	→			←	→
Age groups	N		Positive	Negative	N		Negative	Positive
4	24	19-5	50.16	48.8	23	7-16	45.43	53.50
			(14.32)	(27.66)			(18.87)	(15.63)
5½	25	23-2	41.35	30.5	25	3-22	12.67	49.77
			(18.32)	(13.43)			(15.14)	(16.78)
7	28	28-0	38.89	-	28	3-25	11.33	44.64
			(19.93)				(6.81)	(12.49)
8	27	22-5	28.91	11.20	27	6-21	8.67	36.14
			(19.14)	(10.03)			(7.92)	(18.74)
Total	104	92-12	39.45	30.08	103	19-84	23.26	45.55
			(19.39)	(25.55)			(21.71)	(16.87)

Note 1 *The N is broken down by directional tendencies. For the left side of the road, the figure to the right denotes the number of drawings with a positive bias and the one to the left the number of drawings with a negative one. For the right side of the road the figures denotes the reverse directional tendency.*

Note 2 *The figures in parentheses are standard deviations.*

Developmental trends were present within each directional category for both baselines as children draw less deviant scores with age (left slope positive $F_{3,88} = 4.68$, $p<.01$, negative $F_{2,9} = 4.36$, $p<.05$, right slope positive $F_{3,80} = 4.30$, $p<.01$, negative $F_{3,15} = 9.13$, $p<.001$). Follow-up study (Newman-Keuls, $p<.05$) revealed that the developmental change in the magnitude of mean errors with a positive tilt was rather smooth. Children of all age levels were influenced by perpendicular forces, and only the performance of the 8 year olds was

[15] The analysis on positive biased scores yielded exactly the same pattern of results as the factorial analysis on overall scores. The only significant findings were the main effect of age for both the left and the right slope and the interaction of age with cognitive style for the left slope. The nature of this interaction has been presented in Figure 10.6.

significantly better than the performance of the 4 year olds for the left slope and of the first two age groups for the right slope.

The mean errors with a negative slope revealed a different picture across age groups. The performance of the 4 year olds was significantly worse than the performance of the 8 year olds for the left slope, and significantly worse than of any of their older counterparts for the right slope. Especially for the later, all children older than 4 were influenced by vertical forces. Also, there is some indication that the orientation of the baseline influenced differently the relative strength of the drawing forces at the age of 5½. For the left slope, context-drive forces are of the same impetus at each side of the vertical while, for the right slope, positive errors are susceptible to perpendicular forces while negative errors to vertical forces. However, this claim is purely speculative, considering the small number of negative errors.

The general pattern of the results in Table 10.7 suggests that with respect to age, the directional scores of the younger children are equally deviant from the vertical axis while older children are guided by different forces for negative errors (vertical forces) and positive ones (perpendicular forces). Also there is some indication that the orientation of the baseline affects differently the relative strength of the directional forces. It seems that the left slope reduces the influence of the perpendicular bias within positive-driven children compared to the right slope. On the contrary, the right slope appears to enhance the influence of the vertical factor within the negative-driven children.

The drawings of trees were further divided into correct and incorrect, using the criterion of 11° as a cut off point (Table 10.8). In general, children performed poorly in this task since correct drawings were rare before 8 years of age, while by this age only one third of the children demonstrated ability in using the vertical spatial co-ordinate to orient the trees. The orientation of the baseline did not essentially affect the number of correct drawings, neither did the ability to firmly anchor the trees on the provided baseline change the pattern of the results. Surprisingly the few 'baseline-free' depictions were relatively common among the youngest and the oldest of the children and they were equally present in the accurate and inaccurate drawings of the oldest age group.

Table 10.8 Distribution of correct and incorrect drawings by baseline orientation and age.

	Baseline orientation											
	Left Slope (RO)						Right Slope (LO)					
Age groups	Pass 0°-11°		Fail ≥12°		Total		Pass 0°-11°		Fail ≥12°		Total	
4	1	(1)	23	(2)	24	(3)	0		23	(4)	23	(4)
5½	2		23	(1)	25	(1)	3		22	(1)	25	(1)
7	5	(1)	23		28	(1)	2	(1)	26		28	(1)
8	10	(3)	17	(4)	27	(7)	7	(4)	20	(3)	27	(7)
Total	18	(5)	86	(7)	104	(12)	12	(5)	91	(8)	103	(13)

Note 1 The numbers in parenthesis represent the number of 'baseline-free' cases.
Note 2 RO = Right oblique, LO = Left oblique.

Inaccurate drawings were further classified into angular categories by degree of magnitude of deviation. Due to the predominate positive inclination in children's drawings of trees, this investigation will be restricted only to positively biased orientations to obtain a clearer picture of the drawing patterns. Table 10.9 shows that 66% of the total sample of positive drawings for the left side of the road and 80% for the right side were perpendicular and excessive drawings. In particular to the left slope, 4 year olds drew either perpendicular or excessive drawings, 5½ year olds predominately oriented their drawings perpendicular to the side of the road, while at the age of 7 children demonstrated a sudden shift towards excessive drawings with their rest of responses equally distributed across the remaining angular categories. Finally, at the age of 8 excessive drawings were reduced and children displayed no dominant drawing pattern. A somewhat different pattern is revealed for the right side of the road across the age groups. The frequency distribution at the age of 4 and 8 is clearly unimodal, with mode localised at the excessive category for the youngest and at the perpendicular category for the oldest of the children, while the drawings of the two intermediate age groups have bimodal distributions with modes between the aforementioned angular categories.

Table 10.9 Frequency distribution of positive biased drawings across angular classes by age and baseline orientation.

	Drawing categories of angular deviations				
	Correct	Incorrect			
Age groups	Accurate 0°- 11°	Compromise 12°- 33°	Context-driven 34°-50°	Excessive >50°	Total
Left slope					
4	0	3	7	9	19
$5^{1/2}$	2	3	14	4	23
7	5	6	7	10	28
8	7	5	6	4	22
Total	14	17	34	27	92
Right slope					
4	0	2	4	10	16
$5^{1/2}$	1	2	8	11	22
7	0	5	12	8	25
8	3	4	10	4	21
Total	4	13	34	33	84

A Wilcoxon test was used to compare the distribution of drawings across the angular categories between the two baseline orientations. The analysis showed that out of the 81 drawings of trees with a positive inclination for both edges of the road, in 34 the trees felt into the same class of angular deviation for both edges, in 17 the trees on the right edge were of a less deviant category to those on the right whereas in 30 the opposite was true. As a consequence, within the drawings where the trees for the two edges were of an incongruous angular class, the trees on the left side were of a more advanced drawing strategy compared to those on the right ($z = 1.94$, $p = .053$, excluding the baseline-free cases it reached significance, $z = 2.15, p<.05$).

Finally, averaging individual angular scores from the two slopes, overall performance on Road task correlated with performance on EFT measures [$r(103) = -.36$, $p<.001$] while the correlation estimates by age level revealed that it was only at the age of 7 where a significant relation between FI and verticality performance was established [$r(28) = -.48$, $p<.01$]. With respect to gender, a significant correlation was obtained for boys [$r(39) = -.52$, $p<.001$] but not for girls whose coefficient fell marginally out of significance [$r(64) = -.24$, $p = .056$]. The difference in the magnitude of these correlations was highly significant ($z = 3.91$, $p<.001$)[16].

4.4 Discussion

The general pattern of results indicates that the Road task poses difficulties for the graphic representation of vertical co-ordinate. The low rate of success in the overall sample and the magnitude of angular distortions suggest that, when children are asked to predict the orientation of trees in the sole presence of explicit oblique alignment cues, they fail to graphically represent them as upright. In general, children are directly influenced by the obliquity of the proximal referents, indicated by the dominant positive bias within inaccurate representations (left slope: 90%, right slope: 88%) where perpendicular and excessive orientations are clearly the modal responses.

The change in performance with age was rather smooth. Among the drawings with a positive inclination, children of all age levels erred towards the perpendicular while excessive responses were of similar frequency up to 7 years of age. It was only at the age of 8 where children demonstrated less vulnerability to the obliquity of the proximal cues by significantly reducing the magnitude of angular distortions and the incidence of out-of-sector errors. Adherence to the proximal alignment referents was demonstrated even by negative biased

[16] The correlation coefficients for each slope separately yielded exactly the same results with the analysis of average scores. A slight difference in the magnitude of coefficients was obtained for the

drawings although the developmental shift was more hasty and abrupt. The analysis of those drawings revealed that contextual forces render the trees of the younger children nearly parallel to the edge of the road whereas, from the age of 7, children's drawings were clearly governed by the vertical factor. In summary, these results further support children's reliance on local spatial cues, which, at early stages of development, is so pervasive that allows them to align the trees not only perpendicular but even parallel to the proximal referents; a solution which is rejected by older children as obviously evasive.

In the present study the performance of children at different ages confirms the Piagetian theory in many respects. First, the fact that the developmental change took the form of a smooth decrease of angular deviations rather than of a drastic shift in the rate of success, is consistent with the Piagetian claim that, only by the age of 9, children can demonstrate the development of an internalised and thus conceptual spatial reference system (*Stage IIIB*). The finding that, in general terms, only 8 year olds performed distinctively different to their younger counterparts is also compatible with the Piagetian theory, since some moderate awareness of the spatial concepts is expected at this age (*Stage IIIA)*. However, although children favoured the perpendicular solution, demonstrating their contextual disposition, the overall incidence of excessive errors (positive and negative) suggests that the relationship of parallelism is of comparable privileged status. We can thus reasonably argue that the Piagetian account of the development of the concept of verticality is generally supported when the assessment procedures resemble the ones in the traditional Piagetian tasks (predictive-mnemonic tasks).

In the present study we did not obtain a significant variation in performance as a result of children's gender. Although, numerous investigators have reported gender differences in

---

sample of girls whose both estimates for left and right slope fell well out of significant level (left slope:

verticality performance in favour of boys (Liben 1978; Liben & Golbeck, 1980), these differences are less consistent in early developmental periods. Studies, testing children of a similar age range to the present one, failed to show significant variation in performance as a function of gender (Liben, 1975 & 1981; Perner *et al.*, 1984). Further empirical support to the previous claim comes from Thomas & Lohaus, (1993) who, testing children of a wider age-range (6-17), claimed that the between-sex difference is somewhat uncertain during the early years but well established by late childhood (10 years) and early adolescence.

With respect to the FD/FI variable, although children more susceptible to field effects, as suggested by their cognitive style, are more influenced by the tilt of road's edges, their drawings are not significantly more deviant to those of their FI counterparts. Having obtained a significant correlation between the Road task and the EFT measures, it is fair to say that field effects impose a considerable influence on performance. It is puzzling however the fact that FI children cannot directly apply their skills to outperform FD ones.

A possible explanation can be sought by analysing the task's requirements. The impact of the field effects in the present task is fairly large, considering that the only explicit alignment cues provided are in direct conflict with the vertical co-ordinate and that the child is working in the absence of a pictorial example. It has been previously suggested that the demand characteristics of spatial tasks change the pattern of differential performance between cognitive style groups (DeLisi, 1983) and that the impetus of field effects reaches its maximum in the absence of a rule (Lohaus *et al.*, 1996). Keeping in mind that children of a younger age are not expected to have fully internalised the vertical axis and readily apply the correct drawing strategy for its representation, it is possible that this high demanding task

---

$r = -.19$, $p = .13$, right slope: $r = -.23$, $p = .07$).

masks children's resistance to field effects, resulting in the obtained equivalent performance between FD and FI children.

Children also performed comparably across the two sloping edges of the road, and the orientation of the oblique baselines had no significant effect in the magnitude of angular distortions. This finding is in line with the result from the nursery group in the Chimney task. However, no generalisation can be made on a wider age range, since previous studies with concrete material have not reported results separately for each baseline (Liben, 1975; Liben, 1981; Freeman, 1983; MacKay *et al.*, 1972; Perner *et al.*, 1984; Wilson & Wilson, 1982). However, empirical evidence from abstract tasks proves that the orientation of an oblique baseline (left, right) has no significant effect on children's representation of spatial co-ordinates (Bayraktar, 1985, Exp.Id; Bremner, 1984; Bremner & Taylor, 1982; Ibbotson & Bryant, 1976).

Despite the absence of a baseline-orientation effect, angular distortions were somewhat reduced when working on an oblique baseline with an orientation to the right (left slope) as opposed to one with an orientation to the left (right slope). This superiority was revealed in the analysis of angular scores for all factors at all levels, and it was more pronounced when the requested tree was attached on the road's edge. Bremner's study (1984) is the only one which found a small baseline difference in favour to the left oblique, testing 4 year olds, whereas the general pattern of results from abstract tasks suggests that children around the age of 5 start drawing a non perpendicular target line with better accuracy on right obliques.

There has been some evidence from perceptual recognition tasks that two opposite oblique lines are not of equal difficulty and that the right oblique is recognised faster than the left (Olson & Bialystok, 1983). If faster recognition implies greater accuracy in judging the linear inclination of the baseline, the construction of a target line forming an angle with it will be accordingly affected. However, this suggestion is rather tentative and limited in the sense that it focuses primarily on the operation of a perceptual process in graphic production. The fact is that, for such task, executive constraints, determined both by the baseline and the target line (vertical) –like anchorage and preferred movement in stroke making– operate as well. The differential effect of these graphic variables for each slope can contribute to the variation in verticality performance across the left and the right side of the road.

The directional errors for each edge support the last suggestion which shifts the focus to the orientation of the obtained lines instead. Results show that the left edge of the road reduces the magnitude of linear inclinations with a positive direction while the right edge reduces the magnitude of linear inclinations with a negative direction. As Figure 10.7 reveals, these target lines have a different degree of obliquity but they share a common feature; they are all left obliques (red lines). On the contrary, the obtained lines with a negative inclination for the left edge and a positive for the right edge of the road are all right obliques (black lines).

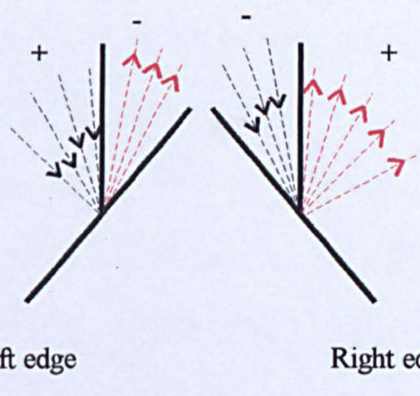


Figure 10.7 Preferred stroke movement of directional errors for each side of the road.

Research into the formal aspects of stroke making suggests that right handers are more likely to work from the left to the right, and to trace left obliques downwards to the right and right obliques upwards to the right as the functional aspect of movement prescribes direction towards the preferred hand (Pemberton, 1990; Van Sommers, 1984). If we now consider the fixed-anchoring principle which prescribes the stroke's start at the point of intersection, the preferred stroke direction is congruent with the anchorage constraint only for the right obliques. The compatibility between these two graphic constraints will make the target lines with a negative slope for the left edge and a positive for the right to grow outward from the baseline, a case which should attenuate responsiveness to field effects.

Considering that the majority of trees had a positive bias for both edges of the road, the right edge of the road is more likely to suppress accuracy if the executive constraints exacerbate the field effects. Support for this suggestion is given by the fact that the superiority of the left edge in reducing the magnitude of absolute angular deviations and promoting the use of a more advanced strategy within positive biased drawings was statistically reliable when baseline-free cases were excluded from both types of analyses.

The last assertion sounds completely speculative, but it may not be too wide of the mark, for it also accommodates some variations in the results obtained only from the left edge of the road. The findings from this slope yielded an interaction between age and cognitive style as well as a less clear developmental pattern in the distribution of strategies for positive biased drawings. As research on executive constraints indicates, when there is a conflict between graphic rules, there is in general greater variability in production strategies. This diversity is also evident cross-sectionally as children of different age give different priority to different rules when incompatibility arises (Goodnow & Levine, 1973; Nihei, 1983; Pemberton, 1985).

Concluding, when children up to the age of 8 are requested to draw from memory a naturally vertical object on slanted surfaces they are contextually bound, failing to employ the external and invariant spatial referents. Instead they produce drawings perpendicular or even parallel to the slanted surface, although the latter is seen mostly up to the age of 7. In the sole presence of explicit cues which compete with the natural co-ordinates, FI children reduce the magnitude of deviations but don't perform significantly different from FD ones.

Having concluded the analysis and the discussion of the results for each task separately, it is evident that, although these tasks assess the representation of the same spatial concept, they elicited different performance. In the following chapter, a comparative analysis will be conducted across the three tasks and the overall performance on verticality assessment will be related with performance on horizontality tasks and with various other ability measures.

CHAPTER XI

SPATIAL AXES: A COMPREHENSIVE LOOK

1. OVERVIEW

So far, we have analysed children's performance on tasks separately, assessing their ability to graphically represent the spatial concepts. The first purpose of the present chapter is to systematically compare children's performance across the verticality studies in order to discern the locus of difficulty, to isolate the contributing factors and to examine how these affect children of different abilities and characteristics.

The second purpose is to study joint performance on the horizontality and verticality studies, focusing mainly on intercorrelational patterns. Another aspect central to Piaget and Inhelder's theoretical argument was that the acquisition of vertical is synchronous with the acquisition of the horizontal and that performance on all these studies tap the same underlying ability to conceptualize space with an invariant co-ordinate system. If performance on the two measures is highly correlated then the long-held treatment of these test materials as operatively equivalent will be justified.

The first analysis was conducted on mean deviation scores and frequency scores of drawing patterns across the vertical tasks. To examine whether an overall pattern of differential performance was present in each subject, individual responses were also analysed. Since there was no significant baseline orientation effect in the Road task, and performance was significantly correlated across the two slanting edges within age and gender, an averaged score was obtained from the left and right slope. This score was subsequently classified into the four drawing categories (accurate to excessive).

2. RESULTS

2.1 Verticality studies

2.1.1 Deviation scores

The mean angular deviation scores were analysed by a 4 (age) × 2 (gender) × 2 (cognitive style) × 3 (task) ANOVA with a repeated measures on the last factor. The total sample consisted of 96 drawings. The occasional equivocal productions of 4 year olds caused incomplete sets of drawings which were discarded from the analysis, reducing their sample size to 16 children. The mean deviations are shown in Figure 11.1 and the results from the analysis are presented in Table 11.1. The between-subjects analysis revealed a main effect of age and cognitive style suggesting that verticality performance improves with age and that FI children were more accurate than FD children. The pairwise comparisons ($p<.05$) showed that the youngest children produced significantly more deviant scores than all of the older age groups while the oldest children produced the least deviant ones compared to all of their younger counterparts, with the mean scores of two intermediate age groups not statistically different from each other. No other test of between-subjects effects yielded significant results.

Table 11.1 ANOVA summary table on mean angular deviation scores.

Effects	*F*	d.f.	*p*
Between-subjects Effects[a]			
Age	24.67	3, 80	<.001
Cognitive style	4.44	1, 80	<.05
Gender	.23	1, 80	n.s.
Within-subjects Effects[b]			
Task	143.89	2, 160	<.001
Task × Age	2.64	6, 160	<.05
Task × Cognitive style[1]	1.75	2, 160	n.s.
Task × Gender	.94	2,160	n.s.

Note a All first and second order interactions were not significant.
Note b All second and third order interactions were not significant.

[1] Since the assumption of sphericity for the within subjects factor was not met [Mauchly's W = .66, $\chi^2(2) = 32.1$, p<.001], multivariate analysis of variance (MANOVA) was also conducted as an alternative procedure which does not depend on assumptions about the covariance matrix. Multivariate tests yielded exactly the same results as the traditional repeated-measures method with the only exception the interaction between task × cognitive style which reached significance ($F_{2,79} = 3.93$, $p<.05$).

The within-subjects analysis showed a main effect of task, and all of the subsequent pairwise comparisons were significant, indicating that all three tasks elicited reliably different levels of performance. However, task differences were not of the same magnitude within each age level, as revealed by the significant interaction between task and age, plotted in Figure 11.1. It is apparent that the Road task seriously hindered children's ability to represent the vertical. It is also clear that the pattern of developmental progression changed differently across the tasks.

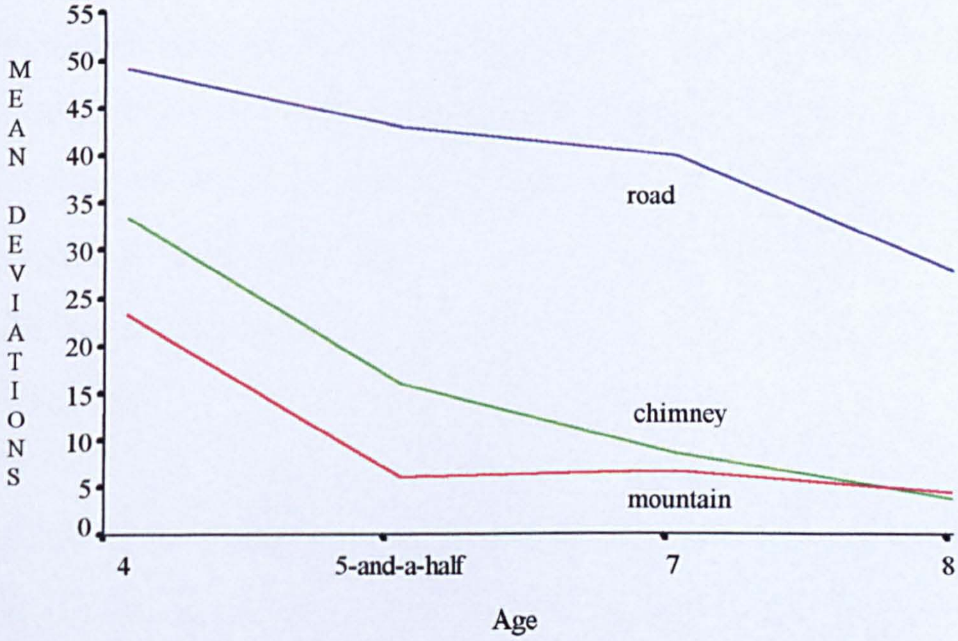


Figure 11.1 Children's performance across the three verticality tasks by age.

With respect to the two copying conditions, task differences are evident between the two younger age groups where the pattern of developmental change is abrupt and of similar magnitude. However, although in the Mountain task performance stabilised after this age, it continued to steadily improve in the Chimney condition, causing between-task differences to disappear from the age of 7. In addition, the fact that for the Road task the developmental change is smooth up to the age of 7 with a marked improvement thereafter, renders performance differences between this task with each of the copying ones to be the largest for the two intermediate age levels and the smallest for the youngest and the oldest of the children. From all of the above, it follows that the overall between-tasks difference is smaller for the two copying tasks and bigger for the Road task compared either of them.

This observation leads us to the analysis of two results which deserve some elaboration. The repeated measure analysis revealed that the first order interactions of task with age, and cognitive style were statistically insignificant while the multivariate analysis yielded an irreconcilable result with respect to the latter interaction. Possibly the very different nature of the Road task, relative to the other two, might have washed out the differential effect of cognitive status and gender on the two copying tasks. This assumption appeared to be well-grounded as, restricting the analysis to the two copying tasks, both of these interactions turned out to be significant (task × cognitive style: $F_{1,86} = 4.05$, $p<.05$, task × gender: $F_{1,86} = 7.71$, $p<.01$).

Considering the cognitive style × task interaction, performance differences between the cognitive style groups were marginal on the Mountain but substantial on the Chimney task, and although both groups were more accurate on the former, FD children showed a greater task differential than FI ones. The fact that the simple effect of cognitive style on the Mountain and the Road task was insignificant indicates that cognitive style differences affect performance only on tasks of intermediate difficulty[2].

The nature of the other significant interaction between gender and task is somewhat different. The direction of between-gender differences was opposite in the Mountain and Chimney tasks as boys had slight smaller deviant scores in the Chimney task but substantially larger ones in the Mountain task compared to girls. As in the previous case, the comparable

[2] It is probably worth saying that the term intermediate is used descriptively and not quantitatively as the differences in the level of difficulty are not equivalent.

performance of the sexes in the Chimney task, along with a clear superiority of the girls in the Mountain task, resulted in larger performance differences within girls and smaller within boys across the two copying conditions.

2.1.2 Frequency scores

Task differences were also investigated comparing the frequency distribution of vertical and non-vertical responses across the three conditions. The Cochran Q test was used to examine whether the matched sets of frequencies of successful and unsuccessful representations in the tasks differ significantly among themselves. As can been seen in Table 11.2, the three verticality tasks produced significantly different rates of success in the whole sample and within each age level. Specific comparisons were then analysed by the McNemar test of change in the drawing response.

In the total sample of drawings, the Mountain task elicited significantly better performance than the Chimney task ($p<.01$), with both of these significantly better than the Road task ($p<.001$). The paired comparisons at each age level yielded the following results. The level of difficulty between the two copying tasks varied across the age levels. The high rate of success among the 8 year old and the low rate of accuracy among the 4 year olds rendered the Mountain and Chimney tasks of equivalent difficulty for the younger and the older of the children in the present sample. On the contrary, a significant change in performance in favour to the Mountain task was obtained for the two intermediate age groups (5½: $p<.001$, 7: $p<.05$). Finally, the two copying tasks were significantly easier than the Mountain task for any age group[3] except for the 4 year olds where only the Mountain task was significantly easier than the Road task ($p<.05$).

Regardless of these age-related differences in the magnitude of between-task difficulty, the overall pattern of results suggests that the Mountain task was the most effective in producing higher rates of vertical orientation output whereas the Road task was the least effective with the Chimney task being of intermediate efficiency. Thus the findings from the analysis of pass/fail scores fully supported the results from the analysis of mean angular scores.

Table 11.2 Number of children drawing accurate orientations (≤11°) in each task by age.

	N	TASKS			Q[1]
		Mountain	Chimney	Road	
Age					
4	30[2]	7 (23%)	2 (7%)	0 (0)%	9.75[b]
5½	25	22 (88%)	10 (40%)	2 (8%)	30.40[c]
7	28	24 (86%)	18 (64%)	3 (11%)	33.43[c]
8	27	25 (93%)	24 (89%)	6 (22%)	34.30[c]
Total	110[b]	78 (71%)	54 (49%)	11 (10%)	100.20[c]

Note 1 *Cochran Q value, b = p<.01, c = p<.001.*
Note 2 *These sample sizes vary compared to those in the analysis of deviation scores. For the purpose of present analysis, the equivocal drawings of 4 year olds were classified as failures to avoid discarding incomplete cases where children succeeded in one task but produced an unrepresentative drawing in the other (e.g. non-linear or inappropriate placed forms). Repeating the analysis on the reduced sample of 16 nursery children who produced representative drawings in all three tasks, we obtained exactly the same results both for the 4 year olds and for the total sample.*

To verify that this order of task difficulty was present in each subject, we analysed individual response patterns on the three tasks. It was found that all but two children could be classified according to four response patterns: failure to produce a vertical orientation in all three tasks(---); success only in the Mountain task (V--); success in the two copying tasks but not in the Road task (VV-) and finally success in all three verticality tasks (VVV). The two children who couldn't fit this classification were one 4 and one 8 year old who succeeded in the Chimney but failed in the Mountain task. Assuming that these two cases are rather

[3] The *p* values for the McNemar tests were as follows: All paired comparisons for the 7 and 8 year olds were significant at $p<.001$. For 5½ year olds the comparison between the Road and Mountain was significant at $p<.001$ and with the Chimney at $p<.01$.

idiosyncratic and not representative of the general trends, they were excluded to obtain a clearer picture of children's overall performance.

Table 11.3 presents the frequency distribution of the four response patterns by age. Kendall's tau_c revealed a very significant correlation between age and ordered pattern of success (tau_c = .56, z = 9.78). The characteristic performance of 4 year olds was that of overall failure (73%) whereas success only in the Mountain task was mostly found among the 5½ year olds. Accuracy in both copying tasks was the modal response of the two older age groups while overall success was mainly obtained by the oldest children.

Table 11.3 Frequency distribution of response patterns according to age.

Response Patterns			Age groups			
			4	5½	7	8
Mountain	Chimney	Road	(n = 29)	(n = 25)	(n = 28)	(n = 26)
-	-	-	22	3	4	1
V	-	-	6	12	6	2
V	V	-	1	8	16	17
V	V	V	0	2	2	6

Note *A "V" denotes a vertical response (0°-11°)*

These findings reveal a clear-cut hierarchy among the tasks. Children who drew the tree vertical on the Road task were able to do the same on the Mountain task as well as to orient a chimney upright on a sloping roof. On the contrary, children who oriented the chimney accurately on a sloping roof were able to accurately orient a tree on a sloping mountainside but not on road's edge. Similarly, children who could draw an upright tree on a mountain could not necessarily do the same when the tree was on a road's edge or a chimney was on a slanting roof. These observations clearly demonstrate that the three tasks involve different processing abilities even though in each case children were requested to draw a vertically oriented form on a non-horizontal baseline.

Next, performance was analysed comparing the distribution of drawing response, classified into accurate, compromise, context-driven and excessive, across the three conditions. Since these drawing solutions reflect a different stage of spatial-processing ability, it was of interest to examine first if these stages are manifested equally across the tasks that vary in terms of performance requirements and second whether the same children performed at similar levels across the tasks.

The tasks elicited significantly different drawing solutions. The Friedman two-way analysis of variance by ranks yielded very reliable differences ($p<.001$)[4] in the overall sample and within each age level, indicating that the Mountain task yielded the more advanced drawing strategies followed by the Chimney task with the least advanced obtained in the Road task. Paired comparisons were then analysed by the Wilcoxon test of signed ranks. Again, significant differences were found within each pair of tasks in the total sample and by age groups[5] except at the age of 8 where the Mountain and the Chimney task did not elicit significantly different drawing responses.

Table 11.4 Pattern of use of drawing strategies across pairs of tasks.

Drawing strategies	Mountain-Chimney (n = 102)[1]	Chimney-Road (n = 100)	Mountain-Road (n = 98)
Congruent	65	21	13
Incongruent			
Adjacent	27 (2)[2]	37 (2)	28 (1)
Discrepant	10 (1)	42 (0)	57 (0)
Z[3]	4.39	7.54	7.94

Note 1 *The sample sizes vary due to rejected cases.*
Note 2 *The numbers in parenthesis denote in how many of the incongruous cases children perform better on the second task compared to the first.*
Note 3 *Wilcoxon matched-pairs signed ranks test, for all the values p>.001.*

[4] All $\chi^2(2)$ values are bigger than 14.00.
[5] The Mountain-Road comparison yielded z values ranging from 3.33 to 4.01, all significant at $p<.001$. The Chimney-Road comparison produced a z value of 2.80 for the youngest of the children significant at $p<.01$, while all other values ranged from 3.99 to 4.51, all significant at $p<.001$. Finally, the Mountain-Chimney comparison yielded z values of 2.17, 3.28, (both $p<.01$) and 2.45 ($p<.05$) for the first three consecutive age levels.

Table 11.4 presents the number of children who produced congruent and incongruent drawing solutions across each pair of tasks. In the first case children performed at exactly the same level making, for example, two compromise or context-driven drawings. Incongruent performance involved cases where the variation was only within adjacent drawing strategies (adjacent: compromise-context driven) and into cases where children alternated between more deviant solutions (discrepant: compromise-excessive).

From the table, it becomes evident that the highest level of consistency in performance was obtained between the Mountain and the Chimney tasks, whereas the lowest level of consistency was observed between the Mountain and the Road task. This further substantiates previous results, according to which, as task differences increase so does the variability of the responses. Collapsing the number of congruous responses with the number of drawings at an adjacent level between any pair of tasks, the percentage of children performing at comparable levels of spatial ability fell from 90% between the two copying task to 58% between the Chimney and the Road, and to 42% between the Mountain and the Road task.

Subsequent analysis by age levels suggested that for the Mountain and the Chimney tasks which produced more stability in performance, all the discrepant responses were produced by the two younger age groups (4 year olds: 8, 5½ year olds: 2). This implies that regardless of the similarities between the two tasks, the lack of fully developed spatial concepts, among the youngest of the children, contributes to larger variability in response as if children are adopting a trial and error approach. On the contrary, as the difference in the demand characteristics of verticality tasks increases, the same degree of response variability was observed across age.

Specifically, for the Chimney and the Road task, the distribution of responses into discrepant and comparable drawing patterns was similar across the age levels [$\chi^2(3) = 3.40, p>.05$]. The same analysis for the Mountain and Road task yielded a significant association between age and pattern of consistency [$\chi^2(3) = 11.7, p<.01$] but the difference was mainly attributed to 8 year olds who produced the least discrepant responses relative to their younger counterparts.

2.1.3 Correlational patterns

Intertask correlation between verticality measures will be considered first. The results from this analysis are presented in Table 11.5. As expected, the magnitude of correlations for the total sample changed between the pairs of tasks. The correlation coefficient was higher for the two copying tasks and lower for the Mountain and Road tasks.

Table 11.5 Intertask correlations[1] with children grouped by age, gender and cognitive style.

	Intertask Correlations					
	Mountain-Chimney		Chimney-Road		Mountain-Road	
Age	N[2]	*r*	N	*r*	N	*r*
4	22	-.11	20	-.08	18	-.05
5½	25	.22	25	.32	25	-.09
7	28	.55[b]	28	.46[a]	28	.49[b]
8	27	.25	27	-.01	27	.02
Gender						
Boys	36	.60[c]	38	.44[b]	36	.28
Girls	66	.52[c]	62	.36[b]	62	.29[a]
Cognitive style						
FD	57	.54[c]	57	.32[a]	54	.33[a]
FI	45	.46[b]	43	.46[b]	44	.19
Total	102	.51[c]	100	.39[c]	98	.27[b]

Note 1 *The correlations are computed on mean angular deviation scores.*
Note 2 *The sample size varies occasionally within each level of the factors due to rejected cases.*
a = p<.05, b = p<.01, c = p<.001

The pattern of correlations by age level is less favourable. It seems that only the performance of 7 year olds correlates significantly across the tasks. The lack of significant correlations within the younger children can be explained on the basis of randomness in their responses due to the lack of a specific spatial rule. What is rather puzzling is the absence of significant correlations among the 8 year olds, who appeared to perform with greater consistency across tasks in the previous analyses. Finally, the correlation coefficients were mainly significant and of similar magnitude for both sexes and cognitive style groups. Only performance on the Mountain and the Road tasks was not reliably related for boys and FI children. However, there were no reliable differences in the magnitude of correlation between the two sexes and between the cognitive style groups (gender: $z = .05$, cognitive style: $z = .72, p>.05$).

2.2 Comparative analysis on axial tasks

Finally, we compared performance on verticality with performance on horizontality measures. Since the verticality assessment yielded three scores and the horizontality two, and since there is empirical support for the distinctively different nature of the Road task compared to the other two verticality tasks, the use of an averaged score from each measure seemed inappropriate. Thus, correlations were computed between all pairs of tasks using the actual deviations scores from each one. Table 11.6 presents the correlations by gender for the total sample which are the most informative and representative of the main pattern. Only the most pertinent of these correlations will be discussed here.

Table 11.6 Correlations between horizontality and verticality measures by age.

	Bowl		Bottle		Mountain		Chimney		Road	
Bowl			*.45^{c}*	*(64)*	.34^{b}	(63)	.42^{c}	(65)	.07	(61)
Bottle	*.47^{b}*	*(37)*			.22	(65)	.38^{b}	(65)	-.04	(63)
Mountain	.37^{a}	(35)	.42^{b}	(36)			*.52^{c}*	*(66)*	*.29^{a}*	*(62)*
Chimney	.47^{b}	(37)	.53^{c}	(38)	*.60^{c}*	*(36)*			*.36^{b}*	*(62)*
Road	.04	(37)	.37^{a}	(39)	*.28*	*(36)*	*.44^{b}*	*(38)*		

Note 1 *The correlations are computed on mean angular deviation scores. The magnitude and the pattern of these results were not changed when performance on verticality and horizontality was computed on pass/fail scores.*

Note 2 *Boys' correlations are below the diagonal; girls' are above the diagonal.*

Note 3 *Cell entries with italics denote not relevant correlation estimates (those within each measure).*

Note 4 *The figures in parenthesis represent the sample size. This varies across pairs of task due to equivocal cases.*

a = p<.05, b = p<.01, c = p<.001

In general, all verticality tasks correlated well with horizontality tasks with the sole exception of the Road task. It was only within boys that a moderate correlation between the Road and the Bottle task was obtained, whereas neither in the total sample nor among girls –in either cognitive group– was there any reliable association between performance on the Road task and the WLTs. So, as long as the main task requirements are similar (all copying tasks) in the assessment of spatial concepts, the Piagetian vertical and horizontal tasks can yield comparable levels of spatial abilities. As Table 11.6 indicates, the correlations between performance on each of the two copying vertical tasks with performance on horizontal tasks were significant and of similar magnitude in both sexes, while the strongest relationships were obtained between the Chimney task and the WLTs.

Considering age, there wasn't any significant correlation between all pairs of tasks for any age level. In a subsequent analysis, the correlations were calculated on two averaged scores, one from each measure, using either angular deviations or angular category. Both procedures yielded results consistent with the previous analysis. The correlations were significant for the overall sample and for each gender, but not for any age level.

The relative difficulty of the horizontal and vertical tasks was assessed by examining the number of children who changed their drawing response across the tasks. First changes of pass and fail scores will be considered and then variations of drawing strategies. In the total sample, when the vertical concept was assessed by the Mountain task, 92% of children succeeded in this but failed in the WLTs. This pattern of change in accuracy was significant[6] (McNemar's test for change) for all age groups except for the 4 year olds, for whom it was found only in the Bowl task. Similarly, 79% and 75% passed the Chimney task but failed the Bowl and Bottle task respectively. The shift in response was statistically reliable only among the 8 year olds (McNemar's test for change, Bowl-Chimney: $p<.01$, Bottle-Chimney: $p<.001$). On the contrary, when the vertical concept was assessed using the Road task, 78% and 89% of all children produced accurate WL orientations in the Bowl and the Bottle respectively, but failed to draw the trees vertical to the road's edge. This shift in response in favour of the WLTs relative to the Road task was prominent among the 7 and 8 year olds (McNemar's test showed that all the comparisons were significant at $p<.05$, except the Bottle-Road one for 7 year olds which was significant at $p<.01$).

Next, we consider if children used drawing solutions which reflect comparable levels of spatial development across the horizontal and vertical measures. Similar to the analysis of strategies between vertical tasks, we classified the pattern of responses into *congruent* – when children used the same strategy across the vertical and horizontal tasks – *adjacent*, – when the variation was only between contiguous deviant classes – and *discrepant* – when they shifted between very different solutions[7]. The results from this analysis are reported in Table 11.7 in percentages for greater legibility, since occasional equivocal responses caused fluctuations in sample sizes across pairs of tasks.

Table 11.7 Pattern of use of drawing strategies between vertical tasks and WLTs.

Vertical →	Mountain		Chimney		Road	
Horizontal →	Bowl	Bottle	Bowl	Bottle	Bowl	Bottle
Patterns						
Congruent	32	35	37	39	25	28
Adjacent	39	34	50	61	33	35
Discrepant	29	31	13	20	42	37

[6] For the 8 year olds both comparisons were significant at $p<.01$ and for the 5½ at $p<.001$. For the 7 year olds, the Bowl-Mountain comparison was significant at $p<.001$ and the Bottle-Mountain at $p<.05$, same as for 4 year olds with respect to the latter.

[7] A *congruent* pattern of response across any pair of vertical and horizontal tasks was one where children produced two drawings falling in the same angular category (i.e. compromise: 12°-33°). A response pattern was *adjacent* when they produced for example an accurate response in one task (0°-11°) and a compromise response (12°-33°) in the other task, whereas it was *discrepant* when the two responses felt in two very deviant angular categories (i.e. one was accurate, 0°-11°, and the other was context-driven, 34°-50°).

The highest percentage of congruent drawing responses was attained by the Chimney and WLTs (37% & 39%) and the lowest by the Road and WLTs (25% & 28%). There was also a remarkable difference in the percentage of discrepant responses between these two vertical tasks with the Bowl task in particular (Chimney/Bowl: 13%; Road/Bowl 42%). The analysis of verticality reveals that the Chimney task is of moderate difficulty compared to the other two. So, it is not the concept of vertical or horizontal invariance *per se* but the nature of the tasks which causes fluctuations in performance on verticality and horizontality assessment. These findings are in agreement with the correlational patterns obtained, where the strongest relationship was found between the Chimney and WLTs.

3. GENERAL DISCUSSION

The systematic comparative analysis of the studies on verticality supported the differential performance across the tasks on many different grounds (mean deviation scores, rate of success, use of drawing strategies). Although in all tasks children were required to make a manual response, representing a vertically oriented form on a non-horizontal baseline, children manifested the conceptual grasp of vertical co-ordinate more readily when copying from a model, and when the immediate contextual cues of the perceptual field competed less with the natural co-ordinates. When children were requested to draw trees along the edges of the road from memory, they oriented them at right angles or even parallel to the oblique baselines. When a pictorial example was provided, in the presence of a context of equivalent obliquity (roof), a considerably larger number of the same children produced vertical orientations and the proportion of perpendicular solutions was substantially reduced. In the presence of a less distracting contextual frame (Mountain), children did even better; from the age of 5½ they performed at ceiling level, and perpendicular solutions were mainly non-existent.

Many authors who have found fluctuations in performance as a result of manipulation of task requirements (Abravanel & Gingold, 1977; Ford, 1970; Liben, 1978; MacKay *et al.*, 1972; Perner *et al.*, 1984) saw such results as a challenge to Piagetian theory and initiated further research to explain these inconsistencies. A number of authors began to make a theoretical distinction between operational competence and utilisation of that competence (DeLisi, 1983; Liben & Golbeck, 1980, 1984). They claimed that in some cases failure on the Piagetian tasks reflects an inability to activate or use underlying competence (performance or competence-utilisation hypothesis), rather than the absence of that competence altogether (competence or competence-deficit hypothesis). Similar views were expressed by authors who, focusing on performance factors and methodological issues, suggested that children are not always insensitive to the co-ordinate axes, but certain conditions prevent them from making full use of this knowledge (Freeman, 1980; Pascual-Leone & Morra, 1991). In general terms, the obtained differences in performance from the present studies offer empirical support for the competence utilisation hypothesis and suggest that the decision of whether a Euclidean spatial concept is available or not will depend to some extent on the form and content (nature) of the instrument used to assess it.

A proper analysis of the characteristics of each task will follow to elucidate upon its underlying processes and demonstrate how these affect children of different abilities and attributes. The findings clearly show that the Road task is the most difficult in eliciting good performance. The pattern of developmental change is smooth and it does not involved a radical progression in the use of drawing strategies but rather a reduction in the magnitude of angular deviations, changing the incidence of excessive and context-driven responses. Analysing its task requirements, two aspects are involved: children have to construct an internal representation of the required form and second they have to invent a graphic equivalent, using their depiction skills to transpose this internal model onto paper.

The nature of children's internal model is a function of their perceptual experience and past representational attempts which in turn depend on the level of operatory intelligence (Luquet, 1927/1977; Piaget & Inhelder, 1966/1971). Trees growing upright from the ground in children's environment have 'ecological frequency'. Similarly, ground-lines drawn parallel to the bottom of the page and forms placed at right angle to them have representational prevalence in children's spontaneous productions. Children who are preoccupied with figurative aspects and who have not mastered the operations needed to construct an invariant spatial reference system and to transform the canonical image of objects or spatial relations, will reproduce this static image of their internal model.

Developmental studies of drawing have demonstrated that children, when drawing a familiar scene from imagination (that is, on the basis of available conceptual knowledge) will focus on intrinsic properties and produce 'object-centred' or intellectually rather than visually realistic drawings. Indeed, up to the age of 9, they tend to use depiction devices of a conventional nature to communicate what they know about the objects or spatial relations between them. So the greater the familiarity of the scene to be portrayed, the greater the mental energy required to transcend the impact of its canonical properties and reproduce its view-specific ones. Pascual-Leone's (Pascual-Leone & Goodman, 1979; Pascual-Leone & Morra, 1991) analysis of the processes involved in the spatial Piagetian tasks contains a claim that the over-learned images and schemes of the discrete stimuli hinder representation of the otherwise acquired spatial concepts. Similar views are held by other authors who point out that success on those tasks demands that children overcome the normal orientation of stimuli which often displays their structure more clearly (Arnheim, 1969, p.263; Freeman, 1980, p.156).

A number of studies have shown that children demonstrate different ability to 'inhibit' their knowledge about the structure of the scene. Crook (1985) points out that children make very definite perpendicular drawings in their spontaneous drawings whereas in other contexts, specifically designed to induce view-specific representations, their errors only tend towards the perpendicular. This claim gains support from studies which show that children are more successful in representing these spatial concepts when tested with abstract and non-physical rather than meaningful and physical material (DeLisi 1983, DeLisi *et al.*, 1995; Liben & Golbeck, 1980).

Another point of departure between the Road task and all the other tasks lies in the pictorial significance of the oblique baseline. In the latter cases, children had to copy from a line drawing, sometimes described as an isographic task (Cox, 1992). Freeman (1983) and Mitchelmore (1985) more accurately call it a 2D homograph. In the two vertical copying tasks, the 2D model represents a 3D scene where the inclination of the baseline on the picture-plane (mountainside, sloping roof) is directly mapped into inclination in the visual field. On the contrary, the inclination of the road's edges has a more complex pictorial significance. As Freeman explained

> *the oblique edges on the picture-plane now represent a horizontal contour in the scenic azimuth, the verticality of the tree remains as in the 2D task, and the acute angle between vertical and oblique on the picture-plane now represents a perpendicular in the scene (1983, p. 128-129).*

Michelmore agreed with Freeman calling such a task a 3D homograph which, although it is structurally equivalent to the Chimney task, places higher demands on representational awareness and on the mental work needed in symbolic processing. A number of studies have reported that children's ability to decode the semantic content of edges on the picture surface to treat it as a projection plane, and their awareness of drawing devices for representing 3D properties on a 2D medium, are limited below nine years of age (Golomb & Farmer, 1983; Reith, 1988, 1990).

It is therefore expected that all children in the present age range will have difficulty with the Road task. The fact that a number of children attached the trees on both the outer and the inner side of each of the road edges suggests that they have difficulty comprehending the representational significance of these contours. Children's difficulty with 3D tasks has been reported by Mitchelmore (1975) who, asking children to draw telegraph poles alongside a receding road, found that up to the age of 11 years, children made perpendicular errors.

Further, besides the symbolic complexity of the road edges, their obliquity (in the absence of any other explicit alignment sign) presents a cue-conflict situation. Elliott and Connolly (1978) reported that accuracy in judging the orientation of a tree dropped significantly from 98% in the total absence of a baseline to 67% in the presence of an oblique one. The deleterious effect of obliquity in the picture plane was also demonstrated in a production version of the previous task where Freeman and Betts (in Freeman, 1983) found significantly better performance without a slanting baseline representing road's edge.

Summarising the underlying processes of the Road task, to demonstrate an ability to represent the vertical co-ordinate, children have to inhibit their knowledge about the canonical structure of the scene, they have to be able to decode the pictorial significance of the provided context and then they have to contrive and apply the appropriate graphic device overriding the context's misleadingness and production biases. While the younger of the children are more likely to lack all these skills, the superiority of the older, in any of those abilities, is masked by the cognitive overload this task imposes, resulting in the rather flat performance across children of different age, gender and cognitive style.

By contrast, the pictorial example in the two copying tasks neither necessitates a mental representation of a scene per se, nor does it require children to devise a depiction strategy for its graphic translation. Thus, conceptual and production difficulties can be overcomed if children attend to the model and confront their notions about the required spatial relation or/and 'borrow' the graphic means to depict it. Children in general performed exceptionally well in the two vertical tasks, even from the age of 5½. These results challenge the findings from a number of studies which report that the provision of a model, or the observation of the physical phenomena in question, have relatively little effect on performance. To account for this inconsistency and to reconcile these findings with Piagetian formulations, it is most likely that the effect of a model in earlier studies was undervalued because the model was not a line-drawing (2D) but rather a 3D stimulus (adults: Liben 1978, Thomas *et al.*, 1973). As it has been previously maintained, not all children possess the graphic repertoire needed to translate its visual image onto the picture plane.

However, the dexterity with which children employed the drawing device at hand varied across the copying tasks on the basis of their contextual misleadingness. Thus, although both facilitated the discovery of the accurate depiction strategy, only the less distracting structural properties of the mountain's contour alleviated the effect of perpendicular bias and fostered the application of the correct graphic device. The scarcity of contextual errors in the Mountain task, along with the fact that it did not correlate with performance on the EFTs, supports the belief that antagonistic field effects can mask the use of the accurate strategy. The more compelling evidence, in line with a competence-utilisation hypothesis, is provided by the significant interaction between the copying tasks with cognitive style status. As this interaction suggests, the performance of FD children is disproportionately depressed compared to FI children in the Chimney task, since the obliquity of the roof's edge hinders the former children from gaining full benefit from the pictorial example.

Children of different age levels are also differentially aided by the provision of a pictorial example. It seems that there is no change in performance when the spatial concept is either non-existent or adequately developed, whereas at intermediate stages of development within-age fluctuations are observed. Given that the information about the appropriate spatial referent is embedded in the pictorial example, the high degree of consistency in the responses of 8 year olds, both in terms of success rate and degree of accuracy, suggests that these children possess the requisite conceptual structure to assimilate this information. On the contrary, the absence of a fundamental difference in the performance of 4 year olds – for whom the lack of competing alignment cues only reduces the magnitude of errors – suggests that they fail to notice the crucial information simply because they lack the conceptual system. Thus, consistent with our discussion of the horizontality studies, a competence-deficit interpretation is more appropriate in explaining performance of very young children whose conceptual and performance difficulties are widespread and resistant to interventions.

Children at transitional stages of development are more susceptible to task variation[8]. Both age groups appear to have sufficient flexible or versatile conceptual schemes to accommodate the provided spatial information and perform comparably with the oldest children in the Mountain task. However, they seem to have limited ability to overcome superficial difficulties in performance factors, since in the presence of a more challenging context (roof-Chimney), their ability to use the appropriate spatial referent fell short.

With respect to the gender variable, it has been suggested that performance factors are more likely to suppress girls' skills since no reliable gender differences are found in tasks with reduced performance demands (Liben & Golbeck, 1980; McGuillicuddy-DeLisi *et al.*, 1978). We were surprised to find gender differences in the most facilitating testing condition in favour of girls which resulted in a significant interaction between gender and copying tasks. The fact that girls seem to benefit most from experimental conditions that facilitated the application of spatial concepts is in agreement with the previous contention. However, the modest superiority of boys in the Chimney and Road task, albeit in the predicted direction, is much less than previously reported. The absence of reliable gender differences in verticality studies is in perfect parity with the results obtained from those on horizontality. These findings corroborate previous studies which, using various methodological approaches (McGillicuddy-DeLisi *et al.*, 1978; Thomas & Jamison, 1975; Thomas & Lohaus, 1993; Thomas & Turner, 1991), maintain that the pattern of gender differences is less clear in the early grades, while around the age of 10 significant differences emerge.

The same chronological age was reported as a threshold above which gender differences in spatial tasks can be more confidently attributed to cognitive style status. Many researchers have claimed that gender differences in cognitive style are inconsistent or absent prior to adolescence (Fairweather, 1976; Witkin & Price-Williams, 1974; Wolf, 1971). In the present study no gender differences were found in mean EFTs scores, and although performance on verticality tasks (Chimney and Road) and EFT was significant correlated for boys but not for girls, the contrast was statistically significant only in the Road task. Early research from Pascual-Leone (1969) led him to predict that

> *...only small cognitive (field dependence/independence) differences in performance will be found before the age of 9 or 10 years. Furthermore, larger cognitive style differences should be found in 11-or 12-year-olds and perhaps even in adults as compared with young children. (p. 300)*

In a more recent paper (Pascual-Leone & Morra, 1991) he explains this hypothesis on the basis of children's mental processing abilities. According to his model, children in Piagetian

[8] The performance of 5½ year olds improved significantly in the Mountain task on both estimates of

tasks are confronted with a cognitive conflict in which the incorrect strategy, prompted by misleading factors, competes with the correct strategy which is boosted by mental attentional capacity. His basic argument is that mental attention (M capacity)[9] is still limited below the aforementioned age threshold, obliterating cognitive style differences. Similar views are shared by Overton and Newman (1982), who consider FI more as a utilisation variable. The fact that, in the more cognitively challenging task (Road), the performance of the cognitive style groups is equivalent, along with the finding that performance on the verticality tasks (Chimney and Road) correlates with cognitive style measures in the overall sample but not within each age level corroborates this premise. These suggestions are also in line with the discussion of WLTs. The general results from the assessment of vertical and horizontal concepts show that FI is instrumental to good performance on Piagetian spatial tasks but that its effect is negligible in children who lack the concepts or in cognitively challenging testing conditions.

As has been already mentioned when discussing WLTs, the comparison of performance on vertical tasks and EFTs can be also interpreted in the light of the tests' attributes. Results from the total sample of drawings show that when the degree of misleadingness of the immediate context is not reduced, disembedding ability is one component of performance on verticality tasks. However, although factor-analytic research provides evidence that the EFTs are structurally similar to the Piagetian task (Goodenough *et al.*, 1987; Pascual-Leone, 1969; Witkin, 1978; Witkin & Goodenough, 1981), many authors have explained the differences on the basis that the conflict is mainly figurative in the former but figurative and operative (motor) in the latter (Liben, 1978; Pasual-Leone 1989). This suggests that, in verticality tasks, field effects do not only hinder the child from seeing the correct solution in the visual field but also operate during execution.

---

performance, success rate and magnitude of errors, while that of 7 year olds changed only on the first measure.

Empirical support for the differential processes involved in EFTs and Piagetian tests has been provided by Pennings (1991) who found that children of 7 and 8 year olds adopt different strategies when solving these tasks. With respect to the vertical tasks, it is also important to note that the target form is not embedded in a misleading context but rather attached; an aspect which could possibly have reduced the magnitude of correlations with EFTs. In summary the results from the studies are in line with Pascual-Leone's claim that in Piagetian tasks

> *"field dependence /independence is an important factor, but it leaves unexplained a large proportion of variance" (Pascual-Leone & Morra 1991, p. 246).*

In view of the structural characteristics of the tasks used in the present studies two hypotheses can be proposed to explain the previous assertion. Either disembedding skills cannot be operationalised due to the limited mental capacity of children of the present age-range or FD/FI – as measured by the EFTs[10] – and the ability to graphically represent the spatial co-ordinates are not, as DeLisi (1983) puts it, "isomorphic constructs".

Finally, performance on horizontality tasks correlates well with performance on verticality tasks, and the relative difficulty of these tasks is related to the task variables and not to a developmental lag in the understanding of these concepts, as has been previously claimed. The finding that comparable levels of spatial ability are demonstrated in the copying tasks on each measure supports the claim that a general, conceptual process underlies performance on these various Euclidean tasks. The absence of a gender difference in the magnitude of these correlations is not totally in agreement with other studies (Liben, 1978; Liben & Golbeck, 1980). This inconsistency can be attributed to differences with age ranges studied.

---

[9] Mental power necessary to simultaneously attend to all the task-relevant information.

[10] Similar reservations have been reported with respect to the (portable) RFT, although on different grounds. For a specific discussion of its limitations see Abravanel & Gingold, 1977; Liben 1978; Thomas & Jamison, 1975.

However, performance across the vertical and horizontal tasks is by no means uniform. First inter-task correlations weren't significant within age levels, and there were significant different rates of accuracy across the tasks. Although Piaget treated the present test materials as conceptually and structurally equivalent, subsequent research discounted this view. Many authors (Freeman, 1980; Liben, 1978; Perner *et al.*, 1984) have pointed out that the WLT is a special type of relative-movement problem. In contrast with the vertical tasks in the present study, WLTs involve a mobile reference system where the water level remains unaffected by the container's spatial displacement. Thus, the superiority of the Mountain and Chimney tasks relative to the WLTs might be attributed to this issue of invariance. In fact, when the concept of verticality is assessed with tasks involving implicit motion (PL suspended from inclined trailers) performance is considerably worse compared to that in tasks using immobile reference systems (roofs, mountains) (Liben, 1978; MacKay, *et al.*, 1972). Further, in WLTs the requisite line is fully embedded in the proximal frame, whereas in the vertical tasks it is just attached. Interestingly enough, the WLTs are less difficult than the Road task. Possibly the context in which the WL problem is presented (line-drawing with the water level present) reduces the impact of mobility and renders it easier than a task which, while it doesn't involve a mobile reference frame, presented difficulties on all other aspects.

4. CONCLUSION

In conclusion, the present series of experiments explore the methodological difficulties that stand in the way of devising an instrument that will assess spatial concepts without underestimating children's abilities. Individual differences in performance on Piagetian tasks have led some to believe that the developmental stages in the representation of spatial co-ordinates "...are no more real than the psychologist's need for an economic taxonomy of behaviour" (MacKay *et al.*, 1972, p.232). However, the age-related hierarchy of individual response patterns across the tasks suggests that each stage of development correspond to a range of abilities which prescribes lawful performance. Certain inconsistencies appear when differential task demands alter the extent to which these abilities will be employed.

Considerable variability in the age of attainment of particular operations has been noted by Piaget and Inhelder (1948/1967, p. 410-411). Discussing the issue of horizontal decalage, they maintain that some conditions produce more resistance than others, but that these are hard to predict. The present studies shed light to the processes involved in each task and elucidated the difficulties encountered at each stage of development. The Road task, which demands image generation, representational awareness and FI is beyond the abilities even of 8 year old children. In terms of Piaget's stages, they still haven't reached *Stage IIIb* which requires a fully internalised reference system. But as Piaget has pointed out "...it is doubtful whether failure to *predict* horizontality at this age is by itself proof of inability to conceive of a co-ordinate system" (Piaget & Inhelder, 1948/1967, p. 360). So, the same children, provided with a pictorial example, can *appreciate*[11] the relevance of a system for overcoming representational (conceptual and/or production) constraints easily.

As the demand characteristics reduce the impact of these constraints (over-learned schema, contextual misleadingness, depiction skills), they maximise the likelihood that competencies, if present, will be activated and resistances will be relaxed. As the findings show, conceptual schemes about the canonical orientation of forms exert the least resistance when such factors are substantially eliminated (Mountain task). To the extent that some subjects continue to fail in the face of task modifications of this kind, competence explanations of failure become increasingly more credible. This applies readily to the performance of 4 year olds.

[11] Piaget and Inhelder claim that it a different issue altogether whether children have the conceptual prerequisites for assimilating information about spatial invariants compared to whether children have already acquired knowledge about these particular invariants. If the scope of an experimental task is the former, then the provision of a model is appropriate. Drawing from memory is like asking children to recite the physical principle in question or give operatory explanations of their correct drawings. Empirical findings indicate that not even adults can do this.

CHAPTER XII

AN INCLUSIVE PICTURE OF DRAWING DEVELOPMENT

1. OVERVIEW

So far, a number of issues pertinent to drawing development have been investigated through this thesis by means of various drawing themes and from different theoretical perspectives. Chapters II and III were concerned with the pictorial representation of one of the most popular drawing topics among children, the Human Figure. The theoretical approach adopted was based on the long tradition of using Human Figure Drawings (HFDs) as a measure of intellectual maturity. The chapters investigated various properties of one of the most widely used drawing tests of intellectual ability, namely the Goodenough-Harris Drawing Test (GHDT: Harris, 1963). The test reveals reliable developmental aspects, yet its validity as an index of general intellectual ability is mediocre. Because of this low validity, the traditional notion that children's drawings can be viewed as a direct translation of the content of their mental concepts started to be disputed, and consequently the drawing tests, developed from this framework, began to be distrusted.

The resurgent interest in the psychological study of pictorial representation in the last decades appears to be the result of new theoretical approaches and methodological techniques. Researchers started to appreciate more the demands of the drawing medium and the various abilities involved in the acquisition of drawing competence. The realization that children can develop quite sophisticated knowledge about the properties of an object or scene, yet their graphic renditions may not be as articulated and differentiated led to the distinction between conceptual maturity and representational logic. In due course, new methodological designs were developed aiming to unfold the variables that affect representation. A large body of research is now devoted to the study of drawing strategies and patterns of errors (e.g. Freeman, 1980; Van Sommers, 1984; Willats 1977, 1981, 1995) by means of detailed experimental analysis and manipulation of a number of neglected variables such as the nature of instructions, task demands and characteristics. Such an approach was adopted in the remaining chapters of the thesis, focusing on two major graphic milestones; the realistic depiction of spatial relationships involving occlusion (Chapters IV-VII) and those involving the use of the horizontal and vertical coordinate system (Chapters VIII-XI).

Studying children's drawings within the experimental framework has obvious merits. Particular developmental phenomena –mainly spatial in nature – are broken down into their constituent parts to allow the investigation of specific hypotheses. Still, this pursuit is inevitably domain-specific, as different aspects of drawing competence are examined independently of each other. In the present chapter an attempt will be made to bring together the areas of pictorial representation considered in this thesis, and investigate the various aspects of developmental change within a broader context.

Further, the rigorous experimentation of recent years has produced rich, yet diverse empirical and theoretical material. Researchers, attempting to explain *why* a particular topic is represented in a certain way, have placed different emphases on conceptual, perceptual and organizational processes, depending on the theoretical perspective they endorse. The findings from the studies on occlusion and spatial axes suggest that drawing is not exclusively a problem of knowing, seeing or 'formatting'. The way children of different ages store information for the purpose of pictorial representation seems to be determined by conceptual, visual and procedural knowledge, all of which make some of these internal descriptions more useful (successful) than others for the purpose of picture-making. The study of child art away from the framework of intellectual maturity has overlooked the role of mental abilities in the construction and use of graphic schemes. Hitherto, recent drawing research has been concerned in establishing general developmental trajectories within various domains of graphic competence, yet there is a dearth of research investigating the separate contribution of intelligence along with age.

The aim of the present chapter is first to provide an inclusive account of children's performance on the spatial tasks, administered in this thesis. For this purpose, drawing scales will be devised to assess children's overall ability in representing the particular aspects of pictorial space studied. After examining the developmental properties of these scales, we will investigate how children's performance on the drawing measures of spatial ability relates to their performance on GHDT and Wechsler scales. Such an analysis aims to account for two obvious omissions in drawing research. First, the relationship between drawing ability with performance on standard instruments of intellectual functioning has been systematically investigated only with HFDs. The administration of a satisfactory number of drawing tasks of various complexity in this thesis seems to provide the opportunity to study this relationship in other domains of graphic competence. Second, to the knowledge of the author, no-one has examined how children's performance on spatial tasks compares with estimates of ability as indicated from a non-verbal instrument such as the GHDT. The study of these two neglected aspects has significant advantages. Having examined the relationship between GH and IQ scores in Chapter III, we can compare the degree of involvement of intelligence in GHDT with that in the drawing tasks. The results of this comparison would have clear implications for determining the optimum method of assessment of children's general abilities by means of drawing tasks.

2. METHOD

The data obtained from the same sample of children assessed with various instruments and tasks though this thesis was used for the purpose of the present analyses. One hundred and fifteen children (115) have been tested with the GHDT and Wechsler scales (WISC-III^{UK} and WPPSI) in Chapter III. Each child yielded three GH scores – one for each scale (Man, Woman, Self) – and three IQ scores – Full Scale IQ (FSIQ), Verbal IQ (VIQ) and Performance IQ (PIQ). These scores were used as criterion measures against which performance on the spatial tasks was compared.

Of the 115 children who provided GH and IQ scores, 110 took part in the series of experimental tasks. Drawing competence in these tasks was measured on two dimensions, using a point scale and a structural scale. The point scale credited the use of the correct drawing device with one point for each task that the child has passed. The ability to use the Hidden line elimination (HLE) device for occlusion scenes was assessed with 9 drawing items (Chapters V & VI), whereas the use of spatial axes was assessed with 5 tasks (Chapters IX & X). Thus, the sum of drawing items passed yielded an occlusion point score (OPS), ranging from 0 to 9, an axial point score (APS), ranging from 0 to 5, and a composite point score (CPS), ranging from 0 to 14.

The structural scale was qualitative, designed to assess how deviant were children's strategies from the normal view-specific strategy. Three general approaches were identified in the drawings of partially occluded scenes; the *HLE device* which was credited with 3 points, the *unification* strategy earning 2 points and the *separation* strategy yielding 1 point. Similarly, deviant strategies in the axial tasks were classified into four angular categories, ranging from *accurate* to *compromise*, *context-driven* and *excessive*, and were credited with a score of 4, 3, 2, and 1 respectfully[1]. Therefore, a child using always the appropriate strategy could obtain a maximum structural score of 27 (3×9) for the occlusion tasks (OSS), and a maximum score of 20 (4×5) for the axial tasks (ASS).

3. RESULTS

3.1 Scales reliability

Two preliminary analyses were conducted before investigating specific relationships between various measures of ability. To examine the properties of the point scales (CPS, OPS, APS) for the assessment of children's drawing skills, estimates of internal consistency were computed on the drawing items comprising these scales, and inter-task correlation coefficients were obtained, separately for occlusion and axial tasks. Although, the relationship between axial tasks has been already investigated using angular deviation scores (see Chapter XI, Table 11.6), the present correlational analysis was performed on pass scores to parallel the equivalent analysis on occlusion tasks. The measures of internal consistency regarding content sampling were satisfactory. The reliability coefficients were .86 for all experimental tasks, .82 for occlusion tasks and .66 for axial tasks using the KR-20 formula[2]. The low estimate for the axial tasks was expected as, according to Kline (1993, p.11), 10 items is mainly the minimum requirement for the assessment of internal consistency of a scale. Considering that reliability relates to the test length, and that coefficients should be around .90, it is fair to say that the CPS – based on 14 drawing items– and the OPS – based on 9 – reliably reflect abilities which are relatively homogeneous.

[1] For a detailed description of the scoring criteria see Chapter V, section 2.4 and Chapter IX, section 1.5.
[2] Comparable results were obtained when internal consistency reliability was investigated with the split-half method, using the Spearman-Brown formula. The coefficients were . 85 for the CPS, .79 for the OPS and .61 for the APS.

Spearman inter-task correlation coefficients are reported in Table 12.1 for the total sample. In general, all drawing tasks were significantly related, and in only 5 out of the 36 comparisons between occlusion tasks did they fail to attain significance. Likewise, from the 10 correlational estimates obtained from the axial tasks only one was not statistically significant. Each drawing task yielded a variable range of coefficient estimates with those measuring presumably the same graphic ability. This was expected since these correlations were computed on the total sample of drawings. Following the empirical results from the previous chapters, domain-specific tasks (e.g. occlusion) reveal a different degree of difficulty and age-related sensitivity.

Table 12.1 Inter-task correlation matrices within two domains of drawing competence.

	OCCLUSION TASKS									
	2	3	4	5	6	7	8	9	Range of *rs*	Median *r*
1. Apples	.30^{c}	.38^{c}	.53^{c}	.47^{c}	.34^{c}	.42^{c}	.29^{b}	.26^{b}	.26 – .53	.36
2. Bottles		.31^{c}	.35^{c}	.27^{b}	.34^{c}	.49^{c}	-.03	.31^{b}	-.03 – .49	.31
3. Sunset			.58^{c}	.56^{c}	.26^{b}	.46^{c}	.22^{a}	.21^{a}	.21 – .58	.34.5
4. Balls MO				.63^{c}	.32^{c}	.55^{c}	.31^{c}	.17	.17 – .63	.44
5. Balls FO					.26^{b}	.46^{c}	.40^{c}	.21^{a}	.21 – .63	.43
6. Bus						.35^{c}	.13	.33^{c}	.13 – .35	.32.5
7. Houses							.16	.28^{b}	.16 – .55	.44
8. Vase								.11	-.03 – .40	.19
9. Boat									.11 – .33	.23.5
	AXIAL TASKS									
	2	3	4	5					Range of *rs*	Median *r*
1. Bowl	.28^{b}	.23^{a}	.33^{c}	-.01					-.01 – .33	.25.5
2. Bottle		.25^{b}	.27^{b}	.31^{c}					.25 – .31	.27.5
3. Mountain			.55^{c}	.21^{a}					.21 – .55	.24
4. Chimney				.34^{c}					.27 – .55	.33.5
5. Road									-.01 – .34	.26

Note N = 110, $^{a}p<.05$, $^{b}p<.01$, $^{c}p<.001$

Considering occlusion scenes, the Boat and Bus tasks yielded the narrowest range of inter-task correlations, since they were relatively the most difficult to pass. Although the correlation values for the Sunset task fell within a wide range, yet it yielded the lowest median coefficient (18.5), as it was notably the easiest of all. Generally, the Houses, Balls and Apples tasks obtained a satisfactory range of inter-task correlations and an adequate median coefficient estimate (mid 30s to mid 40s). The same variable pattern was obtained from the axial tasks, although here the values were somewhat lower and the range of correlations more restricted. In general, it can be said that the Chimney (vertical tasks) and the Bottle (horizontal tasks) tasks, correlate most satisfactorily with the other tasks of spatial coordinates.

Finally, to investigate the comparability between the various point scores derived from these scales, Pearson product moment correlation coefficients were obtained between OPS, APS and CPS. OPS and APS correlated with each other at .67, and with CPS at .96 and .86 respectively (all $p<.001$)[3]. These results suggest that, from the total number of drawing tasks, we can obtain a rich picture of children's graphic competence in handling spatial concepts. Although there is a reasonable overlap between clusters of drawings scores, success in each task or set of tasks calls upon different ancillary abilities and adds new information.

3.2 Developmental aspects

Before investigating the contribution of intellectual abilities to drawing performance as measured by the point and structural scales, it is important to look first at the effect of age on the composite drawing scores and on scores from individual tasks. Point and structural scores from the experimental tasks, as well as raw and standard scores from GHDT and Wechsler scales, were correlated with children's chronological age. The correlation values are reported in Table 12.2. This analysis was considered essential for two more reasons. First, direct comparisons can be made between the GHDT and the drawing scales from experimental tasks on the basis of the magnitude of their correlations with age. Second, this analysis can be useful for deciding what type of scores from the criterion measures subsequent analyses should be based on.

[3] The latter coefficients are higher because they represent part-whole relationships.

Table 12.2 Correlation coefficients between age and various measures of ability.

INSTRUMENTS	N	Raw Scores		Standard Scores
GHDT				
Man	110	.86[c]		.39[c]
Woman	110	.82[c]		.36[c]
Self	110	.84[c]		.37[c]
WISC-III^{UK}/WPPSI				
Verbal (Sum)	110	.14		
Verbal (VIQ)	110			-.25[b]
Vocabulary	110	.42[c]		-.13
Information	110	-.31[c]		-.40[c]
Comprehension/Similarities	110	.02		-.07
Performance (sum)	110	.71[c]		
Performance (PIQ)				-.34[c]
Picture Completion	110	.49[c]		-.45[c]
Block Design	110	.71[c]		-.32[c]
Picture Arrangement/Mazes	110	.53[c]		-.13
Total (sum)	110	.60[c]		
Total (FSIQ)	110			-.35[c]
EXPERIMENTS				
Occlusion point score (OPS)	110	.81[c]		
Axial point score (APS)	110	.72[c]		
Composite point score (CPS)	110	.85[c]		
Occlusion structural score (OSS)	110	.82[c]		
Axial structural score (ASS)	110	.76[c]		
		Point scores	Structural scores	
Occlusion Tasks				
Apples	110	.51[c]	.59[c]	
Bottles	108	.58[c]	.70[c]	
Sunset	110	.60[c]	.59[c]	
Balls MO	108	.60[c]	.62[c]	
Balls FO	107	.59[c]	.55[c]	
Bus	109	.43[c]	.59[c]	
Houses	109	.66[c]	.65[c]	
Vase	110	.31[c]	.31[c]	
Boat	109	.35[c]	.40[c]	
Axial Tasks				
Bowl	102	.36[c]	.45[c]	
Bottle	107	.47[c]	.49[c]	
Mountain	104	.47[c]	.47[c]	
Chimney	106	.60[c]	.67[c]	
Road	103	.25[b]	.40[c]	

Note 1 WISC-III^{UK}/WPPSI: The correlation coefficients reported on the raw scores for the scales (Verbal, Performance, Full scale) are based on the sum of raw scores from the subtests administered (3 Verbal and 3 Performance).

Note 2 The sample size for individual tasks varies as equivocal cases were excluded for this analysis. Point and structural scores for the experimental tasks were computed on the whole sample to avoid excluding children who didn't produce a complete set of drawings.

As expected, both point and structural scores from the spatial tasks correlated significantly with age, and comparable coefficients were obtained for each area of drawing competence (OPS = .81, APS = .72, OSS = .82, ASS = .76). Although the correlational estimates remained significant ($p<.001$), their magnitude varied across particular tasks which reflect the variable pattern of the developmental change in drawing ability, already witnessed at previous chapters. Smaller values were obtained from either easier tasks (Vase) or more demanding ones (Bus, Boat, Road), whereas higher values reflect smoother developmental changes in the representation of the requisite spatial relationship (e.g. Houses, Balls, Sunset, Chimney). The picture remained the same when the correlations were computed on structural scores. What is worth noting though is that the developmental changes were better reflected in the structural scores (strategies) rather than in the point scores for the most difficult tasks (Bus, Bowl, Road).

Two results from this analysis need to be underscored. First, GH raw scores correlated as much with age as composite point scores from the experimental tasks, and comparably with just the overall scores from the occlusion tasks (GHDT = .82-.86, CPS = .85, OPS = .81, OSS = .82). Second, although standard scores from the GHDT and Wechsler scales are correcting for the age factor, nevertheless significant correlations were obtained between these scores and age. Specifically, the correlation estimates for the GH scores were in the high 30s ($p<.001$), a finding which is mainly attributed to the substantially lower scores of the 4 year olds compared with the older age groups (for means see Chapter III, Table 3.3). On the contrary, VIQ, PIQ and FSIQ scores correlated negatively with age (-.25, -.34 and -.35 respectively), as 4 year olds obtained relatively higher scores on these scales compared with their older counterparts (for means see Appendix A, Table I_2).

These results have certain implications for the subsequent analyses of drawing scores from the total sample against criterion estimates. The fact that GH standard scores and those from experimental tasks increased with age is likely to inflate the magnitude of relationship between HFDs and spatial tasks. On the contrary, the negative correlation between IQ scores and age might suppress the degree to which intellectual abilities relate to the drawing performance in experimental tasks. The interference of developmental changes in the scores of criterion measures of ability calls for the elimination of the age factor from these scores when performance of the total sample of children is considered.

In the remaining part of results, the main findings will be reported in three sections. First, the relationship between criterion measures (GHDT, Wechsler scales) and drawing spatial scores will be investigated with correlational analyses, on the total sample and within age levels. Next, the effect of conceptual abilities – as measured by the GHDT and Wechsler scales – on performance in the series of experimental tasks will be further explored with ANOVA. Finally, regression analysis will be carried out to investigate which variable or combination of variables is more powerful in predicting children's IQ scores.

3.3 Correlational analyses

Pearson correlation coefficients were computed between the five spatial scores from the drawing tasks (CPS, OPS, APS, OSS, ASS) with GHDT and Wechsler scales, using both raw and standard scores from the criterion measures. Due to the size of the correlation matrices, a full report of the coefficient estimates is given in the Appendix D, Table 1 for the total sample, and in Tables 2 to 5 for each age group separately. Only the coefficients obtained with standard scores from criterion measures will be reported here which, as expected, were smaller compared to those calculated using raw scores. An account will be given only when there are substantial differences between these two methods of analysis.

3.3.1 Spatial scores and GH scores

For the overall sample, the correlation coefficients between various measures of performance on experimental tasks and GH scales were positive and significant ($p<.001$). Children's composite scores from the spatial tasks (CPS) correlated with all GH scores (Man, Woman, Self) at the range of 50. The coefficients for occlusion scores (OPS) were at the range of 40, whereas those for axial scores fell between low 30s to low 40s. Correlations of comparable magnitude (40-50) were also obtained between GH scores and structural scores for the occlusion and axial tasks.

The comparability of scores changes when the correlational analysis is repeated at each age level. The general picture from Tables 2-5 in Appendix D can be summarized as follows. The relationship between the GH scores and spatial task scores was modest at the age of 4, substantial between 5½ to 7 years of age, and absent at the age of 8. A similar pattern of age differences is found when the analysis is performed using raw scores from the measures, except that the magnitude of correlations is substantially increased for the 4 year old children. As has been previously seen (section 3.2, present Chapter, and in Chapter III, section 3.3.4 and Table 3.3), Harris norms for this age group markedly underestimated children's true drawing abilities which minimizes the comparability between GH standard scores and spatial scores.

Looking at the different measures separately, the highest values were naturally consistently obtained for the CPS. Overall performance in the 14 experimental tasks yielded median coefficients with GH scale scores of .41 ($p<.05$), .64, .68 (both $p<.001$) and .32 ($p>.05$) at each successive age group. Overall, composite spatial scores correlated comparably with the scores from all GH scales at the age of 4, whereas they correlated less with Self scores at the ages of 5½ and 7. Surprisingly, at the age of 8, CPS correlated significantly only with Self scale scores. The fact that exactly the same pattern of results was obtained from the analysis on raw scores from the GHDT substantiates the claim that the obtained variation in the magnitude of the examined relationships is a function of age differences and not an artifact of the raw-to-standard score conversion process.

The results from each area of drawing competence followed the same pattern as the one obtained from the sum of spatial scores. Thus, performance in the occlusion tasks (OPS) correlated less with HFDs at the age of 4, and more at the ages of 5½ and 7, whereas no significant relationship was obtained at the age of 8. Again, the magnitude of correlations remained unchanged across the GH scales for the younger children, whereas performance of the two intermediate age groups at occlusion tasks correlated most with the Man drawing (.66, .62) and least with the Self drawing (.47, .40). Performance in the axial tasks (APS) correlated reliably with GHDT only for the two intermediate age groups. The coefficients were approximately at mid 60s for the 5½ year olds and at mid 50s for the 7 year olds. At the age of 8 a significant relationship was obtained only with the Self drawing again. The analysis on structural scores from occlusion and axial tasks virtually revealed the same pattern of results as the one obtained from the point scores of each domain. The only difference was that this time the ASS of 4 year olds correlated moderately with performance on the Man and Woman scales.

Summarizing the results, the examined relationships were stronger at the age of 5½ and 7 for all types of measures and scores from the spatial tasks. At the age of 4 the moderate relationship between CPS and GHDT is attributed primarily to the significant relationship between OPS with GH scores. Similarly, at the age of 8 the relationship between the measures was rather restricted since only drawing performance on axial tasks was related exclusively to GH score from the Self scale.

3.3.2 Spatial scores and Wechsler scores

At the next stage of analysis, spatial scores were correlated with IQ scores from the Wechsler scales. Performance in the whole series of spatial tasks (CPS) correlated at .39 with FSIQ ($p<.001$), whereas the coefficient obtained for axial scores was somewhat higher than for occlusion scores (APS = .34, $p<.001$; OPS = .29, $p<.01$). The nature of this relationship was further investigated, separately for Wechsler Verbal and Performance scales. Except for the OPS which didn't correlate with VIQ, significant correlations were obtained for the APS and CPS with both Wechsler scales. Overall, the coefficients were consistently higher with PIQ than with VIQ. The significance of these differences was examined using t-tests for dependent correlations (Howell, 1982, p. 243). The findings showed that point scores from axial tasks correlated comparably with PIQ and VIQ scores, whereas for occlusion and composite scores the difference was reliably in favor of PIQ (for OPS, $t_{107} = 2.15, p<.05$; for CPS, $t_{107} = 2.44, p<.01$).

Correlational estimates between drawing scores and subtests from the Wechsler scales were also obtained. All point scales from experimental tasks correlated significantly with the three subtests from the Performance scale between .26 to .46. The best relationship was systematically found to be with Block Design, whereas the magnitude of correlations with the Picture Completion and Mazes/Picture Arrangement were comparable. On the contrary, the only significant correlation with the verbal subtests was with Vocabulary which correlated at mid 20s with APS and CPS, but not with OPS. The coefficients for structural scores from occlusion and axial tasks revealed a similar pattern. No reliable relationship was obtained between structural scores and performance on the Verbal scale and subtests, whereas structural scores correlated with PIQ and FSIQ at mid

20s. Considering subtests of the Performance scale, only Picture Completion didn't correlate with structural scores, whereas the strongest relationship was obtained with Block Design.

Next, the relationship between experimental tasks and IQ scores was examined at each age level. The pattern of general results showed that, at the age of 4, a significant relationship was obtained only between few scores, while at the age of 5½ there was no reliable relationship between the variables. On the contrary, at the age of 7 performance in the experimental tasks correlated significantly with nearly all measures of conceptual maturity, and at the age of 8 the relationship, though still significant, was restricted to certain measures.

In particular, for the 4 year olds, significant relationships (mid 40s) were obtained between performance on occlusion tasks, and in the total series of spatial tasks (CPS) and PIQ from WPPSI. Considering Performance subtests, Block Design correlated with both point and structural scores from the occlusion tasks, whereas the former correlated with the Mazes subtest as well[4]. At the age of 5½ no reliable relationship was found with any sort of ability measure except that OSS correlated moderately with Block Design from the Performance scale[5].

A rather different picture emerged at 7 years of age. Total experimental scores correlated with FSIQ and PIQ around .70, and with VIQ at .50. Further, performance on axial tasks, measured by point and structural scores, correlated significantly with all IQ scores at a range of .57 to .69, while occlusion scores related only with FSIQ and PIQ scores at .48. In addition, significant relationships were found between nearly all Performance subtests and all types of measure from

[4] These results were also found using raw scores from the Wechsler scale. The only difference was that CPS correlated with all the subtests from the Performance scale.

the experimental tasks (OPS, APS, CPS and OSS, ASS)[6]. The coefficients with Verbal subtests were still significant, albeit of a smaller magnitude, whereas only Vocabulary systematically failed to relate to any measure from the experimental tasks. Finally, at the age of 8, point scores from all spatial tasks and from each area of drawing ability correlated with FSIQ between .39 to .50. Overall performance on the experiments also correlated significantly and comparably with VIQ and PIQ; yet a moderate relationship was found only between occlusion scores with VIQ. With respect to the Wechsler subtests, basically Block Design from the Performance scale and Vocabulary from the Verbal scale correlated significantly and systematically with various measures of drawing performance.

3.3.3 Summary of correlational results

A simple summary of the correlational analyses between performance on spatial tasks with criterion measures is as follows. As it was expected, the scores from drawing tasks correlated with GHDT. For the total group, this pattern was consistently found for each domain-specific drawing ability measured by point as well as structural scores and eventually for the whole range of experimental tasks. Variation in the magnitude of correlations was found between the age groups. At the age of 4 a weak positive relationship was found primarily for occlusion tasks. At the two intermediate ages reliable relationships were found between all drawing measures with GH scales, whereas at the age of 8 only axial scores and Self scores were correlated. Further, the degree of the examined relationship also varied across the GH scales at each age level. So, although no scale differential was obtained at the age of 4, experimental scores correlated best with Man scores at intermediate ages, and only with Self scores at the age of 8.

---

[5] Some differences in the pattern of correlations were obtained when raw scores for the Wechsler scale were used. Thus, moderate relationships were found between more drawing scores and the Performance scale and Block Design subtest.

[6] Scores from the occlusion tasks didn't correlate with Block Design subtest.

Considering Wechsler scores, the results from the total sample suggest that drawing scores correlate with FSIQ at the range of .30, systematically and reliably with PIQ and less so with VIQ. Of the subtests of the Wechsler scales, Block Design was notably one which yielded consistently the highest coefficients with the drawing scores. As with GHDT, the pattern of corrrelations changed by age. Thus, at the age of 4 the examined relationship was rather domain-specific as a moderate relationship was found between occlusion scores and Performance scores from WPPSI, whereas at the age of 5½ drawing performance basically didn't correlate with IQ scores. On the contrary, for the 7 year olds the extent to which the various drawing measures correlated with performance on WISC-III^{UK} and the magnitude of these correlations increased notably. Significant relationships were found between all types of drawing scores with all measures of ability, except with the scores from the Vocabulary subtest. Finally, 8 year olds yielded significant correlations between drawing scores from the entire set of experimental task and separately from each drawing domain with FSIQ. Also, overall experimental scores correlated comparably with VIQ and PIQ, but this changed within each field of drawing competence. Yet, only the Vocabulary and Block Design subtests correlated systematically with the drawing measures.

3.4 Factorial analyses

Having examined how drawing performance on spatial tasks compares with performance on non-verbal and conventional measures of intellectual ability, it seemed desirable to broaden the scope of the investigation. The contribution of conceptual maturity to the drawing scores from the experimental tasks was further explored by stratifying children at different levels of ability on the basis of their IQ estimates, derived from GHDT and Wechsler scales. If performance on the series of drawing experiments depends on children's intellectual skills, then children of different conceptual abilities should demonstrate different degree of drawing competence.

The sample was divided into two groups; children whose IQ estimates were below the median value were placed in the low-ability group and those whose IQ estimates were above the median value were placed in the high-ability group. In view of the significant correlations between age and standard scores from criterion measures, it was decided to take as a cut-off point the median score of each age group rather than the median of the total sample. In the analysis of drawing scores by IQ estimates, age and gender were also included as factors in order to investigate their particular effect separately, as well as any possible interactions between the variables. The ANOVAS involving GH standard scores as an ability factor will be reported first.

Since GHDT is comprised of 3 scales (Man, Woman, Self), three estimates of ability were used. The distribution of children to a low and a high ability group on the basis of their scores on each scale is presented in Table 12.3. The dependent measures were five: three sets of point scores (CPS, OPS, APS) and two types of structural scores (OSS, ASS). Initially, a 4 (age) × 2 (ability) × 2 (gender) ANOVA was conducted for each dependent variable, separately for each GH scale.

Table 12.3 Distribution of children into ability groups based on GH scores by age.

		GH-Man SS			GH-Woman SS			GH-Self SS		
Age	*N*	Median	<M	>M	Median	<M	>M	Median	<M	>M
4	30	87	16	14	85	17	13	84	17	13
5½	25	107	14	11	99	13	12	95	13	12
7	28	105	15	13	106	16	12	102	14	14
8	27	105	15	12	100	14	13	102	15	12
Total	110		60	50		60	50		59	51

The findings were conspicuous for the absence of main effects of gender and of any interactions involving gender on the dependent drawing measures. In view of these results this factor was eliminated from the ANOVAs. The findings from the 4 (age) × 2 (ability) ANOVAs revealed a consistent pattern, as they yielded significant main effects of age and ability for all dependent measures of drawing performance regardless of the GH scale used to classify children into ability groups (all $p<.001$). Due to the large number of ANOVAs performed, the values of F statistics with their corresponding level of significance are summarized in Appendix D, Table 6, whereas the means from these analyses are presented hereinafter in Table 12.4.

Table 12.4 Mean scores from drawing scales by age and ability level across GH scales.

	Man SS		Woman SS		Self SS		
	<Med	>Med	<Med	>Med	<Med	>Med	Total
A.CPS (0-14)							
4	2.1	2.7	1.5	3.5	1.9	2.9	2.4
5½	5.1	8.2	5.1	7.8	5.3	7.7	6.5
7	7.3	10.0	7.4	10.1	7.6	9.6	8.6
8	10.0	11.2	10.4	10.8	9.9	11.3	10.6
Total	6.1	8.0	6.1	8.0	6.2	7.9	7.1
B. OPS (0-9)							
4	1.7	2.3	1.2	3.0	1.5	2.5	2.0
5½	3.9	5.9	3.8	5.8	4.2	5.4	4.8
7	5.3	6.8	5.4	6.8	5.4	6.6	6.0
8	7.1	7.9	7.5	7.5	7.3	7.7	7.5
Total	4.5	5.7	4.5	5.8	4.6	5.6	5.1
C. APS (0-5)							
4	.4	.4	.3	.5	.4	.4	.4
5½	1.2	2.3	1.3	2.1	1.2	2.2	1.7
7	2.0	3.1	2.0	3.3	2.1	2.9	2.5
8	2.9	3.3	2.9	3.3	2.6	3.7	3.1
Total	1.6	2.3	1.6	2.3	1.6	2.3	1.9
D. OSS (0-27)							
4	14.4	15.5	13.6	16.6	14.3	15.7	14.9
5½	18.9	22.6	18.8	22.3	19.5	21.7	20.5
7	22.3	24.2	22.6	24.0	22.3	24.1	23.2
8	24.8	25.6	25.2	25.1	25.1	25.2	25.2
Total	20.1	22.0	20.1	22.0	20.3	21.7	21.0
E. ASS (0-20)							
4	9.3	9.2	8.6	10.2	8.3	10.5	9.3
5½	13.6	15.4	13.7	15.2	13.4	15.6	14.4
7	14.8	17.2	14.8	17.4	15.3	16.6	15.9
8	16.7	17.6	16.7	17.5	16.3	18.1	17.1
Total	13.6	14.9	13.4	15.1	13.3	15.2	14.2

Note 1 CPS = Composite point score, OPS = Occlusion point score, APS = Axial point score, OSS = Occlusion structural score, ASS = Axial structural score.

Note 2 <Med = Drawing mean scores of children whose GH score was below the criterion median value >Med = Drawing mean scores of children whose GH score was above the criterion median value.

Note 3 Total drawing mean scores at each age level are those obtained from the univariate analysis on the dependent drawing measure. These means varied slightly across the GH scales.

Inspection of the Table (column on totals) reveals clear developmental differences in the mean drawing scores for all task measures, suggesting that with age children gradually pass more drawing tasks from the point scales (CPS, OPS, APS) and use more appropriate drawing devices to represent spatial relationships (OSS, ASS). Post-hoc Newman-Keuls tests ($p<.05$) were also carried out to investigate where the age differences were located. The analyses showed that each age group performed significantly different from the others, with a smooth increase of score points from the age of 4 to the age of 8, except for ASS where the two older groups yielded comparable mean axial structural scores. Further, children in the low-ability group, as indicated by their GH score, yielded drawing scores in the experimental tasks approximately 1 to 2 points lower compared to those in the high-ability group, and these differences were highly significant in all analyses ($p<.001$).

Finally, looking at the Table 12.4 it appears that when point scores were analyzed against performance on the Man scale, mean differences between ability groups were more pronounced at the two intermediate age levels. Yet, an interaction between age and ability reached significance only in the analysis of CPS using Man scores as criterion measures ($F_{3,102} = 3.00$, $p<.05$). A rather different pattern of interaction was obtained in the analysis of OPS and OSS against scores for Woman scale (OPS: $F_{3,102} = 2.87$; OSS: $F_{3,102} = 2.90$, both $p<.05$). Here the differences in mean occlusion scores (point and structural) were substantial and comparable to the differences in ability obtained from woman figure drawings up to the age of 7, but they were absent at 8 years of age.

Next, the same type of analyses was conducted, using IQ scores from the Wechsler scales. Table 12.5 presents the distribution of children into high and low-ability groups at each age level on the basis of their FSIQ, VIQ and PIQ. Again, a 4 (age) × 2 (ability) ANOVA[7] was conducted for each of the five dependent variables (CPS, OPS, APS, OSS, ASS), separately for IQ type. Pertinent *F* values from these analyses with their corresponding level of significance can be found in Appendix D, Table 6.

Table 12.5 Distribution of children into ability groups based on IQ scores by age.

		FSIQ score			VIQ score			PIQ score		
Age	*N*	Median	<M	>M	Median	<M	>M	Median	<M	>M
4	30	116	16	14	112	18	12	114	16	14
5½	25	112	14	11	110	13	12	111	14	11
7	28	101	14	14	99	14	14	97	14	14
8	27	103	14	13	100	14	13	104	14	13
Total	110		58	52		59	51		58	52

Naturally, the findings relating to the main effects of age didn't add any new information to the previous analyses; age had a significant effect on all five dependent drawing measures and post-hoc comparisons revealed significant differences between any two age groups, except for ASS where no differences were found between the two older groups (see Table 12.6). Generally, the level of conceptual ability, measured by children's FSIQ, had a significant effect on the dependent measures. Looking at Table 12.6, children falling into the low-ability group yielded mean drawing scores approximately 1 point lower compared to the more intellectually competent. These differences were statistically significant for all drawing measures except for OSS where the difference just fell out of significance ($F_{1,102} = 3.61, p = .06$). The same pattern was also obtained with PIQ as index of non-verbal ability. Children yielding higher PIQ scores performed better compared to those with lower PIQ scores for all drawing scales. Finally, verbal abilities appeared to have a selective effect on drawing scores. Performance on occlusion tasks – measured either by point or structural scores – was unaffected by the level of VIQ, whereas children with higher VIQ obtained somewhat higher scores on the drawing tasks of spatial

[7] Age was included to examine any interactions and gender was excluded based on previous results.

coordinates compared to the less competent ones in this respect. Yet, the main effect of VIQ level fell marginally out of significance when the composite scores from all experimental tasks were analyzed (CPS: $F_{1,102} = 3.52$, $p = .06$).

Table 12.6 Mean scores from drawing scales by age, type and level of IQ score.

	FSIQ		VIQ		PIQ		
	<Med	>Med	<Med	>Med	<Med	>Med	Total
A.CPS (0-14)							
4	1.9	2.9	2.6	2.0	1.6	3.2	2.4
5½	6.0	7.0	6.2	6.8	6.1	6.9	6.4
7	7.4	9.7	7.8	9.4	7.4	9.7	8.6
8	9.9	11.2	9.9	11.2	10.3	10.9	10.6
Total	6.3	7.7	6.6	7.3	6.6	7.7	6.9
B. OPS (0-9)							
4	1.5	2.5	2.3	1.5	1.3	2.7	2.0
5½	4.5	5.1	4.5	5.0	4.6	5.0	4.8
7	5.6	6.5	5.7	6.4	5.4	6.6	6.0
8	7.2	7.8	7.2	7.8	7.6	7.4	7.5
Total	4.7	5.5	4.9	5.2	4.7	5.4	5.0
C. APS (0-5)							
4	.4	.4	.3	.5	.3	.5	.4
5½	1.5	1.9	1.6	1.8	1.5	1.9	1.7
7	1.9	3.2	2.1	3.0	2.0	3.1	2.5
8	2.7	3.5	2.7	3.5	2.7	3.5	3.1
Total	1.6	2.2	1.7	2.2	1.6	2.2	1.9
D. OSS (0-27)							
4	14.3	15.6	15.7	13.7	13.9	16.1	14.9
5½	20.2	20.9	20.2	20.8	20.1	21.0	20.5
7	22.6	23.9	22.6	23.8	22.5	23.9	23.2
8	24.8	25.5	24.8	25.5	25.2	25.1	25.2
Total	20.5	21.5	20.8	21.0	20.4	21.5	21.0
E. ASS (0-20)							
4	8.6	10.0	9.0	9.7	8.5	10.1	9.3
5½	14.3	14.6	14.2	14.7	14.1	14.8	14.4
7	14.6	17.2	15.0	16.9	14.8	17.1	15.9
8	16.5	17.7	16.5	17.7	16.6	17.5	17.1
Total	13.5	14.9	13.7	14.7	13.5	14.9	14.2

Note 1 CPS = Composite point score, OPS = Occlusion point score, APS = Axial point score, OSS = Occlusion structural score, ASS = Axial structural score.

Note 2 <Med = Drawing scores of children whose IQ scores were below the criterion median value
>Med = Drawing scores of children whose IQ scores were above the criterion median value.

Note 3 Total drawing mean scores at each age level are those obtained from the univariate analysis on the dependent drawing measure. These means varied slightly across the different types of IQ.

Finally, the ANOVAs revealed only one moderate interaction between age and FSIQ level in the analysis of point scores from axial tasks (APS: $F_{3,102} = 2.97, p<.05$). The interaction showed that intellectual abilities had no effect on performance on axial tasks up to the age of 5½. From this age onwards children with higher FSIQ perform better, and this difference was more pronounced at the age of 7.

A closer look at children's IQ scores from the Wechsler scales showed that there was a number of children who yielded conspicuously different VIQ and PIQ scores. To account for the possibility that these cases might have affected the previous results in any respect, the analyses on drawing scores were run again excluding 40 cases where IQ discrepancies exceeded 11 points in any direction[8]. Confining the analysis to the sample of children with similar verbal and performance abilities changed the previous results mainly in relation to PIQ. Thus, the level of PIQ failed to show an effect on drawing performance, assessed with OPS, OSS and ASS (OPS: $F_{1,52} = 1.54$, OSS: $F_{1,52} = .26$, ASS: $F_{1,52} = 3.42$, all $p>.05$). Only for the effect of VIQ level on CPS did it just reach significance ($F_{1,52} = 4.11$, $p = .048$). Still, the main effect of FSIQ level retained its significance, as in the previous analyses on the total sample.

Finally, to investigate whether the direction of difference in VIQ and PIQ scores had any effect on drawing performance, the five drawing scores were analyzed using an unweighted 4 (age) × 2 (type of IQ difference) ANOVA[9]. Differences favoring PIQ had a moderate, yet significant effect on CPS and OPS, as children whose PIQ was significant higher to their VIQ performed better on occlusion tasks and on the total series of experiments (CPS: $F_{1,42} = 4.50$, OPS: $F_{1,42} = 5.18$, OSS:

[8] Wechsler Manuals were consulted to determine the magnitude of difference required between Verbal and Performance IQ for statistical significance (Wechsler, 1967, p. 23-24, Table 9; Wechsler, 1992, WISC-III^{UK}, p. 66 & 266, Table B1).
[9] 17 children had VIQ>PIQ and 33 had PIQ>VIQ.

$F_{1,40} = 5.75$, all $p<.05$)[10]. However, no effect of type of IQ discrepancy was obtained for the drawing scores from axial tasks.

In conclusion, the main results from the factorial analyses revealed the following consistent patterns. First, clear developmental changes were obtained from all drawing spatial scales. With age, there was a gradual improvement in the use of the appropriate device as well as in the choice of alternative solutions, present in the whole series of experimental tasks and within specific areas of drawing competence. Second, differences in conceptual abilities, assessed either by a non-verbal measure (GHDT) or by a conventional test of intellectual maturity (Wechsler) affected performance in the series of drawing tasks[11]. Children who yielded a GH score or a FSIQ above the median of same age children passed more drawing tasks and adopted a better drawing strategy compared to the less intellectually competent. These effects were stronger when GH scores were used. Comparing the separate contribution of verbal and performance abilities from the Wechsler scales, it appeared that verbal abilities had a selective effect, as they interfered moderately with performance only on axial tasks when the total sample of drawings was analyzed. Finally, gender had no effect on any dependent drawing measure.

3.5 Regression analyses

Finally, three regression analyses was carried out, with age, GH scales and experimental point scales (CPS, OPS and APS) as independent variables, and IQ scaled scores as dependent ones. Table 12.7 presents the results from the regression analyses separately on FSIQ, VIQ and PIQ, using the forward method.

[10] The means from these analyses were (a) CPS PIQ>VIQ: 7.6 vs. VIQ>PIQ: 6.1, (b) OPS PIQ>VIQ: 5.4 vs. VIQ>PIQ: 4.2 (c) OSS PIQ>VIQ: 21.6 vs. VIQ>PIQ: 19.5.

[11] At a later stage, the ANOVAS were repeated stratifying children into ability groups on the basis of the median score of the total group for each criterion measure rather than the one obtained at each age level.

Table 12.7 Results from forward regression on IQ scores.

Dependent Factor	Model/Predictors	R	R^2	Residuals
A. FSIQ	1. Age	.35	.12	13.97
	2. Age + CPS	.54	.29	12.62
B. VIQ	1.Age	.25	.06	14.11
	2. Age + CPS	.35	.12	13.72
C. PIQ	1. Age	.33	.11	16.26
	2. Age + CPS	.55	.30	14.45
	3. Age + CPS + GH-WS	.58	.33	14.22

Note 1 Age is entered as chronological age.
Note 2 CPS = Composite point score, GH-WS = Goodenough-Harris Woman standard score

Looking at R^2 values, it is evident that a very small proportion of the variance in IQ scores can be explained by the chosen variables. However, among all the factors, age was consistently the only one which could explain most of the variability in IQ scores on its own, whereas, next to age, CPS was always added to the models as the one which could produce an appreciable increment in R^2. No other factors were entered for FSIQ and VIQ, but for PIQ, GH-scores from the Woman scale were entered in the model as a third factor. Although it is not surprising that VIQ was the least explained by the selected factors, it is rather unexpected that composite spatial scores from the experimental tasks had relatively better predictive validity than the GH scores.

4. DISCUSSION

In the present study an attempt was made to reconcile the different areas of drawing competence studied in this thesis, to assess their efficacy in revealing developmental changes and conceptual abilities, and finally to investigate the nature of abilities underlying the representational process at different stages of development. Issues concerning children's overall performance on spatial

The analyses revealed identical results as for the main effects of age and ability across all different criterion measures, providing further support for the previous results.

tasks will be discussed first followed by an account of the main results from the comparative analyses.

4.1 The experimental spatial scales

The developmental character of the drawing scales derived from the experimental tasks was borne out by the significant effect of age on the spatial scores in the ANOVAs. Distinctive developmental trends were evident in the (a) number of drawings tasks children passed from the entire set of experiments and from each particular area of drawing competence, and in the (b) overall level of their sophistication in the use of drawing devices to represent the requisite spatial relationships. In this respect, the spatial scales could reliably discriminate children between 4 to 8 years of age.

Additional evidence was obtained from the correlational analyses. Considering that the spatial scales consist of a notably smaller number of drawing items (tasks) compared to GH scales, and yet are equally sensitive to developmental changes, they constitute a promising and reliable method for the assessment of children's ability to handle the structural aspects of drawing. A similar range of correlational estimates was also obtained from the tasks comprising each domain-specific spatial scale. Yet, looking at the magnitude of task-specific correlations, it appears that developmental changes in drawing ability are more readily tapped by occlusion tasks than by axial tasks[12]. Following this, it appears that, for children between 4 to 8 years of age, the pictorial representation of spatial relationships involving occlusion is relatively easier than the portrayal of angular relationships requiring a reference system of invariant axes. This is in line with the Piagetian theory of spatial development (Piaget & Inhelder, 1948/1967), presented in Chapter VIII, and with the empirical findings from the axial tasks, discussed in Chapters IX-XI. Since the mastery of topological and projective concepts precedes the acquisition of Euclidean concepts, solving graphically problems of linear inclination by the use of an external axial system should lag somewhat behind the pictorial representation of spatial overlap by means of occlusion.

4.2 Spatial scores and GHDT

The correlational analysis of composite scores from spatial tasks and GH scores from HFDs showed that, for the whole age range studied, the relationship is significant, yet its magnitude is average (.50). This picture didn't change when spatial scores were compared either with raw GH scores or IQ equivalents. Obviously, the spatial scales and the GH scales sample partially overlapping skills, and this could explain the moderate relationship between these measures of drawing competence. The HF drawing test measures a variety of representational skills, ranging from degree of detailing, to proportions, motor coordination and spatial integration of features, whereas the experimental tasks tap children's dexterity in depicting particular spatial layouts from a vantage point.

However, if we accept that GHDT provides a crude index of children's intellectual abilities – probably of a non-verbal nature – the results seem to suggest that performance in the spatial tasks is not totally dependent on age, but on the level of intellectual functioning as well. This assertion is fully supported from the analysis of spatial scores by age and IQ level as indicated by GHDT. The results showed that children in the upper range of IQ estimates from GH were more competent in incorporating the view-specific properties of the spatial arrays studied here. The

[12] From the tasks on spatial coordinates, only the Chimney task was relatively most sensitive to age-related differences, whereas comparable coefficients were obtained with a number of tasks from occlusion scenes

absence of interaction between age and GH-IQ level, virtually for all spatial scores, indicates that the level of conceptual maturity affects the development and the representation of the pertinent spatial concepts evenly at all ages studied here.

Nevertheless, the correlational analyses within each age level suggest that individual GH scores do not always parallel spatial scores. This has certain implications for the nature of skills the GHDT measures at different age levels. Before discussing the results from this analysis in detail, it is worth noting some inconsistencies observed when raw scores instead of standard scores from GH scales were used in the pertinent analyses. First, these inconsistencies were observed in the magnitude rather than the pattern of relationship, and they were obtained mainly with the 4 year olds. The substantial increase in the magnitude of correlations when the analysis was performed on raw scores substantiates the findings in Chapter III which showed that GH norms for pre-school children underestimate their conceptual abilities. The present results suggest that this bias is also present when representational abilities are assessed by another drawing measure. Further support for this bias is also provided by the unexpected positive correlation between GH standard scores (IQ equivalents) with age, mainly attributed to the lower scores of pre-school children compared to their older counterparts.

Returning to the issue of the age-related variation in the degree of relationship between spatial scores with GH scores, if one could sketch it, it could form an inverted U; for the younger and older children the relationship is moderate and largely dependent on the domain of spatial competence measured by the experimental tasks, whereas for the two intermediate ages, it is stronger and more general. Evaluating this finding along with the results from the ANOVAs the following conclusion can be drawn. Knowing under which range of ability a child's IQ estimate

(e.g. House, Balls, Sun, Bottle, see also Table 12.2).

from GHDT falls, one can reliably differentiate low and high performers in spatial tasks at each age level. Still, correlational evidence shows that a GH score is a better predictor of performance on spatial tasks for children of intermediate ages than it is for the younger and older age group. What seem to account for these seemingly irreconcilable results is the changing nature of relationship between spatial scales and GH scales at different ages. To be able to interpret these findings, it seems necessary at this point to reiterate the findings relating to the performance of the age groups within and between the two spatial domains of drawing ability studied here.

Experimental evidence showed that the spread of performance across domain-specific tasks varied at different age levels. The overall findings from occlusion tasks, presented in Chapter VII, showed that the performance of 5½ and 7 year olds was extremely variable, whereas for the 8 year olds view-specific responses prevailed in the majority of tasks. On the contrary 4 year olds, being cognitively and graphically restricted, displayed the least task variation. Similar conclusions were drawn in Chapter XI where the general results from tasks on spatial coordinates were discussed. However, Euclidean tasks, being more difficult than occlusion tasks, didn't elicit successful responses from the vast majority of the 4 year olds. The global impression that one gets from these findings parallels what the literature, concerned with the emergence of visual realism, suggests; children at intermediate ages of drawing development are the most susceptible to task manipulations aiming to enhance view-specificity, and it is at these ages where the traditional stage theories have the least explanatory power.

What these qualifications indicate is that the variation in the range of spatial scores across age groups obviously affects the degree of relationship with GH scores. Yet, beside the issue of score variability one could justifiably argue that the development of graphic formulas to represent the human figure, spatial overlap, and Euclidean concepts progress at a different pace across the age groups studied, and that the range of representational abilities that the HF test samples, varies accordingly. At the age of 4, children have already developed a conceptual and graphic schema for the human form which, according to the findings in Chapter III, mainly captures its most defining features without these being structurally integrated. When it comes to their level of competence in representing particular spatial relationships, 4-year-old's limitations in dealing with pictorial space are more pronounced.

Empirical evidence suggests that children of this age can comprehend that the incomplete forms, resulting from the use of the HLE device, is a pictorial means to represent spatial overlap (Beal & Arnold, 1990; Hagen, 1976). On the contrary, they don't understand the physical principles related to the Piagetian water-level and vertical tasks, since they lack an invariant system of coordinate axes. This is in line with early assertions that some spatial aspects are learned earlier than others, and that the detection and reproduction of 'incompleteness' in a form appear earlier than the reproduction of angles (Harris, 1963, p.184). Indeed, the present findings show that, while a few of the 4 year olds can occasionally succeed in occlusion tasks, nearly all produce inaccurate orientations in axial tasks. Thus, in the absence of variability of axial point scores, obviously no relationship can be obtained with GH scores. However, even with the low variability of occlusion scores, a moderate relationship is obtained, suggesting that children with the more differentiated HFDs are more likely to differentiate the requisite spatial relations pictorially. This provides partial support to Goodenough's claim that, although for young children the quantitative aspects of a form (features) may be grasped earlier than its spatial aspects, the more intellectually competent a child is, the greater the overlap between these two types of encoding (1926, p.73-74).

With age, children's conceptual and graphic schemes undergo major alteration. Item analysis on GH scales in Chapter III shows that HFDs from the age of 5½ become more elaborate both in terms of content and structure. Not only do pictorial forms begin to lose their geometric properties but spatial relationships between features start to become graphically explicit. As a consequence, from this age onwards there should be greater overlap in the abilities measured between GHDT and tasks or tests assessing spatial skills. Empirical findings from occlusion and axial tasks have shown that children between 5½ to 7 years of age have some limited ability to represent view-specificity. This suggests that they don't totally lack the requisite spatial concepts, only that these are not readily utilized (competence-utilization hypothesis). As the significant correlation between GH scores and spatial scores suggests, the more competent children are in the area of intellectual functioning tapped by the GHDT, the fewer the occasions are which constrain them in utilizing their spatial concepts, and thus the more tasks they pass.

Paradoxically, a different picture was obtained from the 8 year olds. In parallel with previous findings which suggest that the use of the HLE device is pictorially mastered around the age of 8 (Cox, 1978; Freeman, *et al.*, 1977), this group performed at ceiling in occlusion tasks. Similarly, in agreement with the Piagetian theory, which asserts that a fully internalized system of invariant axes is not fully acquired by the age of 8, performance in axial tasks was adequately variable. Nevertheless, their individual GH scores could neither reflect the level of domain-specific spatial abilities nor their overall level of performance in the whole series of experimental tasks. What is even more puzzling is that the only axial scores which correlated with GH scores did so exclusively with GH-Self scores.

The review of the GHDT literature in Chapter II has showed that the utility of the test in the assessment of drawing ability reaches a plateau at the age of 10 approximately. The empirical findings of the present study also showed that from the age of 7 to the age of 8 there was a smaller increment of score points and a negligible appearance of novel features, whereas reliable developmental differences in the frequency of GH items were obtained only for those relating to the structural aspects of small graphic units like nose, neck, fingers (see Chapter III, Section 3.4.1 and Appendix A, Tables A_M and A_W). In view of these findings, two hypotheses can be proposed. It is possible that 8 year olds can readily demonstrate a certain level of competence when they are tested with simple spatial arrays which necessitate the coordination of a small local group of objects but they cannot demonstrate the same degree of competence in the portrayal of more complex themes, such as the HFDs. Alternatively, it may be that the scorable items in the GH scales are neither sensitive enough nor pertinent to children's abilities in the representation of pictorial space from age 8 onwards. Thus, investigating the examined relationships with older children should show which of these speculations is correct.

Considering the moderate correlation between axial scores and Self scores, it is likely that the performance differential across the GH scales could have accounted for this inconsistency. The Self scale yielded both a lower mean score and the narrowest range of scores compared to the other scales. In this respect, the distribution of axial scores might have been more congruent with the distribution of Self scores than with those from the other scales. Again, this is speculative, and investigating this relationship with set of items from GH scales sampling different aspects of HFDs might be informative (e.g. see Strommen, 1987).

In summary, the magnitude of overlap between the skills measured by the GHDT and the spatial scales varies with age. The drawing competence used in the representation of the particular spatial relationships studied here is more readily tapped by the GHDT between the age of 5½ to 8. The developmental separation in the acquisition of the HLE device and Euclidean concepts seems to be more pronounced with the younger and the older children. As a consequence, at these ages GH scores correlate only with domain-specific spatial scores. On the contrary, at times where these spatial skills are developing, GH scores are a better predictor of performance. However, the drawing test doesn't seem to represent the degree of spatial competence achieved by the 8 year olds. Possibly, from this age onwards, the human figure as a drawing topic has limited utility in uncovering children's abilities in the representation of pictorial space.

4.3 Spatial scores and Wechsler scores

Analysis of variance of spatial scores by range of ability again revealed that ability status has an effect on graphic competence evenly within all age groups. Drawing scores from spatial tasks were also correlated with IQ scores from the conventional measure (Wechsler scales). As expected, the magnitude of coefficients was somewhere lower than that obtained when these scores were plotted against IQ equivalents from GHDT. However, what is worth noting is that, for the overall sample, composite spatial scores from experimental tasks and standard drawing scores from GHDT correlated comparably with criterion scores of intellectual ability. The empirical findings from Chapter III showed that, although there was some moderate variation in the validity coefficients for the three GH scales, overall, GH scores correlated significantly with FSIQ in the mid 30s, with PIQ at 40s and with VIQ around 20. Looking at the analogous values for the spatial scores, a striking similarity is revealed both in the pattern and in the magnitude of these coefficients[13]. These findings deserve attention since they demonstrate that the skills involved in the series of spatial tasks relate to conventional measures of ability to the same extent as those measured by a more general drawing topic, such as the GHDT.

[13] The observed congruity in the pattern of relationships was upheld in subsequent analyses where raw scores from GHDT were correlated with PIQ, VIQ and FSIQ to match the correlation analysis with spatial scores. The magnitude of correlation coefficients remained unchanged (GHDT-VIQ: .14-.24, GHDT-PIQ: .33-.40, GHDT-FSIQ: .28-.39).

This uniformity was maintained even when the drawing scores were analyzed against criterion scores from Wechsler sub-tests. The analyses revealed that both spatial and GH scores correlated reliably with all the subtests from Performance scale, whereas only Vocabulary from the Verbal scale correlated with performance on the drawing measures. Taken together, these results indicate that for children between 4 to 8 years of age, drawing abilities of a nonspecific or a spatial character are associated moderately with general intellectual abilities and more so with non-verbal skills compared to verbal skills. This conclusion still holds when each area of spatial competence is examined separately except that verbal abilities are involved only in performance on tasks of spatial coordinates and not on tasks of occlusion.

The conspicuous regularity in the way performance with various drawing instruments correlates with children's IQ fully supports the conclusions drawn in Chapter III. The moderate validity of GHDT as a measure of intellectual maturity results from the way it operationalizes intelligence. The present results show that the drawing tasks measure a narrow scope of cognitive operations. Mental components such as verbal abstraction are not as readily tapped as spatial concepts and visuo-motor abilities. However, the magnitude of correlations between drawing tests and PIQ suggests that different aspects of non-verbal abilities are tapped by the 'performance component' of Wechsler scales compared to the drawing scales.

Further, early conjectures (Harris, 1963) and empirical findings from Chapter III have indicated that the contribution of various mental abilities to drawing performance changes with age. Similar conclusions are drawn from the results of correlational analyses separately for each age group. At the age of 4, the relationship is restricted, as only PIQ is associated mainly with performance on occlusion tasks, whereas at the age of 7 reliable relationships are obtained between all spatial scales and all types of IQ scores. This picture not only parallels the one obtained from the correlational analyses between spatial scores and GH scores for these two age groups but also between GH scores and IQ estimates.

The congruent pattern in which the two sets of drawing scores correlate with criterion IQ scores was maintained when the analysis was extended to the subtest scores from the Wechsler scales. The findings in Chapter III showed that at the age of 4, GH scores correlate only with the Mazes subtest from the Performance scale, a relationship also found with occlusion scores, except that the latter correlate with Block Design as well. On the contrary, at the age of 7 all drawing scores from the GHDT and from the spatial point scales correlate with all Performance subtests and with the Information and Similarity subtests from the Verbal scale (see also Appendix A, Table V_2).

The comparability of scores across the various measures follows a less consistent pattern at the age of 5½ and 8. Surprisingly enough, at the age of 5½ performance on experimental spatial tasks didn't correlate with children's IQ scores from the Wechsler scales, although their GH scores correlated moderately with PIQ and FSIQ. The opposite picture was obtained at the age of 8, where, despite the absence of relationship between GH scores and IQ scores, overall performance on spatial tasks correlated moderately with all types of IQ (FSIQ, VIQ, PIQ). As has been previously maintained with respect to 8 year olds, the limited validity of the HFDs in the assessment of cognitive operations may be due to the idiosyncrasies of the GH scales or to the nature of the topic *per se*. The human form may impose certain limits to the extent that it can be differentiated on the basis of structure compared to content (details). Possibly only artistically gifted children will overcome these limitations, and this might explain the levelling off of the GHDT by the time children lose interest in details. The results of regression analysis on IQ scores also provide some support that drawing topics or scenes with a greater emphasis on spatial composition might constitute a more promising method in the assessment of conceptual abilities.

At least on the basis of the properties of the drawing tests/tasks used in this thesis, it is fair to say that the content of the drawing measures doesn't affect the nature of the mental processes involved in the course of picture production for children of 4 and 7 year of age. For pre-school children, drawing performance is mainly determined by visual perception and motor-coordination skills. On the contrary, by the age of 7 both verbal and non-verbal abilities are equally significance for the level of competence in pictorial representation. Partial support is also provided from the performance of 8 year olds. Still, at this age, the spatial tasks – either due to their content or to the underlying processes – appear more appropriate to the changing nature of children's cognitive structures which have started to acquire more abstract properties. However, the performance of 5½ year olds appears irreconcilable with this pattern, outlining the differential nature of abilities involved through the course of drawing development.

5. OVERALL CONCLUSIONS

The assessment of drawing ability by means of spatial tasks appears as reliable – and in some respects even more promising – as the use of the human figure. Continuous and reliable developmental changes have been revealed by children's overall performance in the series of spatial tasks and in each domain-specific area. But, for children between 4 to 8 years of age, arrays involving pictorial overlap are more age-sensitive than those requiring the use of an invariant system of spatial coordinates. Performance on spatial tasks compares favorably with performance on GHDT, yet the magnitude of relationship is stronger at intermediate ages. The developmental separation in the emergence of the pertinent spatial concepts compared to the human form seems to account for the restricted nature of these relationships at the age of 4. The factors responsible for similar results at the age of 8 need further investigation. Possible areas of further research are to extend the present study to older children, to examine the studied relationship separately for the elaborate and the structural dimensions comprising the GH scales (see Strommen 1987 and Chapter III, section 3.2.3), or even to use human figure tasks specifically requiring structural differentiation of larger graphic units (figures in action or within a scene).

The results also show that the representational process tapped by the spatial tasks appears to be reasonably sensitive to level of intellectual functioning as indicated by children's IQ scores from conventional and non-verbal measures. Further, the effect of ability on the degree of view-specificity in the drawings of spatial arrays is uniform across the age groups studied here. However, correlational findings provide strong evidence that a drawing scale, regardless of its content, is more closely linked to performance IQ than to verbal IQ. This conclusion is tempered by the finding that the contribution of various mental skills to drawing performance changes with age. At the age of 4, perceptual-motor skills mainly determine drawing performance, whereas from the age of 7, verbal abilities are equally involved. The irreconcilable performance of 5½ year olds across various measures needs further investigation. Perhaps faulty sampling procedures can account for their results or other factors relating to school entrance. In conclusion, pictorial representation of space holds some promise as an alternative domain for measurement of children's specific and general abilities. The construction of a drawing scale sampling age-sensitive tasks from various spatial domains seems to reconcile the traditional and modern approach to the study of children's drawings.

CHAPTER XIII

GENERAL DISCUSSION AND CONCLUSIONS

In most areas of development, the formula is simple: youngsters get better, more skilled, more sophisticated with age. But such a unilinear portrait does not do the arts justice. In some way young children are especially intimate with the arts; and the story of artistic development is replete with declines, zigs and zags, rather than following an automatic upward progression
Howard Gardner, 1982, p. 83-84

...it is more difficult to establish regular stages of development in the case of artistic tendencies than it is in that of other mental functions.
Jean Piaget, 1953, p.22.

This thesis has investigated a number of issues relevant to drawing development with various drawing themes and different perspectives. The main objective was to examine the effectiveness of different drawing methods in the assessment of cognitive development and intellectual maturity of children between 4 to 8 years of age. At this point it will ease matters if we consider first the main findings from the areas of drawing competence studied through this thesis before general conclusions are drawn.

1. SUMMARIES OF MAIN RESULTS

1.1 GHDT

Regarding the test norms, the previously claimed upward bias in Harris standardisation norms was not upheld for children of 5½ to 8 years of age. Conversely, a clear conclusion from the present study is that Harris norms require revision for pre-school children and for girls, especially when they are assessed with the Man scale (MS). In line with previous results, the test is liable to underestimate the abilities of children with these characteristics. Therefore, before revising the norms, it is mandatory – more than in any other case – to administer the GHDT as part of a test battery, or to rely on raw scores when assessing the intellectual abilities of pre-school children. Likewise, it is equally essential to obtain two opposite-sex HFDs (Man/Woman or Man/Self) from girls in order to reliably evaluate their abilities. This latter recommendation is supported on many grounds; a finding which clearly demonstrates that the test has certain troublesome features for the assessment of girls.

First, in parallel with previous results, gender differences on raw mean scores in favour of girls are present for the WS only, and not for the MS. This finding doesn't completely justify Harris method of developing separate raw-to-standard score conversion tables for the sexes, suppressing girls' raw score in each scale. Further, the present study showed that these differences are consistently absent when the analysis is confined to two reduced sets of items, common to both scales (core and elaborate). Does this suggest that the differential performance of the sexes on WS is just an artifact of scale construction and not a sign of real differences in ability? Possibly girls might simply score more items peculiar to WS. Conversely, it might also be the case that the set of items particular to each scale measures operatively different abilities.

Careful item analysis of the two scales (MS/WS) reveals that both in the overall sample and at each age level, gender differences are not pronounced and only present on a limited number of items. Further, finding that gender-specific features don't appear any more frequently among girls than among boys rules out previous claims of girls' superiority in detailing and sexual differentiation of the HF. Therefore, at least on the basis of the present results, it appears that scale construction offers a more credible explanation for the obtained gender differences on WS. Yet, it is the author's firm belief that a definite answer to the previous question can be given only if future research adopts factor analytic techniques on

GH scales. This method will allow the investigation of the nature of abilities and the extent in which these are sampled by clusters of items common and particular to each scale.

The need for such methodological approach is further underscored by the examination of test's internal consistency. The present study showed that the three GH scales tap overlapping skills, but that these are not truly parallel. One source of evidence is provided by the magnitude of the inter-scale reliability coefficients (around .80). Partial support is also provided by the substantial scale differences in the mean scores, found mainly with girls. The fact that this pattern was consistent in all types of analyses, once more counts against the interchangeable use of the scales for girls. However, finding that the direction of the obtained scale differential varied with different methods of analysis (complete/reduced set of items; raw/standard scores) suggests that a number of factors can be held responsible.

Two of them have been already mentioned; one is related to the construction of the raw-to-standard scores conversion tables which creates scale differences in favour of the MS when standard scores are used, and difference in the opposite direction when raw scores are used instead. The latter can also result from the different features of the scales. However, the fact that scale differences were more pronounced when girls' scores from drawings of Man are compared with those of Self indicates that possibly having to draw two same-sex drawings in a consecutive order (woman → self) undermines the figural elaboration of the latter one. Therefore, another issue which needs to be addressed in future developmental research with the GHDT is the fixed order of testing. This aspect should be examined within gender and age level, as it might be auxiliary at younger ages (practice effect) and inhibitory at older ages (loss of interest).

On the positive side, the GHDT appears to be a stable and reliable instrument. All estimates of reliability are comfortably high, with inter-scale coefficients in the high 70s, inter-item coefficients in the high 80s and intra-scorer coefficients in the high 90s. Certain allowances need to be made though for the latter figure. Examination of scorer's consistency item by item revealed systematic discrepancies in scoring, despite the high consensus between any pair of composite scores from each child. This finding fully corroborates previous claims about certain item-specific ambiguities in the scoring criteria which render the nature of disagreement consistent. The present study has elucidated these problematic aspects and offered scoring suggestions which can be of help to researchers as well as to practitioners.

Another favourable feature of the instrument is its efficacy in discriminating children of 4 to 8 years of age. Not only there was a reliable increment of composite score points at each successive age group, but item analysis revealed that the developmental changes in HFDs go beyond the simple accretion of features. While at the age of 4 only the barest minimum of core features are present, at the age of 5½ a dramatic increase takes place both in the frequency of occurrence of items as well as in the appearance of novel features. A greater concern for realism is evident as children begin to score items from the elaborate and structural dimension of the scales. Although additions are fewer at older ages, children demonstrate a continuous concern for structural elaboration which is evident on larger (limbs, body) as well as smaller (fingers, neck, eyes) graphic units.

Still, the critical question is whether the GHDT is a valid instrument of intellectual functioning. Before answering this question, let us consider the main findings of this study. Regardless of the GH scale used, standardised drawing scores were found to be consistently and reliably lower than Wechsler IQ scores. Further, the obtained 10 points score differential between FSIQ and drawing scores is comparable with the prevalent view that an overall of

12 point discrepancy lies between the tests' scores. In addition, the present results lend support to previous claims that the magnitude of the obtained discrepancy varies considerably at different ages and levels of intelligence. The conclusion is that the GHDT tends to underestimate children at lower age levels (pre-school children) and in the upper ranges of intelligence (IQ>101).

Correlational evidence also suggests that the relationship between the drawing test and conventional instruments of intelligence is not quite firm. First, the validity coefficients from the total sample, albeit significant, are mediocre (mid-30s) and notably higher with the performance component rather than with the verbal component of the Wechsler scales. Yet, this conclusion needs to be tempered by noting that the magnitude of the examined relationship varies as a function of age. At the age of 4, more likely due to the upward bias in GH norms, performance in the drawing test doesn't relate to children's IQ, however there is some indication that the test reflects abilities of a non-verbal nature. On the contrary, at 5½ years of age the relationship is reliable, yet it is still primarily determined by performance skills. A different picture is obtained at 7 years of age as the test taps abilities of a non-verbal and verbal nature equally well, whereas the GHDT loses its validity as a test of intellectual maturity at the age of 8.

In view of these findings it becomes evident that an answer to questions about the test's validity is not a straightforward one. It is definitely a poor predictor of Wechsler IQ scores. So, on this ground, psychologists should not rely solely on it. On the other hand, overall correlations of the obtained magnitude don't appear any lower than those obtained from other brief instruments or subtests of more comprehensive batteries of intellectual functioning (Jensen, 1980). So its use as a screening measure or as an additional measure within the psycho-educational test battery is reasonably legitimate. Besides this general conclusion, one should argue about its usefulness on the basis of children's age. At present, it is a more valid instrument of intellectual functioning at intermediate stages of development.

A more general issue addressed in Chapter III is that the aforementioned features are not particular to the GHDT, but intrinsic to this method of assessment. A number of authors in defence of the utility of the drawing technique in intellectual evaluation, attribute its modest relationship with conventional intelligence tests to content differences. It is only natural that a drawing test, by using a non-verbal medium, should assess a narrower range of abilities than those measured by extensive batteries of intelligence. Theorists arguing along this line emphasize that instruments such as the Wechsler scales measure two types of intelligence, fluid and crystallized, whereas the drawing test is limited in terms of the latter (see Bardos, 1993). Others maintain that the content of a task/test is less critical than the processes involved. The latter view is characteristic of theorists who adopt an information-processing conception of intelligence and assert that most of non-verbal intelligence tests (including drawing tests) employ a narrow concept of intelligence. According to them they primarily assess simultaneous processes at the expense of planning, attention, and successive cognitive processes (see Naglieri & Das, 1990; Naglieri *et al.*, 1989).

Because the present study was directed towards certain practical issues, it assessed the validity of GHDT against the most commonly used measure of intelligence. Nevertheless, a fruitful area of future research is the assessment of its utility in relation to other brief measures of intellectual functioning or instruments which tap particular operations. To return to a previous suggestion, such investigations should adopt factor analytic methods. In this way we would be in a better position to reflect on the empirical basis of the test's structure before any serious attempt is made to revise it. Further, since much modern theorising about drawing development has an information-processing framework, we can justifiably argue that the drawing mode *per se* can provide the opportunity to assess a wide range of operations. When Harris revised the Goodenough DAM test, this potential wasn't fully appreciated, obviously affecting its content and purpose. Bearing in mind the test's continuous popularity in the psycho-educational setting as well as the promise that the drawing medium holds in cognitive assessment, it seems important to carry on work with GHDT in order to rectify its limitations and realise its potential.

1.2 Partial Occlusion

Besides HFDs, another way in which drawings change during early childhood involves the pictorial representation of spatial relationships in an array of objects. One facet of this problem has been how children represent pictorially the concealed properties of partially occluded objects. Despite the general consensus that the developmental shift takes the form of moving from object/array-centred drawings – where the occluded form is drawn as complete – to view-centred arrangements – where its hidden features are omitted –, there is a certain disagreement as to when this shift takes place and what factors influence this progression. These issues were addressed with a series of experimental tasks, where children had either to copy or draw from memory occluded scenes of two-item arrays.

Recent claims that the visually accurate conceptualisation and representation of spatial relations involving occlusion depend on objects' physical properties and arrays' structural features (e.g. Teske *et al.*, 1992) receive full support by the obtained between-task differential in the frequency of the view-specific response and in the occurrence of drawing strategies. Further, the variable pattern of developmental change across tasks and the fluctuation in the use of devices at each age level upheld recent suggestions that there isn't a perfect parity between children's age and depiction strategy. A synopsis of children's performance is as follows. The youngest of children have limited abilities to radically change their approach across tasks and the oldest have acquired procedural knowledge and well-established skills which are easily transferred to different situations. On the other hand, children at intermediate ages are sufficiently flexible to alter their representational priorities, indicating that the pertinent skills are available but not always accessible.

The fact that 4 year olds are least susceptible to task manipulations and most liable to represent the occluded forms as complete indicates that their spatial schemata don't incorporate information about the visual appearance of objects. By being trapped in object-centred descriptions, their performance is better explained along the lines of 'competence-deficiency' hypothesis, and in this sense they can be characterised as intellectual realists. Contrary, the rather stable performance of the 8 year olds, who used the view-specific device in the majority of cases, suggests that their concepts have been transformed to integrate spatial features which define forms and between-form relationships from a specific point of view. This is consistent with the Piagetian premise that the ability to perform mental transformations is symptomatic of operational thought which characterise children of this age. However, this doesn't guarantee every prediction about the performance of the 8 year olds under any condition. But, keeping in mind that in 6 out of 9 tasks children were presented with line drawing models, one can comfortably claim that, by the age of 8, children are reliably using the HLE device in copying tasks.

Turning to the performance of the 5½ and 7 year olds, all the evidence lends credit to competence-utilisation or performance-deficit hypotheses. Their variable drawing style suggests that they have the basic knowledge to produce visually realistic drawings, but that various factors facilitate or hinder children in recognising the need to give representational precedence to appearance rather than reality. And it is indeed at this age range that traditional stage theories have the least credibility. In view of this developmental picture, we can conclude that the progression from intellectual realism and object-centred drawings to that of visual realism and view-specific ones doesn't take place in an abrupt fashion. In agreement with current views (Arrowsmith *et al.*, 1994; Costall, 1995) these stages should be interpreted only as generalised descriptions and their accompanying phenomena simply as tendencies. In this frame, drawing development should be redefined on the basis of children's adaptive flexibility to representational demands.

Later approaches to children's drawings have also laid stress on the cross-sectional study of children's drawing devices as a means to gain a better understanding of the nature of developmental change. At the time when the pictorial representation of partial occlusion received most attention, Freeman *et al.*, (1977) proposed that children acquire the pertinent graphic device (HLE) through a developmental accretion of drawing rules. Present results refute this claim since there was no systematic age-related transition from one strategy to another that cuts across different contexts.

An unswerving impression from the studies is that the representation of the occluded form as complete and spatially separate from its occluder (segregation) is the only true depiction bias which undermines the pictorial representation of partial occlusion and the only device which reveals a reliable age-related trend. The pattern of its decline with age is largely dependent upon the properties of the scenes, as it is steeper and abrupt with easy tasks and smoother and gradual with the most difficult ones. Besides these age-related changes, as the scope for error increases in more demanding tasks, segregation output also increases with an ensuing decrease of other alternative strategies.

Subsequently, the explication of the forces guiding its appearance becomes a pressing issue. A number of researchers have suggested that segregation is dictated by young children's tendency to preserve the formal integrity of the occluded form (Light, 1985; Crook, 1985). If this conceptual bias or socially-driven attitude (informative intention) guides young children's representational attempts, we would expect all the alternative devices to appear with comparable frequency. The findings in Chapters V & VI show that this is not the case. It has been also suggested that segregation has a preferred status over other alternative devices because it is the only strategy which captures the internal structure of a scene. However, this premise originates from studies using arrays arranged in depth where there is projective but no true overlap between constituent elements. The studies in Chapter VI with arrays of structurally integrated items were designed to clarify this issue. The fact that spatial separation prevailed among the younger children – and even when no loss of the form's integrity was involved (Houses task) – provides the most compelling evidence against the latter supposition.

Interpreting these results in conjunction with cross-sectional data, we can conclude that segregation is an all-purpose drawing convention for the 4 year olds, symptomatic of their "synthetic incapacity" (Luquet, 1927/1977). Since it appears in children's early representational attempts, it is an over-learned mode which older children occasionally regress to whenever a depiction task requires greater cognitive processing and physical dexterity from them. Finding this formula in other drawing tasks (i.e. human figure), suggests that as the degree of mental computations increases so does the likelihood of primitive solutions (i.e. tadpole, armless figures).

This issue has direct implications for the optimum method of assessing children's abilities. For example, the present studies show that challenging tasks do not readily reveal developmental changes if one assess drawing competence only by the examination of children's graphic solutions or by counting successful responses. Since demanding themes eradicate children's competence to resist segregation at intermediate stages of development, a limited approach to judging performance could mask other developmental differences and fail to discriminate children younger than 8 years of age. A case which elucidates this point, is the Bus task. Here the spatio-temporal order of production revealed more age-related information than the conventional way of assessment. Therefore it seems desirable to adopt a comprehensive method of assessment, particularly when we obtain a relapse into primitive graphic devices from children between 5 to 7 years of age.

Returning to the incidence of graphic strategies, apart from segregation, their use is not age-related but instead is array-contingent[1]. A factor which has a direct impact on the use of drawing devices is the degree of spatial occlusion between the forms. The analysis of children's performance across tasks revealed that when the concealment is minimum (Bottles) overlap is substituted by attachment, when it is maximum, overlap is substituted by partial occlusion (Balls-FO), whereas in moderate cases attachment is substituted by overlap (Balls-MO, Apples, Sunset).

A number of authorities, mainly arguing from a cognitive approach, have explained how various structural aspects of a scene affect the way children will encode a given array (Crook, 1985; Cox 1986; Morra *et al.*, 1996). The present results allow us to adopt a similar stance and offer two suggestions in terms of how this factor operates. In general, the greater the spatial overlap between the forms the more likely it is that the occluded form will be incomplete and that its hidden features will not be depicted. This issue seems to be of special significance when children are tested with model arrays of similar forms, since increased visual contrast between items promotes view-specificity (Davis, 1983 & 1985a; Davis & Bentley, 1984). Thus, increasing the spatial overlap between similar items underscores their

[1] An allowance need to be made here with respect to enclosure configurations which, as they appeared mainly among the 4 year olds, are in a more limited sense age -dependent.

formal differences and prevents encoding by similarity. Children do not simply use one scheme (complete form) to represent the occluded and occluding item but two different ones (complete and incomplete). The results from the Balls tasks (MO/FO) fully support this argument.

At this point it is important to keep in mind that some authors have claimed that testing children with a similar-objects array constitutes a rather challenging (Cox, 1986) – not to say misleading (Morra *et al.*, 1996) – situation. Yet, research on partial occlusion is replete with studies following this paradigm. The comparative analysis of children's performance here shows that apart from the aforementioned factor, two more variables (either in isolation or in combination) appear to interfere with the accessibility of the appropriate strategy in similar-item arrays. One is whether the occluding form is pre-drawn on the children's paper and the other is whether they can distinguish the items by attributes other than relative completeness (e.g orientation in the Bottle task). Therefore, when we assess the acquisition of the HLE device with a model array of similar items, it is advisable first to present the array with maximum partial occlusion while at the same time preserving the similarity of the forms in full, and second to ask children to reproduce the whole array. However, the studies here were not designed to manipulate these variables directly. Their involvement appears in the course of drawing general conclusions from the present results and relating them to previous conjectures. Considering that these variables have not been directly investigated before, these findings emphasize the need to examine them in future with rigorously controlled studies.

Another issue addressed in the present studies was whether the use of copying tasks is itself a facilitating condition so far as HLE is concerned. The answer is definitely affirmative if we compare the results from the present tasks (Apples, Balls) with those of previous studies using exactly the same material, but with a 3D model or with a mnemonic task (Braine *et al.*, 1993; Chen & Holman, 1989; Cox, 1978; Freeman *et al.*, 1977; Klaue, 1992; Smith & Campbell, 1987). Clearly, the output of partial occlusion is substantially higher in copying tasks, particularly at intermediate ages. However, if we compare children's performance between copying and mnemonic tasks in the present series of studies, regardless of the material used, we obtain a variable pattern. It follows that neither working from memory or from a 3D model is always inhibitory, nor is copying form a line drawing always facilitating. The fact that an answer to the previous question cannot be given independently of the material used indicates that it is the formal features of the scene which codetermine the amount of work required from the child.

Similar conclusions are drawn when investigating previous claims that using 'ecologically valid' scenes fosters partial occlusion output by directing children's attention to the representational relevance of omitting hidden information (Cox, 1985 & 1993). Again, the value of scenes wherein the idea of hiding is intrinsically salient is moderated by production difficulties (Sunset and Vase tasks vs. Boat task). The comparative analysis of children's performance across the occlusion studies suggests that a third variable always interferes in promoting or confining the auspicious effect of 'ecological validity'; namely the level of graphic complexity of the requisite drawing. When the formal simplicity of the occluded form is preserved and the mental computations are reduced, mnemonic tasks (Vase) or arrays of a much lesser 'ecological value' (Balls)[2] can be equally or even more effective in promoting the use of the appropriate strategy.

Therefore, if we accept that not all objects are equally easy to draw as integral forms, subsequently we have to admit that neither are they equally easy (or difficult) to draw as incomplete forms. The elimination of hidden lines is not a discrete ability; the representation

[2] In the paradigm of 'one ball behind another' the issue of hidden-feature omission is not as intrinsically salient as in the case of a flower inside a vase.

of some occluded figures calls for advanced technical competence and organisational skills (i.e. Bottles, Boat tasks). Assessing the acquisition of the HLE device with such forms certainly imposes a cognitive overload on younger children and renders the pertinent device inaccessible. Present results show that this is the case for children at intermediate stages of development: under such conditions they regress to primitive solutions. It follows that if the point of assessment is to reveal children's grasp of the requisite spatial concept (aside from technical drawing competence), we should keep the material as structurally simple as possible.

The final note concerns two subject variables. Neither gender, nor cognitive style has any effect on performance in occlusion tasks. The findings for gender are consistent with those of the few studies which investigated this variable (Morra *et al.*, 1996; Radkey & Enns, 1987; Teske *et al.*, 1992). However, contrary to Morra's conclusion, FI was not associated with the use of the HLE device in a similar-item array. Possibly differences in the materials used to assess drawing competence (the dimensionality of arrays presented) and in the instruments employed for the assessment of cognitive style contributed to the irreconcilable results between the studies. In view of this, the association of cognitive style with performance on arrays with similar objects needs further investigation.

In conclusion, the studies reveal distinct developmental changes in the way children handle graphically the concealed properties of partially occluded objects. Children at the age of 8 begin to use reliably the view-specific strategy whereas at the age of 4 visually correct responses are rather sporadic. A number of experimental variables appear to have an auspicious or antagonistic effect on performance of children between 5½ and 7 years of age. The method adopted does not allow assessment of their individual impact, but this is hardly the case outside the experimental setting. From the pattern of developmental change it is evident that they affect differently children at different stages of development. Until more advanced methodological approaches allow us to investigate the way in which various factors operate in isolation and combination at different age levels, it is desirable to assess children with more than one task and to judge their performance with criteria other than the crude pass/fail.

1.3 Spatial axes

Similar conclusions are drawn from the studies of axial tasks. Drawing stage theories, and the Piagetian model in particular, is confirmed in some respect but discredited in others. When the concept of horizontality is assessed with WLTs, in agreement with Piaget, it is not found until the age of 8. The same applies to verticality, though only when the testing procedures resemble those in the traditional Piagetian variant (predictive-mnemonic task). However, overall results from the axial tasks first refute the Piagetian premise that the two reference systems guide performance and second call in question subsequent claims that the developmental change lies in the shift from being susceptible to a contextual and thus varied spatial frame to being adherent to the external and thus fixed frame of coordinate axes.

The analysis of the magnitude of linear inclinations shows that in a considerable number of drawings these are neither accurate (0°-11°) nor context driven (34°-50°), but either fall between these categories (12°-33°) or are quite deviant (>50°). The cross-sectional examination of error patterns also reveals that the transition from context-driven to accurate drawings is not an abrupt one, but that all types of deviant response exist at each age level. Thus, in parallel with the occlusion studies, developmental change is better conceptualised in terms of the relative frequency of these response patterns rather than their presence or absence. To complete the argument, recent research into axial graphic problems suggests that the fact that children often make compromises between the proximate/varied and distal/invariant frame indicates that they are caught in a conflict between the invariant axes and other forces acting in opposition (perceptual, conceptual, geometric). Differences in the magnitude of absolute and directional angular scores, as well as fluctuation in the frequency of inaccurate drawing solutions by age suggest that the developmental progression reflects children's growing ability to resist the combined or separate effects of these antagonistic factors. Before expanding on this issue the performance of children at different ages will be considered first.

Overall evaluation of the developmental data from WLTs is as follows. Competence deficit interpretations are readily applied to the performance of children up to the age of 5½ who appear to have limited understanding of the horizontality of the water level. In support of this conclusion is their inability to benefit from the provision of a pictorial example and the prevalence of incorrect drawings among them. Serious reservations need to be made though with respect to the assessment of preschool children. These children not only were totally impervious to the additional facilitative trial (two-vessel presentation), but also half of them produced scribbles or equivocal drawings and even had problems making the level of the water flat in upright containers. All these count against the use of WLTs for the assessment of horizontality in pre-school children as they appear to be beyond their basic drawing skills.

On the other hand, the performance of 7 and 8 year olds is best explained along the lines of a competence-utilisation hypothesis. First, half of them produced accurate drawings and their errors were mainly compromise solutions between variant and invariant axial systems (12°-33°). Second, their performance is markedly better compared to studies testing same-age children with 3D models wherein water was either present or absent. It follows that children of this age range do not totally lack the pertinent concept but certain antagonistic forces hamper its application in the drawing situation.

The developmental picture from the vertical tasks is generally similar to that from the WLTs, although here some ability in the use of the appropriate reference system is present from the age of 5½. These tasks appear more effective in demonstrating errors in Piaget's account. Children perform notably better than he predicts in tasks where they have simply to copy the target line from a 2D model (Mountain, Chimney tasks). These tasks increased the rate of success from the age of 5½ and reduced the magnitude of deviant strategies in younger children. This refutes the Piagetian claim that at stages where children presumably lack the necessary conceptual spatial scheme, they cannot benefit from testing conditions where the physical facts are made known to them (demonstration, copying tasks).

The most significant finding concerns another obstinate bias similar to the segregation strategy in occlusion studies; namely the perpendicular bias or young children's tendency to orient target lines at right angles to the baseline. The manipulation of task requirements aiming to curb this tendency was successful in markedly suppressing right angle solutions in the Chimney task from the age of 5½ and in the Mountain task from the age of 4. This amply demonstrates children's flexibility in the use of reference systems and in parallel with occlusion studies provides compelling evidence against the association of particular strategies with certain stages of development.

On the other hand, the performance of children in the Road task, which required a predictive response, is congruous with Piagetian claims. As with the WLTs, the pictorial representation of the vertical is not fully demonstrated here even at the age of 8. In particular, this task radically increases the output of context-driven drawings at all age ranges while the pattern of developmental change is more in terms of reduction in the magnitude of angular deviation rather than in the radical transformation of the axial rule used. The overall differences in performance within the vertical tasks and between the axial tasks, as well the variation in the extent of this differential with age, point to the following conclusions. When a decision about the use of a Euclidean spatial concept is made from children drawings, due attention should be paid to the attributes of the tasks employed to assess it, especially from the age of 5½ and above. As in the case of occlusion, more cognitively challenging situations disable the use of underlying competence and flatten developmental change by activating over-learned and immature drawing devices.

Although the similarity of developmental sequences in the acquisition of the vertical and horizontal is central to Piaget's account, some later studies suggest a developmental lag between these spatial concepts (Perner *et al.*, 1984; Thomas & Lohaus, 1993) with the horizontal concept delayed in expression. But this lag may be an artifact of the material used to assess the pertinent concepts. WLTs involve a mobile reference system wherein the target line is fully embedded. Neither of these two attributes are present in the vertical tasks used here, which possibly explains the fact that the majority of 8 year olds succeed in the copying tasks assessing verticality (Mountain, Chimney) but only half of them demonstrate accuracy in copying the WL.

The elucidation of the task variables which influence competence utilisation will be addressed first with respect to vertical tasks. As with partial occlusion studies, researchers have given due attention to certain stereotypic solutions or biases in order to explain the source of errors. Thus, some have attributed right angle solutions to a concept-driven tendency to represent the canonical attributes of forms or between-form relationships. These cognitive accounts typically claim that young children draw the water line perpendicular to the walls of tilted containers or make the chimneys and tress perpendicular to sloping baselines because these are the orientations frequently encountered in their familiar surroundings (Arnheim 1969; Pascual-Leone & Morra, 1991). If this line of argument was valid, then perpendicular drawings should be of comparable frequency across the vertical tasks or at least across the two copying tasks. But the significant fluctuation of perpendicular drawings between the tasks indicates that (a) if there is a cognitive component in this bias then its operative strength is determined by other factors (children's level of spatial awareness, dimensionality) and that (b) this bias is not a unitary phenomenon.

Piagetian tasks have been often criticised for having a misleading structure, since the obliquity of the proximal frame of reference is in conflict with the true environmental coordinates. One source of evidence comes from studies which manipulate the degree of misleadingness by eliminating all the straight lines from the structural axes of the proximal frame (Abravanel & Gingold, 1977; DeLisi *et al.*, 1995; Russell & Maxwell, 1980; Thomas & Jamison, 1975). The apparent easiness of the Mountain task compared to the Chimney and Road task lends support to this argument. It also demonstrates that the perpendicular error, apart from its conceptual origin, has a geometric one. Since an oblique context involves the production of acute (or obtuse) angles and since young children 'perpendicularize' angularity, the elimination of obliquity in the Mountain task suppresses right angle solutions even at the age of 4. This provides the most convincing piece of evidence that the locus of this bias is not conceptual – at least not entirely.

However, results inconsistent with this view were obtained in the WLTs. Children's comparable performance across the Bottle and the Bowl tasks discounts the presumed facilitating effect of non-rectangular vessels in the use of horizontal axis. A possible explanation offered is that such 'field' effects might have not been totally obliterated by the present material, since the elliptical shape of the bowl, albeit devoid of straight lines, has a discernible orientation and extension unlike the previously used circular containers. Although previous studies have demonstrated that the presence of any straight line on a vessel's shape attenuates the propitious effect resulting from the absence of obliqueness (Willemsen & Reynolds, 1973; Vasta *et al.*, 1994), the present results show that vessels with detectable orientation can equally compete with the use of the natural coordinates.

In summary, the developmental picture obtained from axial tasks is analogous to the one from occlusion tasks. The tasks show that in general children are more resourceful than stage theories would suggest. Still, pre-school children don't seem to possess the requisite spatial concepts, since the most facilitating conditions simply affect their choice of strategy without increasing accuracy. From this age and above the axial concepts are beginning to emerge but they aren't fully acquired even at the age of 8. As the number and the depth of processes required in a drawing task increase, these concepts become harder to employ by children between 5½ and 8 years of age. Thus, children of this age range can appreciate the relevance of the graphic rule when provided with a pictorial example. When disembedding skills and geometric forces are reduced children perform at ceiling from the age of 5½. When distracting contexts are added accurate responses are of average frequency from the age of 7 and well established at the age of 8, but if the aspect of movement is also added in these contexts (WLT), both 7 and 8 year olds perform at only moderate levels of accuracy. When the axial concepts are assessed with the least contextualised task, which requires children to generate a mental image, to independently devise a strategy and to implement it graphically in the sole presence of distracting cues, there is no real development change and view-specific drawings are meagre even at the age of 8.

1.4 Cognitive style

The issue of contextual misleadingness in traditional Piagetian tasks now brings us to the Field dependent/Field independent subject variable. Interpreting the overall results from EFT, we can claim that cognitive style has an effect in drawing performance. As it was expected performance on EFT correlated reliably with GHDT (.40) due to their common perceptual/spatial component which render them closer to the Performance than the Verbal component of conventional tests of intelligence. The involvement of Field independence in the representation of spatial relationship is also confirmed from the factorial analyses separately on the two spatial scales. Field independent children passed more occlusion tasks and produced less deviant inclinations in axial tasks. However certain qualifications need to be made at this point.

First, field articulation didn't affect the use of HLE device when the relationship was investigated on each occlusion task separately. This result was particularly unexpected for similar-objects arrays (Balls, Apples, Bottles) as Morra *et al.*, (1996) found that perceptual dissembedding skills are involved in arrays of similar objects due to their increased degree of misleadingness. It was suggested that differences in the dimensionality of the materials used could have contributed to these irreconcilable results. Morra *et al.* tested children with 3D material whereas children in the present studies were assessed with 2D scenes. Possibly mnemonic and 3D tasks are more effective in revealing cognitive style differences. If children possess the requisite graphic strategy, cognitive style would facilitate its application only in conditions which require image generation and more mental effort. This suggests that cognitive style may serve an activation function and that this function is obliterated in "simplified" tasks where performance functions are reduced. Support for this hypothesis comes from the axial tasks.

Since it has been suggested that the ability to decentre from the alignment cues of the proximal frame is at the heart of these tasks, variation in disembedding skills should influence performance. The present results corroborate this claim. Except for the Mountain task, there is a reliable positive relationship between axial tasks and cognitive style. This finding is consistent with task differences. Since the Mountain task yielded the highest rate of accuracy and its axes compete least with natural coordinates, we can reasonably claim that eliminating the degree of structural conflict between proximal and distal frames facilitates use of the correct spatial rule. Besides the correlational pattern, additional evidence for the involvement of disembedding skills in axial tasks comes from the analyses of children's overall performance by cognitive style status for the vertical and the horizontal tasks separately. These analyses demonstrate that FI children produce significantly smaller angular distortions than FD children.

However, certain reservations need to be made. Careful examination of these relationships separately for each task and at each age level suggests that the ability to overcome field effects is more of a 'utilization' variable. This implies that provided children have some understanding of the spatial rule and that the task is of moderate complexity, FI children are likely to perform better than FD ones when they are asked to apply this rule in a drawing task. In particular, the uniform absence of a significant relationship between measures of cognitive style and performance on axial tasks at the age of 4 supports previous claims that, due to the absence of a rule at young ages, FI is not a consistent factor in Piagetian tasks (Pascual-Leone & Morra, 1991). Further, there is some indication that axial tasks which impose a cognitive overload on children will obliterate cognitive style differences. In particular, although correlational evidence shows that the distracting context in the Chimney and Road tasks calls for disembedding skills, FI and FD children performed comparably across all types of analyses in the Road task. On the contrary, in the Chimney task the former group outperformed the latter when the analysis was confined to data obtained from the same baseline (left slope of roof).

Finally, with respect to gender differences, the present studies fully support previous claims that these differences are not well-established in early developmental periods (Lohaus *et al.*, 1996; Thomas & Lohaus, 1993; Thomas & Turner, 1991). Although in WLTs boys performed better than girls, this difference was significant only in the analysis of pass/fail responses, whereas in vertical tasks the differences weren't reliable. What seems to account for the absence of differential performance between the sexes in the axial tasks is the analogous absence of reliable differences in cognitive style. In parallel with the present results, a number of researchers have maintain that gender differences in disembedding abilities are not found prior to adolescence (Pascual-Leone, 1969; Witkin & Price-Williams, 1974; Wolf, 1971).

2. OVERALL CONCLUSIONS

The above review suggests that explaining drawing development is not a cut and dried matter. The present work echoes the views held by Gardner (1982) and Piaget (1953) in the opening quotations of the present chapter. The establishment of steadfast developmental trajectories falls short in the field of pictorial representation. Children are not resolutely anchored either in the stage of intellectual realism or in that of visual realism. The present results underscore the need to adopt a more flexible notion of 'stage' as a landmark which involves a range of abilities which yield view-specific drawings with variable degree of consistency as well as alternative solutions of a different scale of credibility. When the stage theories were developed, inferences about children's representational competence were mainly drawn from their spontaneous drawings. Under these conditions, the absence of view-specificity was interpreted as stemming from children's limitations to conceptualise a scene from a fixed point of view. On the contrary, current experimental drawing research is primarily based on copying tasks (2D or 3D).

Although some have criticised this methodological approach as unnatural for young children, since they rarely draw by looking at a model (e.g. Costall, 1995), in fact these procedures make us realise that young children are not totally impervious to the view-specific properties of models before them. Children have some awareness that scenes are relative to a point of view before the previously claimed age threshold of 8 years (Piaget and Inhelder, 1956) and sometimes they can give representational precedence to their fleeting visual appearance. Intrinsically, the experimental context along with the standard request "draw <u>this</u> exactly as it looks from where you are sitting" renders the task at hand more of a social act. As was maintained at the introduction of this thesis, drawing – like language (or better speech) – serves a communicative purpose which changes as a function of age but also as a function of the setting in which it is embedded (see Van Sommers, 1984, chapter 11). Besides all other considerations regarding task demands, any experimental situation emphasises the need to produce an effective pictorial message in using the strategy of visual realism.

Research on communication abilities as well as recent investigation in the area of pictorial reasoning suggest that children before the age of 7 don't know what constitutes an adequate representation. Robinson (Robinson & Robinson, 1981), addressing this issue in the context of speech, refers to this ability as "metacommunicative knowledge" and claims that although children up to the age of 6 have well-developed verbal skills, they are limited in their understanding of what renders a message comprehensible by an audience. Similarly, Korzenik (1975) demonstrated empirically that prior to the first grade children mainly blame the beholder for communication failures caused by their ambiguous drawings, and are less likely to change them, believing that their productions are self-evident (see also Golomb, 1992). However, a dramatic conceptual shift in children's course of reasoning takes place by the age of 6. Similar views are presented by current attempts to link pictorial representation with the development of a theory of mind. Freeman (1995) suggests that before the age of 7, children don't have a grasp of how their picture appears to a beholder and cannot comprehend that ambiguous pictures allow a range of different interpretations.

What all this implies is that visual realism is also a result of the child's growing awareness of the needs and 'mind' of a viewer. Children's increased sensitivity with age to the experimental manipulation of the context and the content (instructions), and the variability between their 'private' drawings and their 'public' drawings made under request, suggest the following conclusion. Just as the characterisation of intellectual (or visual) realism is relative rather than absolute so are children's abilities in social cognition (see also Light, 1985). The increased emphasis on the social context of child art in recent times has opened exciting and fruitful lines of research. Keeping in mind that drawing performance is multidetermined, perhaps further study in the domain of pictorial reasoning will shed light on the course of development of pictorial representation.

Besides this theoretical issue, the overall evaluation of children's performance on experimental tasks provides clear evidence that with age children become more able to extend and use their knowledge in more demanding situations. Children's early drawing development seems to consist more of the acquisition of discrete graphic formulae for specific scenes rather than of the mastery of generic drawings skills which can be deployed for drawing a range of pictures falling under a presumably unitary topic; for example partial occlusion or spatial axes. Van Sommers discussing the complexity of picture making advised us that

> *Just because we come up with a single product in the end, we should not be deceived into thinking that that product has its origins in a single intact image, or even a single concept (1995, p. 54)*

Ample evidence, especially from children between 5½ to 7 years of age, suggests that the operative spatial schemata might already exist but that they are hard to employ when a number of other operations are required for their graphic implementation. Modern information-processing theories elucidate the plethora of processes involve in picture making. For instance, in the absence of a model, children need to generate a figurative scheme, to transform it by applying an operative scheme and to mediate relevant procedural schemata from previous drawing attempts. Or otherwise they must simultaneously attend to the visual attributes of a model while monitoring the course of production.

In fact there is growing body of research investigating the role of children's mental processing abilities in performance on spatial graphic tasks (Dennis 1992; Morra, 1995). Based on the theoretical formulations of neo-Piagetian frameworks, these researchers claim that the limitations of children's attentional capacity create a bottleneck and that this developmental variable will determine the magnitude of overall success across tasks of various structural complexity. These studies yield convincing evidence for a reliable relationship between age, mental processing abilities and spatial problem solving. Further, the relationship between mental attention and drawing performance remains reliable after age is partialled out. Additional support comes also from studies, which investigated the variable of attention with various occlusion tasks and WLTs (Morra *et al.*, 1996; Pascual Leone & Morra, 1991). Even if in principle each of these spatial areas involves a single conceptual scheme, nevertheless performance on these tasks is determined by the structure of the models along with children's age.

It has been outwith the remit of the current thesis to compare the rate of success on spatial tasks with attentional capacity, etc. Yet, based on the general impression from the present results it seems that this variable holds some promise in explaining the magnitude of age-related variability in performance as well as some variance in drawing competence within age. Thus, an obvious step to take in future research is systematically investigating its validity in accounting for the changes which take place in the domain of pictorial representation.

A final comment needs to be made with respect to the relationship between intelligence and drawing performance. The results from Chapter XIII suggest that the ability to solve pictorially spatial problems is an indication of intellectual maturity to the same extent as it is the ability to draw a much general topic like the human figure. Further, performance on the non-verbal test of intelligence (GHDT) as well as on conventional measures (Wechsler scales) is related to performance on spatial tasks, since the more intellectually competent children outperform the less competent. However, judging from the moderate strength of these relationships, it is evident that intellectual maturity accounts only partially for drawing competence. In addition, solving pictorial problems of a spatial and more general nature (HFDs) does not implicate equally all the domains of intelligence. In general, performance abilities are more heavily influenced in drawing tasks/tests, whereas verbal abilities operate only selectively in certain tasks and in certain age groups.

Earlier conjectures (Harris, 1963) that the GHDT might tap different abilities at different ages receive full support from the present study on spatial tasks. Clearly at the age of 4 drawing competence is mainly influenced by non-verbal skills independently of the drawing technique/theme used. On the contrary, at the age of 7, verbal and non-verbal skills equally affect pictorial representation, again regardless of the nature of the drawing area studied. An inconsistent picture is obtained at the ages of 5½ and 8. Here, the content of the drawing measure seems to determine the nature and the extent to which intellectual abilities are involved in drawing performance. Thus, for the 8 year olds the spatial tasks appear more intellectually challenging compared to the GHDT whereas the opposite picture is obtained from the 5½ year olds.

Another piece of evidence which validates the notion that the abilities the GHDT measures vary as a function of age comes from its variable relationship with spatial task performance at different age levels. The results in Chapter XIII suggest that different domains of pictorial representation demonstrate different degree of drawing competence at various ages. This developmental lag is also present within tasks of a purely spatial nature. For example in parallel with the present results, Pemberton (1990) found that axial tasks remained the most difficult even up to the age of 10, corroborating the Piagetian claim that Euclidean concepts are the last to be acquired. However, returning to the initial argument, drawing competence on the GHDT is more closely related to performance on spatial tasks at intermediate ages whereas no reliable relationship between the two types of drawing scores is present at the age of 8. Two explanations were offered for the later result. Either the human figure as a drawing topic is of limited efficacy in capturing the growing spatial skills of older children or the GH scales do not adequately sample items relating to structural aspects of figural elaboration.

The practical implication of these results is the following. Since overall performance on spatial tasks is largely dependent on children's age and reasonably sensitive to intellectual functioning, this justifies Piaget in placing the study of children's drawings in the domain of their growing understanding of space. Research in this domain appears as a more powerful framework to describe the multifaceted changes in drawing development. There are some encouraging results from the present study indicating that the pictorial representation of space can provide an alternative method for the assessment of cognitive abilities. Perhaps devising a drawing scale sampling tasks of a diverse spatial nature and complexity we will be able to strengthen and extend the validity of intellectual assessment through drawing tasks beyond middle childhood.

In conclusion, the discovery and the development of pictorial world by the child remain an area of immense consequence for his/her cognitive growth. The realisation that there are many theoretical and empirical lines of enquiry to pursue is by no means discouraging. Some propositions are put forward by the present study and future research will show which are more fertile. Meanwhile it is necessary to look at the process and products of children's drawing activity with an informed and sensitive eye bearing in mind that

> *"The process of drawing...is a complex one in which the child brings together diverse elements of his environment to make a meaningful whole. In the process of selecting, interpreting and reforming these elements, he has given us more than a picture, he has given us a part of himself".*
>
> *Lowenfeld (1947, p. 1)*

REFERENCES

Abell, S. C., Heiberger, A. M. & Johnson, J. E. (1994). Cognitive evaluation of young adults using human figure drawings: An empirical evaluation of two methods. *Journal of Clinical Psychology*, **50**, 900-905.

Abell, S. C., Horkheimer, R. & Nguyen, S. E. (1998). Intellectual evaluations of adolescents via human figure drawings: An empirical comparison of two methods. *Journal of Clinical Psychology*, **54(6)**, 811-815.

Abell, S. C., Von Briesen, P. D. & Watz, L. S. (1996). Intellectual evaluations of children using human figure drawings: An empirical investigation of two methods. *Journal of Clinical Psychology*, **52(1)**, 67-74.

Ables, B. S. (1971). The use of the Draw-A-Man Test with borderline retarded children without pronounced pathology. *Journal of Clinical Psychology*, **27**, 262-263.

Abravanel, E. & Gingold, H. (1977). Perceiving and representing orientation: effects of the spatial framework. *Merrill-Palmer Quarterly*, **23**, 265-278.

Adams, J., & Lieb, J. J. (1973). Canter-BIP and Draw-A-Person Test performance of Negro and Caucasian Head-start children. *Psychology in the Schools*, **10**, 299-304.

Adler, P. (1970). Figure drawing techniques. *Journal of Consulting and Clinical Psychology*, **35**, 52-57.

Aikman, K. G., Belter, R. W., & Finch, A. J. (1992). Human figure drawings: Validity in assessing intellectual level and academic achievement. *Journal of Clinical Psychology*, **48(1)**, 114-120.

Amponsah, B. & Krekling, S. (1994). Directional biases in adult's performance on a multiple-choice water-level task. *Perceptual and Motor Skills*, **79**, 899-903.

Anastasi, A. (1990). *Psychological testing*, 6th ed. New York: Macmillan Publishing Company.

Ansbacher, H. L. (1952). The Goodenough Draw-A-Man Test and primary mental abilities. *Journal of Consulting Psychology*, **16**, 176-180.

Arnheim, R. (1969). *Visual thinking*. Berkley: University of California Press.

Arnheim, R. (1974). *Art and visual perception: A psychology of the creative eye. The new version*. Berkeley & Los Angeles: University of California Press.

Arnheim, R. (1980). The puzzle of Nadia's drawings, *The Arts in Psychotherapy*, **7**, 79-85.

Arrowsmith, C. J., Cox, M. V. & Eames, K. (1994). Eliciting partial occlusions in the drawing of 4- and 5-year-olds. *British Journal of Developmental Psychology*, **12(4)**, 577-584.

Ashton, A. (1997). *Do hiding scenarios encourage young children to use partial occlusion to depict identical objects one behind the other?* Poster presented at the BPS Annual Meeting, Developmental Section, Loughborough, September.

Ayres, A. J. & Reid, W. (1966). The self-drawing as an expression of perceptual-motor dysfunction. *Cortex*, **2**, 154-265.

Bardos, A. N. (1993). Human figure drawings: Abusing the abused. *School Psychology Quarterly*, **8(3)**, 177-181.

Bardos, A. N., Softas, B. C. & Petrogiannis, K. (1989). Comparison of the Goodenough-Harris and Naglieri's Draw-a-Person scoring systems for Greek children. *School Psychology International*, **10(3)**, 205-209.

Barnes, E. (1897). *Studies in Education, Vol. 1*. California: University of Stanford Press.

Barrett, M. D., Beaumont, A. V. & Jennett, M. S. (1985). Some children sometimes do what they have told to do: task demands and verbal instructions in children's drawings. In N. H. Freeman & M. V. Cox (eds.), *Visual order: The nature and development of pictorial representation*. Cambridge: University of Cambridge Press.

Barrett, M. D. & Bridson, A. (1983). The effect of instructions upon children's drawings. *British Journal of Developmental Psychology*, **1**, 175-178.

Barrett, M. D., Sunderland, A. & Lee, F. (1988). *The effect of instructions upon children's drawings of scenes containing occluded objects*. Poster presented at IIIrd European Conference of Developmental Psychology, Budapest, July.

Barsky, R. D. & Lachman, M. E. (1986). Understanding of horizontality in college women: Effects of two training procedures. *International Journal of Behavioral Development*, **9**, 31-43.

Bayraktar, R. (1985). Cross-cultural analysis of drawing errors. In N. H. Freeman & M. V. Cox (eds.), *Visual Order: The nature and development of pictorial representation*. Cambridge: Cambridge University Press.

Beal, C. R. & Arnold, D. S. (1990). The effect of instruction on view-specificity in young children's drawing and picture selection. *British Journal of Developmental Psychology*, **8**, 393-399.

Beard, R. M. (1964). Further studies in concept development. *Educational Review*, **17**, 41-58.

Beery, K. E. (1967). *Developmental Test of Visual–Motor Integration: administration and scoring manual*. Chicago: Follett.

Beilin, H. (1989). Piagetian theory. In R. Vast (ed.), *Annals of child development*, **6**, 85-131. Greenwich, CT: JAI Press, Inc.

Beilin, H., Kagan, J. & Rabinowitz, R. (1966). Effects of verbal and perceptual training on water level presentation. *Child Development*, **37**, 317-329.

Bender, L. (1938). A visual motor Gestalt test and its clinical use. *New York American Orthopsychiatric Association Research Monograph*, **No 3**.

Bender, L. & Wolfson, W. Q. (1943). The nautical theme in the art and fantasy of children. *American Journal of Orthopsychiatry*, **13**, 462-467.

Benton, A. L. (1974). *Revised Visual Retention Test: Manual*. San Antonio, TX: Psychological Corporation.

Beyer, F. S. & Nodine, C. F. (1985). Familiarity influences how children draw what they see. *Visual Arts Research*, **11(2, Issue 22)**, 60-68.

Braine, L. G., Schauble, L., Kugelmass, S. & Winter, A. (1993). Representation of depth by children: Spatial strategies and lateral biases. *Developmental Psychology*, **29 (3)**, 466-479.

Bremner, J. G. (1984). Errors towards the perpendicular in children's copies of angular figures: A test of the bisection interpretation. *Perception*, **13**, 117-128.

Bremner, J. G. (1985a). Provoked use of height in picture as a depth cue in young children's drawings. *British Journal of Developmental Psychology*, **3(1)**, 95-98.

Bremner, J. G. (1985b). Figural biases and young children's drawings. In N. H. Freeman & M.V. Cox (eds.), *Visual Order: The nature and development of pictorial representation.* Cambridge: University of Cambridge Press.

Bremner, J. G. & Andreasen, (1997). A change in viewpoint promotes use of height in picture as a depth cue in 5 to 7 year olds' drawings of a simple depth relationship. *Journal of Experimental Child Psychology*, **66**, 180-192.

Bremner, J. G. & Batten, A. (1991). Sensitivity to viewpoint in children's drawings of objects and relations between objects. *Journal of Experimental Child Psychology*, **52(3)**, 375-394.

Bremner, J. G. & Morre, S. (1984). Prior visual inspection and object naming: Two factors that enhance hidden feature inclusion in young children's drawings. *British Journal of Developmental Psychology*, **2(4)**, 371-376.

Bremner, J. G. & Taylor, A. J. (1982). Children's errors in copying angles: Perpendicular error or bisection error? *Perception*, **11**, 163-171.

Brown, E. V. (1977). Reliability of children's drawings in the Goodenough-Harris "Draw-A-Man Test". *Perceptual and Motor Skills*, **44**, 739-742.

Brown, E. V. (1990). Developmental characteristics of figure drawings made by boys and girls ages five through eleven. *Perceptual and Motor Skills*, **70**, 279-288.

Bryant, P. E. (1974). *Perception and understanding in children: An experimental approach.* London: Methuen.

Buck, J. N. (1948). The HTP test. *Journal of Clinical Psychology*, **4**, 151-159.

Bühler, K. (1930). *The mental development of the child.* London: Kegan Paul.

Burns, C. J. & Velicer, W. F. (1977). Art instruction and the Goodenough- Harris drawing test in fifth-graders. *Psychology in the Schools*, **14**, 109-112.

Burns, R. C. & Kaufman, S. H. (1970). *Kinetic Family Drawings.* New York: Brunner/Mazel.

Buros, O. K. (ed.) (1953). *Fourth mental measurements yearbook.* Highland Park, N.J.: Gryphon Press.

Burt C. (1921). *Mental and scholastic tests.* London: King and Son.

Burton, E. (1995). Representing representation: Making a computer draw like a child: An MA Computing in Design Research project. *Computer Graphics Education*, **29**, 28-32.

Carvajal, H., McVey, S., Sellers, T., Weyand, K. & McKnab, P. (1987). Relationships between the scores of the General Purpose Abbreviated Battery of the Staford-Binet IV, Peabody Picture Vocabulary Test-Revised, Columbia Mental Maturity Scale, and Goodenough-Harris Test. *Psychological Records*, **37(1)**, 127-130.

Chen, M. J. (1985). Young children's representational drawings of solid objects: a comparison of drawing and copying. In N. H. Freeman & M. V. Cox (eds.), *Visual order: The nature and development of pictorial representation.* Cambridge: University of Cambridge Press.

Chen, M. J. & Cook, M. (1984). Representational drawings of solid objects by young children. *Perception*, **13**, 377-385.

Chen, M. J. & Holman, J. (1989). Emergence of drawing devices for total and partial occlusion: A longitudinal study. *Perception*, **18(4)**, 445-455.

Claparède, E. (1907). Plan d' experiences collectives sur le dessin des enfants. *Archives de Psychologie*, **6**, 276-278.

Clark, A. B. (1897). The child's attitude towards perspective problems. In E. Barnes (ed.), *Studies in Education*, **Vol. I**, Stanford: University of Stanford Press.

Coates, S. (1972). *The Preschool Embedded-Figure Test-PEFT.* California: Consulting Psychologists Press, Palo Alto.

Cooke, E. (1886). Our art teaching and child nature. *London Journal of Education*, January, 12-15.

Costall, A. (1985). How meaning covers the traces. In N. H. Freeman & M. V. Cox (eds.), *Visual Order: The nature and development of pictorial representation.* Cambridge: University of Cambridge Press.

Costall, A. (1995). The myth of the sensory core: the traditional versus the ecological approach to children's drawings. In C. Lange-Küttner & G. V. Thomas (eds.), *Drawing and looking: Theoretical approaches to pictorial representation in children.* Hertfordshire: Harvester Wheatsheaf.

Costall, A. (1997). Innocence and corruption: conflicting images of child art. *Human Development*, **40**, 133-144.

Cox, M. V. (1978). Spatial depth relationships in young children's drawings. *Journal of Experimental Child Psychology*, **26**, 551-554.

Cox, M. V. (1981). One thing behind another: Problems of representation in children's drawings. *Educational Psychology*, **1(4)**, 275-287.

Cox, M. V. (1985). One object behind another: Young children's use of array-specific or view-specific representations. In N. H. Freeman & M. V. Cox (eds.) *Visual Order:*

The nature and development of pictorial representation. Cambridge: Cambridge University Press.

Cox, M. V. (1986). *The child's point of view: The development of cognition and language.* Brighton, Sussex: Harvest Press.

Cox, M. V. (1992). *Children's Drawings.* Harmondsworth: Penguin books.

Cox, M. V. (1993). *Children's drawings of the human figure.* Hove: Lawrence Erlbaum Associated Ltd.

Coyle, R. T., Clance, P. R. & Joesting, J. (1977). Kinesthetic enrichment and Goodenough-Harris Draw-A-Man scores of black children from lower socioeconomic backgrounds. *Perceptual and Motor Skills*, **45**, 201-202.

Croake, J. W., Keller, J. F. & Caltin, N. (1973). WPPSI, Rutgers, Goodenough, Goodenough-Harris IQ's for lower socioeconomic black, preschool children. *Psychology*, **10**, 58-65.

Crook, C. (1984). Factors influencing the use of transparency in children's drawing. *British Journal of Developmental Psychology*, **2(3)**, 213-221.

Crook, C. (1985). Knowledge and appearance. In N. H. Freeman & M. V. Cox (eds.), *Visual order: The nature and development of pictorial representation.* Cambridge: University of Cambridge Press.

Cundick, B. P. (1970). Measures of intelligence on southwest Indian students. *Journal of Social Psychology*, **81**, 151-156.

Datta, L. E. (1967). Draw-A-Person Test as a measure of intelligence in pre-school children from very low income families. *Journal of Consulting Psychology*, **31**, 626-630.

Davis, A. M. (1983). Contextual sensitivity in young children's drawings. *Journal of Experimental Child Psychology*, **35**, 478-486.

Davis, A. M. (1984). Noncanonical orientation without occlusion: Children's drawings of transparent objects. *Journal of Experimental Child Psychology*, **37(3)**, 451-462.

Davis, A. M. (1985a). The canonical bias: young children's drawings of familiar objects. In N. H. Freeman & M. V. Cox (eds.), *Visual Order: The nature and development of pictorial representation*, (pp. 229-245). Cambridge: University of Cambridge Press.

Davis, A. M. (1985b). Conflict between canonicality and array-specificity in young children's drawings. *British Journal of Developmental Psychology*, 3(4), 363-372.

Davis, A. M. & Bentley, M. (1984). Young children's interpretation of the task demands in a simple experimental situation: An example from drawings. *Educational Psychology*, 4(3), 249-254.

De Goes, C. & Martlew, M. (1983). Young children's approach to literacy. In M. Martlew (ed.), *The psychology of written language*, Chichester: John Wiley.

DeLisi, R. (1983). Developmental and individual differences in children's representation of horizontal coordinate. *Merrill-Palmer Quarterly*, 29, 179-196.

De Lisi, R., Locker, R. & Youniss, J. (1976). Anticipatory imagery and spatial operations. *Developmental Psychology*, 12, 298-310.

DeLisi, R., McGillicuddy-DeLisi, A. V. & Mowrer-Popiel, E. (1995). Measuring grade-school children's ability to represent the Euclidean horizontal coordinate. *Journal of Genetic Psychology*, 156(2), 231-240.

Dennis, S. (1992). Stage and structure in the development of children's spatial representations. In R. Case (ed.), *The mind's staircase*, 229-245. Hillsdale, NJ: Erlbaum.

Dennis, W. (1942). Performance of Hopi Indian children on the Goodenough Draw-A-Man Test. *Journal of Comparative Psychology*, 34, 341-348.

Deregowski, J. B. & Strang, P. (1986). On the drawing of a cube and its derivatives. *British Journal of Developmental Psychology*, 4, 323-330.

Deregowski, J. B., Parker, D. M. & Dziurawiec, S. (1996). The role of typical contours in object processing by children. *British Journal of Developmental Psychology*, 14, 425-440.

Desai, K. (1958). *Developmental stages in drawing a cow.* Unpublished B.A. thesis, Baroda University.

Donaldson, M. (1978). *Children's Minds*. London: Fontana.

Donaldson, M., Grieve, R. & Pratt, F. R. (eds.) (1983). *Early Childhood Development and Education*. Oxford: Blackwell.

Du Bois, M. B. (1939). A test standardized on Pueblo Indian children. *Psychological Bulletin*, 36, 523.

Dunn, J. A. (1967a). Inter-and intra-rater reliability of the new Harris-Goodenough Draw-A-Man Test. *Perceptual and Motor Skills*, **24**, 269-270.

Dunn, J. A. (1967b). Note on the relation of Harris Draw-A-Woman to WISC IQs. *Perceptual and Motor Skills*, **24**, 316.

Dunn J. A. (1967c). Validity coefficients for the new Harris-Goodenough Draw-A-Man Test. *Perceptual and Motor Skills*, **24**, 299-301.

Dziurawiec, S. & Deregowski, J. B. (1992). "Twisted perspective" in young children's drawings. *British Journal of Developmental Psychology*, **10(1)**, 35-49.

Elliot, C. (1983). *The British Ability Scales*. Windsor: NFER.

Elliott, J. & Connolly, K. J. (1978). The effect of visual frame of reference on a judgement of plane stimulus orientation by children. *Perception*, **7**, 139-149.

Eng, H. (1931). *The psychology of children's drawings*. New York: Harcourt, Brace.

Evans, R., Ferguson, N. Davies, P. & Williams, P. (1975). Reliability of the "Draw-a-man" Test. *Educational Research*, **18**, 32-36.

Fabian, V. (1982). *Language development after 5: A neo-Piagetian investigation of subordinate conjunctions*. Unpublished doctoral dissertation, University of California, Berkeley.

Fabry, J. J. & Bertinetti, J. F. (1990). A construct validation study of the Human Figure Drawing Test. *Perceptual-and-Motor-Skills*, **70(2)**, 465-466.

Fairweather, H. (1976). Sex differences in cognition. *Cognition*, **4**, 231-280.

Farylo, B, Paludi, M. A. (1985). Research with Draw-A-Person Test; conceptual and methodological issues. *The Journal of Psychology*, **119(6)**, 575-580.

Fayol, M., Barrouillet, P. & Chevrot, C. (1995). Judgement and production of drawings by 3-to-10-year-olds: Comparison of declarative and procedural drawing knowledge, *European Journal of Psychology of Education*, **10(3)**, 303-313.

Ferreiro, E. (1985). Literacy development: A psychogenic perspective. In D. R. Olson, N. Torrance & A. Hildyard (eds.), *Literacy, language and learning: The nature and consequences of reading and writing*. Cambridge: University of Cambridge Press.

Fine, M. J. & Tracy, D. B. (1968). Performance of normal and EMR boys on the FRPV and GHDT. *American Journal of Mental Deficiency*, **72**, 648-652.

Foorman, B. R (1981). A neo-Piagetian analysis of four-year-olds' performance on Piaget's water-level task. *Perceptual and Motor Skills*, **52**, 631-639.

Ford, L. H. (1970). Predictive versus perceptual responses to Piaget's water-line task and their relation to distance conservation. *Child Development*, **41**, 193-200.

Ford, M. A., Stern, D. N. & Dillon, D. J. (1974). Performance of children ages 3 to 5 on the Draw-A-Person Task: Sex differences. *Perceptual and Motor Skills*, **38**, 1188.

Freeman, N. H. (1972). Process and product in children's drawing. *Perception*, **1**, 123-140.

Freeman, N. H. (1976). Children's drawings: Cognitive aspects. *Journal of Child Psychology and Psychiatry*, **17**, 345-350.

Freeman, N. H. (1980). *Strategies of representation in young children: Analysis of spatial skills and drawing process*. London: Academic Press.

Freeman, N. H. (1983). Picture-plane bias in children's representational drawings. *Australian Journal of Psychology*, **35(2)**, 121-134.

Freeman, N. H. (1995). The emergence of a framework theory of pictorial reasoning. In C. Lange-Küttner & C. Thomas (eds.), *Drawing and looking: Theoretical approaches to pictorial representation in children*. Hertfordshire: Harvester Wheatsheaf.

Freeman, N. H. & Cox, M. V. (eds) (1985). *Visual Order: The nature and development of pictorial representation*. Cambridge: University of Cambridge Press.

Freeman, N. H., Eiser, C. & Sayers, J. (1977). Children's strategies in producing three-dimensional relationships on a two-dimensional surface. *Journal of Experimental Child Psychology*, **23**, 305-314.

Freeman, N. H. & Hargreaves, S. (1977). Directed movements and the body-proportion effect in pre-school children's human figure drawings. *Quarterly Journal of Experimental Psychology*, **29**, 227-235.

Freeman, N. H. & Janikoun, R. (1972). Intellectual realism in children's drawing s of a familiar object with distinctive features. *Child Development*, **43**, 1116-1121.

Freeman, N. H. & Kelham, S. E. (1981). The interference of a baseline with children's orientation judgement need not entirely mask their contextual responsiveness. *Quarterly Journal of Experimental Psychology*, **33A**, 145-154.

Gardner, H. (1980). *Artful scribbles: The significance of children's drawings*. New York: Basic Books.

Gardner, H. (1982). *Art, mind and brain*. New York: Basic Books.

Gayton, W. F., Evans, H. & Wilson, W. (1971). Comparative validity of Harris Point and Quality scales. *Perceptual and Motor Skills*, **33**, 1111-1113.

Gayton, W. F., Tavormina, J., Evans, H. E. & Schuh, J. (1974). Comparative validity of Harris's and Koppitz' scoring systems for human-figure drawings. *Perceptual and Motor Skills*, **39**, 369-370.

Georgas, J. & Georgas, C. (1972). A children's intelligence test for Greece. *Mental Tests and Cultural Adaptation*, 219-220.

Gibson, E. (1969). *Principles of perceptual learning and development.* New York: Appleton-Century-Crofts.

Gibson, J. J. (1979). *The ecological approach to visual perception.* Boston: Houghton Mifflin.

Glenn, S. M., Bradshaw, K. & Sharp, M. (1995). Handedness and the development of direction and sequencing in children's drawings of people. *Educational Psychology,* **15(1)**, 11-21.

Goldman, R. K. & Velasco, M. M. (1980). Towards the development of a rational scale in the use of human-figure drawings as a kindergarten screening measure. *Perceptual and Motor Skills*, **50**, 571-577.

Golomb, C. (1983). On imaginary or real décalage in children's representations: Compositional trends in drawing, completion, and selection tasks. *Visual Arts Research,* **9(2)**, 71-81.

Golomb, C. (1987). The development of compositional strategies in children's drawings. *Visual Arts Research*, **13(2)**, 42-52.

Golomb, C. (1992). *The child's creation of a pictorial world.* Berkeley: University of California Press.

Golomb, C. & Farmer, D. (1983). Children's graphic planning strategies and early principles of spatial organization in drawing. *Studies in Art Education*, **24(2)**, 86-100.

Gombrich, E. H. (1972). *Art and illusion: A study in the Psychology of Pictorial Representation*, 4th ed. London: Phaidon Press.

Gombrich, E. H. (1982). *The image and the eye.* Oxford: Phaidon Press.

Goodenough, D. R., Oltman, P. K. & Cox, P. W. (1987). The nature of individual differences in field dependence. *Journal of Research in Personality*, **21**, 81-90.

Goodenough, F. L (1926). *Measurement of Intelligence by Drawings*. New York: Harcourt, Brace & World.

Goodnow, J. J., (1977). *Children's Drawings*, Cambridge: Harvard University Press.

Goodnow, J. J., (1978). Visible thinking: Cognitive aspects of change in drawings. *Child Development*, 49(3) 637-741.

Goodnow, J. J. & Friedman, S. (1972). Orientation in children's human figure drawings: An aspect of graphic language. *Developmental Psychology*, 7, 10-16.

Goodnow, J. J. & Levine R. A. (1973). The grammar of action: Sequence and syntax in children's copying. *Cognitive Psychology*, 4, 82-98.

Goodnow, J. J., Wilkins, P. & Dawes, L. (1986). Acquiring cultural forms: Cognitive aspects of socialization illustrated by children's drawings and judgements of drawings. *International Journal of Behavioural Development*, 9, 485-505.

Graham, F. K. & Kendall, B. S. (1960). Memory for Designs Test: revised general manual. *Perceptual and Motor Skills*, **11**, 147-188. (Monogr. Suppl. 2-V11).

Grice, H. P. (1975). Logic and conversations. In P. Cole & J. Morgan (eds.), *Syntax and semantics, Vol. 3, Speech acts*. New York: Academic Press.

Griffiths, R. (1945). *A study of imagination in early childhood and its function in mental development*. London: Kegan Paul, Trench, Trubner.

Groves, R. J & Fried, P. A. (1991). Developmental items on children's human figure drawings: a replication and extension of Koppitz to younger children. *Journal of Clinical Psychology*, 47(1), 140-148.

Guthrie, P. J (1994). The effects of development, manipulation of objects, and verbal cues on the spatial representation in young children's drawings. *Visual Arts Research*, 20(1), 50-61.

Hagen, M. A. (1976). Development of ability to perceive and produce pictorial depth cue of overlapping. *Perceptual and Motor Skills*, **42**, 1007-1014.

Hagen, M. A. (1985). There is no development in art. In N. H. Freeman & M. V. Cox (eds.), *Visual order: The nature and development of pictorial representation*. Cambridge: University of Cambridge Press.

Hagen, M. A. (1986). *Varieties of realism: Geometries of representational art*. Cambridge, England: University of Cambridge Press.

Harris, D. B. (1963). *Children's drawings as measures of intellectual maturity: A revision and extension of the Goodenough Draw-A-Man Test*. New York: Harcourt, Brace & World.

Harris, D. B. & Pinder, G. D. (1974). The Goodenough-Harris Drawing Test as a measure of intellectual maturity of youths 12-17: United States. *Vital and Health Statistics*, **138**, Serial No. 11.

Harris, D. B. & Roberts J. (1972). Intellectual maturity of children: Demographic and socioeconomic factors. *Vital and Health Statistics*, **116**, Serial No. 11.

Harris D. B., Roberts, J. & Pinder, G. D. (1970). Intellectual maturity of children as measured by the Goodenough-Harris Drawing Test. *Vital and Health Statistics*, **105**, Serial No. 11.

Hochberg, J.(1978). *Perception,* 2nd ed., Englewood Cliffs, N. J.: Prentice-Hall.

Howard, I. P. (1978). Recognition and knowledge of the water-level principle. *Perception*, **7**, 151-160.

Howell, D. C. (1982). *Statistical methods for Psychology*, 2nd ed. Boston: PWS-Kent Publishing Company.

Hughes, M. & Donaldson, M. (1979). The use of hiding games for studying the co-ordination of viewpoints. *Educational Review*, **31**, 133-140.

Hyde, J., Geiringer, E. & Yen, W. (1975). On the empirical relation between spatial ability and sex differences in other aspects of cognitive performance. *Multivariate behavioral Research*, **10**, 289-309.

Ibbotson, A. & Bryant, P. E. (1976). The perpendicular error and the vertical effect in children's drawings. *Perception*, **5**, 319-326.

Ingram, N. & Butterworth, G. (1989). The young child's representation of depth in drawing: Process and product. *Journal of Experimental Child Psychology*, **47(3)**, 356-369.

Israelite, J. (1936). A comparison of the difficulty of items for intellectually normal children and mental defectives on the Goodenough drawing test. *American Journal of Orthopsychiatry*, **6**, 494-503.

Ivanoff, E. (1909). Recherches experimentales sur le dessin des ecoliers de la Suisse Romande: Correlation entre l'aptitude au dessin et les autres aptitudes. *Archives de Psychologie*, **8**, 97-156.

Ives, W. (1980). The use of orientation in children's drawings of familiar objects: Principles versus precepts. *British Journal of Educational Psychology*, **50**, 295-296.

Ives, W. & Rovet, J. (1979). The role of the graphic orientations in children's drawings of familiar and novel objects at rest and in motion. *Merrill Palmer Quarterly*, **24(4)**, 281-292.

Jensen, A. R. (1980). *Bias in mental testing*. New York: Free Press.

Johnson, D. L. & Johnson, C. A. (1971). Comparisons of four intelligence tests used with culturally disadvantaged children. *Psychological Reports*, **28**, 209-210.

Johnson, J. (1982). *The development of metaphor comprehension: Its mental-demand measurement and its process-analytical models*. Unpublished doctoral dissertation, York University, Toronto.

Johnson, L. R., Perlmutter, M. & Trabasso, T. (1979). The leg bone is connected to the knee bone: Children's representation of body parts in memory, drawing and language. *Child Development*, **50**, 1192-1202.

Kaiser, M., Proffitt, D. & Anderson, K. (1985). Judgements of natural and anomalous trajectories in the presence and absence of motion. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **11**, 795-803.

Kalichman, S. C. (1986). Horizontality as a function of sex and academic major. *Perceptual and Motor Skills*, **63**, 903-906.

Kalichman, S. C. (1988). Individual differences in water-level task performance: A component skill analysis. *Developmental Review*, **8**, 273-295.

Kamphaus, R. W. & Pleiss, K. L. (1991). Draw-a-person techniques: Tests in search of a construct. *Journal of School Psychology*, **29**, 395-401.

Karmiloff-Smith. A. (1990). Constraints on representational change: Evidence from children's drawing. *Cognition*, **34(1)**, 57-83.

Kaufman, A. S. (1972). A short form of the Wechsler Preschool and Primary Scale of Intelligence. *Journal of Consulting and Clinical Psychology*, **39**, 361-369.

Kellogg, R. (1970). *Analysing children's art*, Palo Alto, California: National Press Books.

Kennedy, J. M. (1978). *Pictures and the blind.* Paper presented at the American Psychological Association Meeting, Toronto, in the symposium "First encounters of a pictorial kind".

Kerr, M. (1937). Children's drawings of houses. *British Journal of Medical Psychology*, **16**, 206-218.

Kerschensteiner, D. G. (1905). *Die Entwickelung der zeichnerischen Begabung.* Munich: Gerber.

Kifune, N. (1984). The concurrent validity of the Draw-A-Man Intelligence Test in the cerebral palsied children. *Japanese Journal of Special Education*, **22(1)**, 1-6.

Kirk, U. (1981). The development and use of rules in the acquisition of perceptual motor skills. *Child Development*, **52**, 299-305.

Klaus, K. (1992). The development of depth representation in children's drawings: Effects of graphic surface and visibility of the model. *British Journal of Developmental Psychology*, **10(1)**, 71-83.

Kline, P. (1993). *The handbook of psychological testing*. London: Routledge.

Knopf, I. J. & Richards, T. W. (1952). The child's differentiation of sex as reflected in drawings of the human figure. *Journal of Genetic Psychology*, **81**, 99-112.

Koppitz, E. (1968). *Psychological evaluation of human figure drawings*. New York: Grune & Stratton.

Korzenik, D. (1975). Changes in representation between the ages five and seven. *Psychiatry and Art*, **4**, 95-104.

Kosslyn, S. M., Heldmeyer, K. H. & Locklear, E. P. (1977). Children's drawings as data about internal representations. *Journal of Experimental Child Psychology*, **23(2)**, 191-211.

Kramer, J. J. & Conoly J. C. (1992). *The eleventh mental measurements yearbook*. Nebraska: University of Nebraska Press, Lincoln.

Krascum, R., Tregenza, C. & Whitehead, P. (1996). Hidden-feature inclusion in children's drawings: The effect of age and model familiarity. *British Journal of Developmental Psychology*, **14**, 441-455.

Lamprecht, K. (1906). Les dessins d' enfants comme source hisorique. *Bulletin de l'Académie Royale de Belgique, Classe des Lettres*, **Nos. 9-10**, 457-469.

Lange-Küttner, C. & Reith, E. (1995). The transformation of figurative thought: implications of Piaget and Inhelder's developmental theory for children's drawings. In C. Lange-Küttner & G. V. Thomas (eds.), *Drawing and looking: Theoretical approaches to pictorial representation in children*. Hertfordshire: Harvester Wheatsheaf.

Laosa L. M., Swartz, J. D. & Diaz-Guerrero, R. (1974). Perceptual-cognitive and personality development of Mexican and Anglo-American children as measured by human figure drawings. *Developmental Psychology*, **10**, 131-139.

Laosa, L. M., Swartz, J. D. & Holtzman, W. H. (1973). Human figure drawings by normal children: A longitudinal study of perceptual-cognitive and personality development. *Developmental Psychology*, **8**, 350-356.

Lark-Horovitz, B., Barnhart, E. N. & Sills, E. M. (1939). *Graphic work-sample diagnosis: An analytical method of estimating children's drawing ability.* Cleveland, OH: Cleveland Museum of Art, cited in Lark-Horovitz, B., Lewis, H. & Luca, M. (1973). *Understanding children's art for better teaching*, 2nd ed. Columbus, OH: Charles E. Merrill.

Leeds, A., Dirlam, D. & Brannigan, G. G. (1983). The development of spatial representation in children from five to thirteen years of age. *Genetic Psychology Monographs*, **108(1)**, 137-165.

Lehman, E. B. & Levy, B. I. (1971). Discrepancies in estimates of children's intelligence: WISC and human figure drawings. *Journal of Clinical Psychology*, **27**, 74-76.

Lewis, C., Russell, C. & Berridge, D. (1993). When is a mug not a mug? Effects of content, naming, and instructions on children's drawings. *Journal of Experimental Child Psychology*, **56**, 291-302.

Lewis, H. P. (1963). Spatial representation in drawing as a correlate of development and a basis for picture preference. *The Journal of Genetic Psychology*, **102**, 95-107.

Liben, L. S. (1974). Operative understanding of horizontality and its relation to long-term memory. *Child Development*, **45**, 416-424.

Liben, L. S. (1975). Long-term memory for pictures related to seration, horizontality and verticality concepts. *Developmental Psychology*, **11(6)**, 795-806.

Liben, L. S. (1978). Performance on Piagetian spatial tasks as a function of sex field dependence and training. *Merrill-Palmer Quarterly*, **24**, 97-110.

Liben, L. S. (1981). Copying and reproducing pictures in relation to subject's operative levels. *Developmental Psychology*, **17(3)**, 357-365.

Liben, L. S. (1991a). The Piagetian water-level task: Looking beneath the surface. In R. Vasta (Ed.), *Annals of Child Development* (**Vol. 8**, 81-144). London: Kingsley.

Liben, L. S. (1991b). Adults' performance on horizontality tasks: Conflicting frames of reference. *Developmental Psychology*, **8**, 273-295.

Liben, L. S. & Golbeck, S. L. (1980). Sex differences in performance on Piagetian tasks: Differences in competence or performance? *Child Development*, **51**, 594-597.

Liben, L. S. & Golbeck, S. L (1984). Performance on Piagetian horizontality and verticality tasks: Evidence from small-scale and life-size spatial tasks. *Developmental Psychology*, **20**, 595-597.

Liben, L. S. & Golbeck, S. L (1986). Adults' demonstration of underlying Euclidean concepts in relation to task context. *Developmental Psychology*, **22**, 487-490.

Light, P. (1985). The development of view-specific representation considered from a socio-cognitive standpoint. In N. H. Freeman & M. V. Cox (eds.), *Visual order: The nature and development of pictorial representation*. Cambridge: University of Cambridge Press.

Light, P. & Foot, T. (1986). Partial occlusion in young children's drawings. *Journal of Experimental Child Psychology*, **41(1)**, 38-48.

Light, P. & Humphreys, J. (1981). Internal spatial relationships in young children's drawings. *Journal of Experimental Child Psychology*, **31**, 521-530.

Light, P. H. & MacIntosh, E. (1980). Depth relationships in young children's drawings. *Journal of Experimental Child Psychology*, **30**, 79-87.

Light, P. H. & McEwen, F. (1987). Drawings as messages: The effect of a communication game upon production of view-specific drawings. *British Journal of Developmental Psychology*, **5**, 53-60.

Light, P. H. & Simmons, B. (1983). The effects of a communication task upon the representation of depth relationships in young children's drawings. *Journal of Experimental Child Psychology*, **35**, 81-92.

Lindier, R. (1962). *The Goodenough Draw-A-Man Test: its relationship to intelligence, achievement and cultural variables of Negro elementary school children in Southeast United States*. Unpublished doctoral dissertation, Florida State University.

Linn, M. C. & Kyllonen, P. (1981). The field-dependence-independence construct: Some, one or none. *Journal of Educational Psychology*, **73**, 261-273.

Lobsien, M. (1905). Kinderzeichnung und Kunstkanon. *Zeitschrift Pedagogische Psychologie*, **7**, 393-404.

Lohaus, A., Thomas, H., Kessler, T. & Gediga, G. (1996). Decomposing water-level responses: Field effects as separate influences. *Journal of Experimental Child Psychology*, **63**, 79-102.

Lowenfeld, V. (1947). *Creative and mental growth*. New York: Macmillan.

Lowenfeld, V. (1952). *The nature of creative activity: Experimental and comparative studies of visual and non-visual sources of drawing, painting, and sculpture by means of artistic products of weak sighted and blind subjects and of the art of different epochs and cultures*, 2nd ed. London: Routledge & Kegan Paul. (First published in English in 1939).

Lowenfeld, V. & Brittain, W. L. (1975). *Creative and mental growth*. (6th edition compiled by W. L. Brittain), London: Collier-Macmillan.

Luquet, G. H. (1913). *Les dessins d'un enfant*. Paris: Alcan.

Luquet, G. H. (1927/1977). *Le dessin enfantin*. Paris: Alcan. (third edition reprinted by Delachaux et Niestle, Lausanne and Paris, 1977).

Machover, K. (1949). *Personality projection in the drawings of the human figure.* Springfield, IL: C. C. Thomas.

MacKay, C. K., Brazendale, A. H. & Wilson, L. F. (1972). Concepts of horizontal and vertical: A methodological note. *Developmental Psychology*, **7**, 232-237.

Manuel H. & Hughes, L. (1932). The IQ and drawing ability of young Mexican children. *Journal of Applied Psychology*, **4**, 383-387.

Markham, S. (1954). An item analysis of children's drawings of a house. *Journal of Clinical Psychology*, **10**, 185-187.

Marr, D. (1977). Analysis of occluding contours. *Prc. Royal Society of London*, **B197**, 441-475.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco: W. H. Freeman.

Martlew, M. (1988). Children's oral and written language. In A. D. Pellegrini (ed.), *Psychological basis for early education*, Chichester: John Wiley.

Matthews, (1984). Children's drawings: are young children really scribbling? *Early Child Development and Care*, **18**, 1-39.

McAfee, E. A. & Proffitt, D. R. (1991). Understanding the surface orientation of liquids. *Cognitive Psychology*, **23**, 483-514.

McCarthy, D. (1970). *McCarthy Scales of Children's Abilities*. New York: Psychological Corporation.

McCathy, D. (1924). *Children's drawings*. Baltimore, MD: Williams and Wilkins.

McElwee, E. W. (1932). The reliability of the Goodenough intelligence test used with sub-normal children fourteen years of age. *Journal of Applied Psychology*, **16**, 217-218.

McGillicuddy-DeLisi, A. V., DeLisi, R. & Youniss, J. (1978). Representation of the horizontal coordinate with and without liquid. *Merrill-Palmer Quarterly*, **24**, 199-208.

McGilligan, R. P., Yater, A. C. & Hulsing, R. (1971). Goodenough-Harris Drawing Tests reliabilities. *Psychology in Schools*, **8**, 359-362.

Meehan, A. M. & Overton, W. F. (1986). Gender differences in expectancies for success on Piagetian spatial tasks. *Merrill-Palmer Quarterly*, **32**, 427-441.

Mehryar, A. H. Tashakkori, A., Yousefi, F.& Khajavi, F. (1987). The application of the Goodenough-Harris Draw-A-Man Test to a group of Iranian children in the city of Shiraz. *British Journal of Educational Psychology*, **57**, 401-406.

Milbrath, C. (1998). *Patterns of artistic development in children: Comparative studies of talent*. Cambridge: University of Cambridge Press.

Mitchelmore, M. C. (1975). The perceptual development of Jamaican students. (Unpublished PhD. Thesis, Ohio State University). *Dissertation Abstracts International*, **35**, 7130A.

Mitchelmore, M. C. (1985). Geometric foundations of children's drawings. In N. H. Freeman & M. V. Cox (eds.), *Visual order: The nature and development of pictorial representation*. Cambridge: University of Cambridge Press.

Moore, V. R. (1987). The influence of experience on children's drawings of a familiar and unfamiliar object. *British Journal of Developmental Psychology*, **5(3)**, 221-229.

Morra, S. (1995). A neo-Piagetian approach to children's drawings. In C. Lange-Küttner & G. V. Thomas (eds.), *Drawing and looking: Theoretical approaches to pictorial representation in children*. Hertfordshire: Harvester Wheatsheaf.

Morra, S., Angi, A. & Tomat, L. (1996). Planning, encoding and overcoming conflict in partial occlusion drawing: A Neo-Piagetian model and an experimental analysis. *Journal of Experimental Child Psychology*, **61**, 276-301.

Morra, S., Moizo, C. & Scopesi, A. (1988). Working memory (or the M operator) and the planning of children's drawings. *Journal of Experimental Child Psychology*, **46**, 41-73.

Mortensen, K. V. (1991). *Form and content in children's human figure drawings.* New York: University Press.

Motta, R.. W., Little, S. G. & Tobin, M. I. (1993). The use and abuse of Human Figure Drawings. *School Psychology Quarterly*, **8(3)**, 162-169.

Myer, K.A. & Hensley, J. H. (1984). Cognitive style, gender and self-report of principles as predictors of adult performance on Piaget's water-level tasks. *Journal of Genetic Psychology*, **144**, 179-183.

Naglieri, J. A. (1988). *Draw A Person: A Quantitative Scoring System.* New York: Psychological Corporation.

Naglieri, J. A. (1993). Human figure drawings in perspective. *School Psychology Quarterly,* **8(3)**, 170-176.

Naglieri, J. A. & Das, J. P. (1990). Planning, Attention, Simultaneous and Successive (PASS) cognitive processes as a model for intelligence. *Journal of Psychoeducational Assessment*, **8(3)**, 303-337.

Naglieri, J. A., McNeish, T. J. & Bardois, A. N. (1991). *Draw-A-Person: Screening Procedure for Emotional Disturbance.* Austin, TX: ProEd.

Naglieri, J. A., Prewett, P. N. & Bardos, A. N. (1989). An exploratory study of planning attention, simultaneous and successive cognitive processes. *Journal of School Psychology*, 27(4), 347-364.

Nicholls, A. L. & Kennedy, J. M. (1992). Drawing development: From similarity of features to direction. *Child Development*, **63(1)**, 227-241.

Nihei, Y. (1983). Developmental changes in covert principles for the organization of strokes in the drawing and handwriting. *Acta Psychologica*, **54**, 221-232.

O'Keefe, R., Leskosky, R. J., O'Brien, T. G., Yater, A. C. & Barclay, A. (1971). Influence of age, sex, and ethnic origin on Goodenough-Harris Drawing Test performances by disadvantaged preschool children. *Perceptual and Motor Skills*, 33, 708-710.

Oakland, T. & Dowling, L. (1983). The Draw-A-Person Test: Validity properties for nonbiased assessment. *Learning Disability Quarterly*, 6(4), 526-534.

Olson, D. R. (1972). Language use for communicating, instructing and thinking. In J. B. Caroll and R. O. Freedle, (eds.), *Language, comprehension and the acquisition of knowledge*. Washington, D. C.: Winston and Sons, Inc.

Olson, D. R. & Bialystok, E. (1983). *Spatial Cognition: The structure and development of mental representations of spatial relations*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Ortega, D., Arnold, E. L. & Smeltzer, D. J. (1988). Analysis of children's Human Figure Drawings. *Journal of the American Academy of Child and Adolescent Psychiatry*, **27(2)**, 260-261.

Otani, H. & Leonard, S. D. (1988). Judgements of horizontality of still water: Effects of perceptual and cognitive features. *Perceptual and Motor Skills*, **66**, 235-245.

Overton, W. F. & Newman, J. L. (1982). Cognitive development: A competence-activation/utilization approach. In T. Field, A. Huston, H. Quay, L. Troll, & G. Finley (eds.), *Review of Human Development*, (p.217-241). New York: Wiley.

Paivio, A. (1971). Imagery and verbal processes. New York: Holt, Rinehart & Winston.

Papadakis-Michaelidis, E. A. (1989). *Development of children's drawings in relation to gender and culture*. Unpublished PhD thesis, University of Birmingham.

Pariser, D. (1981). Nadia's drawings: Theorising about an autistic child's phenomenal ability. *Journal in Art Education*, **22**, 20-29.

Pascual-Leone, J. (1969). *Cognitive development and cognitive style: A general psychological integration*. Unpublished doctoral thesis, University of Geneva, Switzerland.

Pasual-Leone, J. (1989). An organismic process model of Witkin's field-dependence-independence. In T. Globerson & T. Zelniker (eds.), *Cognitive style and cognitive development* (p. 36-70). Norwood, NJ: Ablex.

Pascual-Leone, J. & Goodman, D. (1979). Intelligence and experience: A neo-Piagetian approach. *Instructional Science*, **8**, 301-367.

Pascual-Leone J. & Morra, S. (1991). Horizontality of water-level: A neo-Piagetian developmental review. *Advances in Child Development and Behavior*, **23**, 231-276.

Pemberton, E. F. (1985). *The syntax of children's drawings: Rules of sequencing and orientation*. Unpublished doctoral dissertation, Pennsylvania State University, University Park.

Pemberton, E. F. (1990). Systematic errors in children's drawings. *Cognitive Development*, **5(4)**, 395-404.

Pennings, A. H. (1991). Altering the strategies in embedded-figure task and water level tasks via instruction: A neo-Piagetian learning study. *Perceptual and Motor Skills*, **72**, 639-660.

Perner, J., Kohlmann, R. & Wimmer, H. (1984). Young children's recognition and use of the vertical and horizontal in drawings. *Child Development*, **55(4)**, 1637-1645.

Phillips, C. J., Smith B. & Broadhurst, A. (1973). The Draw Man Test: A study of scoring methods, validity and norms with English children at five and eleven years. *Journal of Child Psychology and Psychiatry*, **14**, 123-135.

Phillips, W. A., Hobbs, S. B. & Pratt, F. R. (1978). Intellectual realism in children's drawings of cubes. *Cognition*, **6**, 15-33.

Phillips, W. A., Inall, M. & Lauder, E. (1985). On the discovery, storage and use of graphic descriptions. In N. H. Freeman & M. V. Cox (eds.), *Visual order: The nature and development of pictorial representation*. Cambridge: University of Cambridge Press.

Piaget, J. (1953). The child as artist. In E. Ziegfeld (ed.), *Education and art*. Paris: Unesco.

Piaget, J. (1962). *Play, dreams and imitation in childhood.* London: Routledge & Kegan Paul.

Piaget, J. & Inhelder, B. (1948/1967). *The child's conception of space.* 3rd ed., London: Routledge & Kegan Paul.

Piaget, J. & Inhelder, B. (1966/1971). *Mental imagery in the child.* London: Routledge & Kegan Paul.

Piaget, J. & Inhelder, B. (1969). *The psychology of the child.* London: Routledge & Kegan Paul.

Pihl, R. O. & Nimrod, G. (1976). The reliability and validity of the Draw-A-Person Test in IQ and personality assessment. *Journal of Clinical Psychology*, **32**, 470-472.

Pikulski, J. J. (1972). A comparison of figure drawings and WISC IQs among disabled readers. *Journal of Learning Disabilities*, **5**, 156-159.

Pittenger, J. (1987). *Perceptual information in global-directed action.* Paper presented at the Conference on Speaking, Reading, Thinking and Development, Tampa, FL.

Prewett, P. N., Bardos, A. N. & Naglieri, J. A. (1989). Assessment of mentally retarded children with the Matrix Analogies Test-Short Form, Draw A Person: A Quantitative Scoring System, and the Kaufman Test of Educational Achievement. *Psychology in the Schools*, **26**, 254-260.

Pulos, S. (1997). Adolescent's implicit theories of physical phenomena: A matter of gravity. *International Journal of Behavioral Development*, **20(3)**, 493-507.

Radkey, A. L. & Enns, J. T. (1987). Da Vinci's window facilitates drawings of total and partial occlusion in young children. *Journal of Experimental Child Psychology*, **44(2)**, 222-235.

Randall, T. M. (1980). Training the horizontality concept in a group of nontransitional children. *Journal of Genetic Psychology*, **136**, 213-220.

Rebelsky, F. (1964). Adult performance on the Horizontal. *Perceptual and Motor Skills*, **19**, 371-374

Reisman, J. & Yamokoski, Y. (1973). Can intelligence be estimated from the drawings of a man? *Journal of School Psychology*, **22**, 239-244.

Reith, E. (1987). *Attitude intellectuelle et attitude visuelle dans la copie de modeles tridimensionelles et bidimensionelles chez des enfants de 4 a 10 ans*. Unpublished dissertation, Geneva. Cited in C. Lange-Küttner & E. Reith (1995). The transformation of figurative thought: implications of Piaget and Inhelder's developmental theory for children's drawings. In C. Lange-Küttner & G. V. Thomas (eds.), *Drawing and looking: Theoretical approaches to pictorial representation in children*. Hertfordshire: Harvester Wheatsheaf.

Reith, E. (1988). The development of use of contour lines in children's drawings of figurative and non-figurative three-dimensional models. *Archives de Psychologie*, **56(217)**, 83-103.

Reith, E. (1990). Development of representational awareness and performance in drawing production. *Archives de Psychologie*, **58**, 369-379.

Reith, E. (1994). *Using a da Vinci window to study children's ability to measure and represent visual projection*. Paper presented at the Annual Symposium of the Jean Piaget Society, Chicago, June.

Reith, E., Steffen, C. & Gillieron, C. (1994). Children's drawing of water level: Operatory knowledge, attention to visual image, and depiction skills. *Swiss Journal of Psychology*, **53(2)**, 86-97.

Ricci, C. (1887). *L'arte dei bambini*. Bologna: N. Zanichelli.

Roback, H. (1968). Human figure drawings: Their utility in the psychologist's armamentarium for the personality assessment. *Psychological Bulletin*, **70(1)**, 1-19.

Robertson, A. & Youniss, J. (1969). Anticipatory visual imagery in deaf and hearing children. *Child Development*, **40**, 123-135.

Robinson, E. & Robinson, W. (1981). Ways of reacting to communication failure in relation to the development of the child's understanding about verbal communication. *European Journal of Social Psychology*, **11**, 189-208.

Rouma, G. (1913). *Le language graphique de l'enfant*. Paris: Misch et Thron.

Russell, J. & Maxwell, M. (1980). Children's water-level drawing: A problem of ignoring internal frameworks? *Perceptual and Motor Skills*, **51**, 1001-1002.

Saint Exupéry, A. de (1943). *The Little Prince*. New York: Harcourt Brace Jovanovich.

Saracho, O. N. (1985). A modification in scoring the ABC Scale. *Journal of Personality Assessment*, **49(2)**, 154-155.

Saracho, O. N. (1986). Validation of two cognitive measures to assess field-dependence/independence. *Perceptual and Motor Skills*, **63**, 255-263.

Sattler, J. M. (1982). *Assessment of children's intelligence and special abilities*, 2nd ed. Boston: Allyn and Bacon.

Schuyten, M. C. (1904). De oorspronkkelijke 'ventjes' der Antwerpsche schoolkindern. *Paedologisch Jaarboek*, 5, 1-87.

Scott, L, H. (1981). Measuring intelligence with the Goodenough-Harris Drawing Test. *Psychological Bulleting*, **89(3)**, 483-505.

Self, L. (1977). *Nadia: A case of extraordinary drawing ability in an autistic child*. London: Academic Press.

Self, L. (1983). *Normal and anomalous representational drawing ability in children*. London: Academic Press.

Serwer, B. J., Shapiro, B. I. & Shapiro, P. D. (1972). Achievement prediction of "high-risk" children. *Perceptual and Motor Skills*, **35**, 347-354.

Shepp, B. E. & Schwartz, K. B. (1976). Selective attention and processing of integral and nonitegral dimensions: A developmental study. *Journal of Experimental Child Psychology*, **22**, 73-85.

Sheppard, J. L. (1974). The child's concept of horizontality with water levels: A training study. *Australian Journal of Psychology*, **26**, 191-198.

Sholl, M. J. & Liben, L. S. (1995). Illusory tilt and Euclidean schemes as factors in performance on the water-level task. *Journal of Experimental Psychology*, **21(6)**, 1624-1638.

Short-DeGraff, M. A. & Holan, S. (1992). Self-drawing as a gauge of perceptual-motor skill. *Physical and Occupational Therapy in Pediatrics*, **12(1)**, 53-68.

Short-DeGraff, M. A., Slansky, L. & Diamond, K. E. (1989). Validity of preschoolers' self-drawings as an index of human figure drawing performance. *Occupational Therapy Journal of Research*, **9(5)**, 305-315.

Signorella, M. L. & Jamison, W. (1978). Sex differences in the correlations among field dependence, spatial ability, sex role orientation and performance on Piaget's water-level task. *Developmental Psychology*, **14(6)**, 689-690.

Silverstein, A. B. (1968a). Validity of a new approach to the design of WAIS, WISC and WPPSI short forms. *Journal of Consulting and Clinical Psychology*, **32**, 478-479.

Silverstein, A. B. (1968b). Validity of WPPSI short forms. *Journal of Consulting and Clinical Psychology*, **32**, 229-230.

Silverstein, A. B. (1970). Reappraisal of the validity of WAIS, WISC and WPPSI short forms. *Journal of Consulting and Clinical Psychology*, **34**, 12-14.

Silverstein, A. B. (1971). A corrected formula for assessing the validity of WAIS, WISC and WPPSI short forms. *Journal of Clinical Psychology*, **27**, 212-213.

Simner, M. L. (1981). The grammar of action and children's printing. *Developmental Psychology*, **17(6)**, 866-871.

Simner, M. L. (1985). School readiness and the Draw-A-Man Test: An empirically derived alternative to Harris' scoring system. *Journal of Learning Disabilities*, **18(2)**, 77-82.

Sinha, M. (1970). A study of the Harris Revision of the Goodenough Draw-A-Man test. *British Journal of Educational Psychology*, **40**, 221-222.

Smedslund, J. (1963). The effect of observation on children's representation of the spatial orientation of water surface. *Journal of Genetic Psychology*, **102**, 195-201.

Smith, J. M. (1987). *Construct and criterion-related validity of the Draw A Person: A Quantitative Scoring System for normal, reading disabled, and developmentally handicapped children.* Unpublished master's thesis, Ohio State University. In R. C. Reynolds & R. W. Kamphaus (eds.). *Handbook of psychological and educational assessment of children: Intelligence and achievement.* London: The Guilford Press.

Smith, L. & Campbell, J. (1987). One thing behind another: A school based study of children's drawing abilities. *Educational Psychology*, **7(4)**, 283-293.

Smith, L. B. & Kemler, D. G. (1977). Developmental trends in free classification: Evidence from a new conceptualization of perceptual development. *Journal of Experimental Child Psychology*, **24**, 279-298.

Smith, P. & Brown, J. (1982). *Eliciting drawings of partial occlusion from young children.* Paper presented at the Conference on the Graphic Representation, University of York, April.

Snyder, R. T. & Gaston, D. S. (1970). The figure drawing of the first grade child – item analysis and comparison with Koppitz norms. *Journal of Clinical Psychology*, **26**, 377-383.

Sommerville, D. & Cox, M. V. (1988). Sources of error in the Piagetian water-level task. *Perception*, **17**, 249-254.

Spoerl, D. T. (1940). The drawing ability of mentally retarded children. *Journal of Genetic Psychology*, **57**, 259-277.

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, **87**, 245-251.

Stetsenko, A. (1995). The psychological function of children's drawing: a Vygotskian perspective. In C. Lange-Küttner & G. V. Thomas (eds.), *Drawing and looking: Theoretical approaches to pictorial representation in children*, Hertfordshire: Harvester Wheatsheaf.

Stricker, G. (1967). Actuarial, naïve, clinical, and sophisticated clinical predication of pathology from figure drawings. *Journal of Consulting Psychology*, **13**, 492-494.

Strommen, E. (1988). A century of children drawing: The evolution of theory and research concerning the drawings of children. *Visual Arts Research*, **14(2)** [28], 13-24.

Strommen, E. F. (1987). The Goodenough-Harris Draw-A-Person Test and drawing development. *Visual Arts Research*, **13(2)**, 22-28.

Strommen, E. F & Smith, J, F. (1987). Internal validity and bias considering the Draw-A-Person Test. *Educational and Psychological Measurement*, **47**, 731-736.

Struempfer, D. J. (1971). Validation of two quality scales for children's figure drawings. *Perceptual and Motor Skills*, **32**, 887-893.

Strümpfer, D. J. W. & Mienie, C. J. P. (1968). A validation of the Harris-Goodenough Test. *British Journal of Educational Psychology*, **38**, 96-100.

Sully, J. (1895). *Studies of childhood.* London: Longman, Green and Co.

Sundberg, N. D. (1961). The practice of psychological testing in clinical services in the United States. *American Psychologists*, **16**, 79-83.

Sutter, E. G. & Bishop, P. C. (1986). Further investigation of the correlations among the WISC-R, PIAT and DAM. *Psychology in the Schools*, **23**, 365-367.

Swensen, E. (1968). Empirical evaluations of human figure drawings; 1957-1966. *Psychological Bulletin*, **70(1)**, 20-44.

Taylor, M. & Bacharach, V. R. (1982). Constrains on the visual accuracy of drawings produced by young children. *Journal of Experimental child Psychology*, **34**, 311-329.

Terman, L. M. & Merrill, M. A. (1960). *Stanford-Binet Intelligence Scale: Manual for the third revision, Form L-M.* Boston: Houghton Mifflin.

Teske, J. A., Waltz, C. R. & Shenk, C. C. (1992). The effect of object alignment on the representation of depth in young children's drawings. *Journal of Experimental Child Psychology*, **54(1)**, 19-36.

Thomas, G. V. (1995). The role of the graphic strategies and skills. In C. Lange-Küttner & G. V. Thomas (eds), *Drawing and looking: Theoretical Approaches to Pictorial Representation in Children.* Hertfordshire: Harvester Wheatsheaf.

Thomas, G. V. & Silk, A. M. J. (1990). *An introduction to the psychology of children's drawings.* New York: University of New York Press.

Thomas, H. & Jamison, W. (1975). On the acquisition of understanding that the still water is horizontal. *Merrill-Palmer Quarterly,* **21**, 31-44.

Thomas, II., Jamison, W. & Hummel, D. D. (1973). Observation is insufficient for discovering that the surface of still water is invariantly horizontal. *Science*, **181**, 173-174.

Thomas, H. & Lohaus, A. (1993). *Modelling growth and individual differences in spatial tasks.* Monographs of the Society for Research in Child Development, Serial No. 237. Chicago: University of Chicago Press.

Thomas, H., & Turner, G. F. W. (1991). Individual differences and development in water-level task performance. *Journal of Experimental Psychology*, **51**, 171-194.

Tramill, J. Edwards, P. & Tramill, J. (1980). Comparison of the Goodenough-Harris Drawing Test and the WISC-R for children experiencing academic difficulties. *Perceptual and Motor Skills*, **50**, 543-546.

Van Esch, L. (1978). *The measurement of individual processing capacity.* Unpublished master's thesis, York University, Toronto.

Van Sommers P. (1984). *Drawing and Cognition.* Cambridge: Cambridge University Press.

Van Sommers, P. (1989). A system for drawing and drawing-related neuropsychology, *Cognitive Neuropsychology*, 6, 117-164.

Van Sommers, P. (1991). Where writing starts: The analysis of action applied to the historical development of writing. In J. Wann, J. A. Wing & N. Sovik (eds.), *Development of graphic skills*, London: Academic Press.

Van Sommers, P. (1995). Observational, experimental and neuropsychological studies of drawing. In C. Lange-Küttner & G. V. Thomas (eds.), *Drawing and looking: Theoretical approaches to pictorial representation in children.* Hertfordshire: Harvester Wheatsheaf.

Vane, J. R. (1967). An evaluation of the Harris revision of the Goodenough Draw-A-Man Test. *Journal of Clinical Psychology,* **23**, 375-377.

Vasta, R., Belongia, C. & Ribble, C. (1994). Investigating the orientation effect on the water-level task: Who? when? and why? *Developmental Psychology*, **30(6)**, 893-904.

Vasta, R. & Liben, L. S. (1996). The water-level task: An intriguing puzzle. *Current Directions in Psychological Science*, **5**, 171-177.

Verworn, M. (1917). *Zur Psychologie der primitiven Kunst.* Jena: Fischer, (First published 1908).

Wallon, M. P. & Baudoin, C. (1990). A new criterium in the assessment of the behaviour of the drawing child. *British Journal of Educational Psychology*, **60(3)**, 338-348.

Ward, R. & Eliot, J. (1987). A study of migrant children. *Psychological Reports*, **60**, 120-122.

Wechsler, D. (1967). *Manual for the Wechsler Preschool and Primary Scale of Intelligence.* New York: The Psychological Corporation.

Wechsler, D. (1992). *WISC-III^{UK} manual.* New York: Psychological Corporation.

Wells, D. G. & Pedrini, D. T. (1971). Relationships among Wechsler Adult Intelligence Scale, Goodenough-Harris and Peabody Picture Vocabulary Tests with institutionalised retarded adults. *Perceptual and Motor Skills*, **33**, 227-232.

Werner, H. (1948). *Comparative psychology of mental development* (Rev. ed.). New York: International University Press.

White, T. H. (1979). Correlations among the WISC-R, PIAT, and DAM. *Psychology in the Schools*, **16**, 497-501.

Willats, J. (1977). How children learn to draw realistic pictures. *Quarterly Journal of Experimental Psychology*, **29**, 367-382.

Willats, J. (1981). What do the marks in the picture stand for? The child's acquisition of systems of transformation and denotation. *Review of Research in Visual Arts Education*, **13**, 18-33.

Willats, J. (1985). Drawing systems revisited: the role of denotation systems in children's figure drawings. In N. H. Freeman & M. V. Cox (eds.), *Visual order: The nature and development of pictorial representation*. Cambridge: University of Cambridge Press.

Willats, J. (1987). Marr and pictures: An Information processing account of children's drawings. *Archives de Psychologie*, **55**, 105-125.

Willats, J. (1992). The representation of extendedness in children's drawings of sticks and discs. *Child Development*, **63(3)**, 692-710.

Willats, J. (1995). An information-processing approach to drawing development. In C. Lange-Küttner & G. V. Thomas (eds.), *Drawing and looking: Theoretical approaches to pictorial representation in children*, Hertfordshire: Harvester Wheatsheaf.

Willemsen, E. & Reynolds, B. (1973). Sex differences in adults' judgement of the horizontal. *Developmental Psychology*, **8**, 309.

Willis, G. W. (1987). Retention/promotion decisions: selective use of data? *Perceptual and Motor Skills*, **64**, 287-290.

Willsdon, J. A. (1977). A discussion of some sex differences in a study of human figure drawings by children aged four-an-a-half to seven-and-a-half years. In G. Butterworth (ed.), *The child's representation of the world*. New York: Plenum Press.

Wilson, B. & Wilson, M. (1982). The persistence of the perpendicular principle: Why, when and where innate factors determine the nature of the drawings. *Review of Research in Visual Art Education*, **15**, 1-18.

Winner, E. (1982). *Invented Words: The psychology of the arts*, Cambridge, MA: University of Harvard Press.

Wisniewski, J. J. & Naglieri, J. A. (1989). Validity of the Draw A Person: A quantitative scoring system with the WISC-R. *Journal of Psychoeducational Assessment*, 7, 346-351.

Witkin, H. A. (1978). *Cognitive styles in personal and cultural adaptation*. (Vol. XI, 1977 Heinz Werner Lecture Series). Worcester, MA: University of Clark Press.

Witkin, H. A. & Goodenough, D. R. (1981). *Cognitive style: Essence and origins*. New York: International Universities Press.

Witkin, H. A., Oltman, P. K., Raskin, E. R. & Karp, S. A. (1971). *A Manual for the Embedded-Figure Test*. California: Consulting Psychologists Press, Palo Alto.

Witkin, H. A. & Price-Williams, D. (1974). Social conformity and psychological differentiation. *International Journal of Psychology*, **9**, 11-29.

Witt, J. C. (ed.) (1993). *School Psychology Quarterly*, [8(3), p. 159-237]. London: Guilford Press.

Wittig, M. A. & Allen, M. J. (1984). Measurement of adult performance on Piaget's water horizontality task. *Intelligence*, **8**, 305-313.

Wolf, D. P. (1997). Reimagining development: Possibilities from the study of children's art. *Human Development*, **40**, 189-194.

Wolf, V. C. (1971). Age and sex differences as measured by a new non-verbal visual perceptual test. *Psychological Science*, **25**, 85-87.

Yule, W., Lockyer, L. & Noone, A. (1967). The reliability and validity of the Goodenough-Harris Drawing Test. *British Journal of Educational Psychology*, **37**, 110-111.

Zhi, Z., Thomas, G. V. & Robinson, E. J. (1997). Constraints on representational change: Drawing a man with two heads. *British Journal of Developmental Psychology*, **15**, 275-290.

APPENDIX A

General notes:

1. Tables GH-M/GH-W	Description and location of items of the GH scales.
2. Table S-D	Description of GH items loading on Strommen's dimensions.
3. Table KDI	Description of Koppitz Developmental Items.
4. Tables R_x	Tables reporting various estimates of reliability.
5. Tables M_x-S_{wx}	Tables reporting percentage of children passing each item by scale (Man/Woman/$Self_M$/$Self_W$) and age group.
6. Tables A_x	Tables reporting items which reliably discriminate children at successive age groups for each scale.
7. Tables I_x	General information from the descriptive analysis on children's performance at various criterion measures of ability.
8. Tables V_x	Tables reporting various estimates of validity.

Table GH-M Scoring items in Goodenough-Harris Man Scale.

1.	Head present	24.	Fingers present	50.	Proportion: face
2.	Neck present	25.	Correct number of fingers	51.	Proportion: arms I
3.	Neck, 2D	26.	Detail of fingers correct	52.	Proportion: arms II
4.	Eyes present	27.	Opposition of thumb shown	53.	Proportion: legs
5.	Eye detail: Brow or lashes	28.	Hands present	54.	Proportion: limbs in 2-D
6.	Eye detail: pupil	29.	Wrist or ankle shown	55.	Clothing I
7.	Eye detail: proportion	30.	Arms present	56.	Clothing II
8.	Eye detail: glance	31.	Shoulders I	57.	Clothing III
9.	Nose present	32.	Shoulders II	58.	Clothing IV
10.	Nose, 2D	33.	Arms at side or in activity	59.	Clothing V
11.	Mouth present	34.	Elbow joint shown	60.	Profile I
12.	Lips, 2D	35.	Legs present	61.	Profile II
13.	Nose & lips in 2D	36.	Hip I (crotch)	62.	Full face
14.	Chin & forehead shown	37.	Hip II	63.	Motor coordination: lines
15.	Projection of chin shown;	38.	Knee joint shown	64.	Motor coordination: juncture
	chin clearly differentiated	39.	Feet I: any indication	65.	Superior motor coordination
	form lower lip	40.	Feet II: proportion	66.	Directed lines/form: head
16.	Line of jaw indicated	41.	Feet III: heel	67.	Directed lines/form: trunk
17.	Bridge of nose	42.	Feet IV: perspective	68.	Directed lines/form: legs/arms
18.	Hair I	43.	Feet V: detail	69.	Directed lines/form: facial
19.	Hair II	44.	Attachment of arms/legs I		features
20.	Hair III	45.	Attachment of arms/legs II	70.	Sketching technique
21.	Hair IV	46.	Trunk present	71.	Modelling technique
22.	Ears present	47.	Trunk in proportion, 2D	72.	Arm movement
23.	Ears present: proportion and	48.	Proportion: head I	73.	Leg movement
	position	49.	Proportion: head II		

Table GH-W Scoring items in Goodenough-Harris Woman Scale.

1.	Head present	27.	Elbow joint shown	51.	No transparencies in the figure
2.	Neck present	28.	Fingers present	52.	Garb feminine
3.	Neck, 2D	29.	Correct number of fingers	53.	Garb complete without
4.	Eyes present	30.	Detail of fingers correct		incongruities
5.	Eye detail: Brow or lashes	31.	Opposition of thumb shown	54.	Garb a definite "type"
6.	Eye detail: pupil	32.	Hands present	55.	Trunk present
7.	Eye detail: proportion	33.	Legs present	56.	Trunk in proportion, 2D
8.	Cheeks	34.	Hip	57.	Head-trunk proportion
9.	Nose present	35.	Feet I: any indication	58.	Head: proportion
10.	Nose, 2D	36.	Feet II: proportion	59.	Limbs: proportion
11.	Bridge of nose	37.	Feet III: detail	60.	Arms in proportion to trunk
12.	Nostrils shown	38.	Shoe I: feminine	61.	Location of waist
13.	Mouth present	39.	Shoe II: style	62.	Dress area
14.	Lips, 2-D	40.	Placement of feet to figure	63.	Motor coordination: juncture
15.	"Cosmetic lips"	41.	Attachment of arms/legs I	64.	Motor coordination: lines
16.	Nose & lips in 2D	42.	Attachment of arms/legs II	65.	Superior motor coordination
17.	Chin & forehead shown	43.	Clothing indicated	66.	Directed lines/form: head
18.	Line of jaw indicated	44.	Sleeve I	67.	Directed lines/form: breast
19.	Hair I	45.	Sleeve II	68.	Directed lines/form: hip
20.	Hair II	46.	Neckline I		contour
21	Hair III	47.	Neckline II: collar	69.	Directed lines/form: arms taper
22.	Hair IV	48.	Waist I		
23.	Necklace or earrings	49.	Waist II	70.	Directed lines/form: calf of leg
24.	Arms present	50	Skirt "modelled" to indicate	71.	Directed lines/form: facial
25.	Shoulders		pleats or draping		features
26.	Arms at side or in activity				

Table S-D Items common to GH Man and Woman scales in Strommen's dimensions (1987).

Core Dimension			Elaborate Dimension		
Item	MS	WS	Item	MS	WS
Head present	1	1	Neck	2	2
Eyes present	4	4	Neck, two dimensions	3	3
Nose present	9	9	Eye detail: brow or lashes	5	5
Mouth present	11	13	Eye detail: Pupil	6	6
Hair I	18	19	Nose, two dimensions	10	10
Hair II	19	20	Bridge of nose	17	11
Fingers shown	24	28	Lips, two dimension	12	14
Arms present	30	24	Nose/Lips two dimensions	13	16
Legs present	35	33	Hair III	20	21
Feet I: any indication	39	35	Hair IV	21	22
Attachment to limbs I	44	41	Correct number of fingers	25	29
Trunk present	46	55	Detail of fingers correct	26	30
Trunk two dimensions	47	56	Opposition of thumb shown	27	31
Motor coordination: lines	63	64	Hands present	28	32
			Shoulders I	31	25
			Arms at side	33	26
			Hip present	36	24
			Feet II: proportion	40	36
			Attachment of limbs II	45	42
			Head/trunk proportion	48	57
			Face proportion	50	58
			Arm/trunk proportion	51	60
			Motor coordination: juncture	64	63

Table KDI The thirty developmental items in Koppitz's Draw-A-Person Test.

1.	Head	16.	Arms correctly attached to shoulders
2.	Eyes	17.	Elbows
3.	Pupils	18.	Hands
4.	Eyebrows or eyelashes	19.	Fingers
5.	Nose	20.	Correct number of fingers
6.	Nostrils	21.	Legs
7.	Mouth	22.	Legs in two dimensions
8.	Two lips	23.	Knees
9.	Ears	24.	Feet
10.	Hair	25.	Feet in two dimensions
11.	Neck	26.	Profile
12.	Body	27.	Clothing: one item or none
13.	Arms	28.	Clothing: two or three items
14.	Arms in two dimensions	29.	Clothing: four or more items
15.	Arms pointing downwards	30.	Good proportions

Table R1 Intrascorer reliability coefficients for GH Scales by age and gender.

			Man	Woman	Self
Age level	Sex	N	*r*	*r*	*r*
4	Boys	14	.98	.97	.99
	Girls	20	.99	.97	.95
	Combined	34	.99	.97	.96
5 ½	Boys	10	.99	.97	.99
	Girls	16	.99	.97	.98
	Combined	26	.99	.98	.98
7	Boys	8	.99	.99	1.00
	Girls	20	.98	.99	.98
	Combined	28	.98	.99	.99
8	Boys	10	.99	.99	.99
	Girls	17	.99	.99	.98
	Combined	27	.99	.99	.99
Total		115	.99	.99	.99

Note The intrascorer reliability coefficients are calculated on raw scores.

Table R2 Inter-scale correlation coefficients for the GH Scales by age level and gender.

Age		N	Man-Woman	Man-Self	Woman-Self
4 year olds	Boys	14	.60[a]	.80	.61[a]
	Girls	20	.95	.63[b]	.56[a]
	Combined	34	.87	.66	.54
5 ½ year olds	Boys	10	.76[a]	.67[a]	.32[c]
	Girls	16	.92	.93	.86
	Combined	26	.86	.82	.70
7 year olds	Boys	8	.91[b]	.81[a]	.69[c]
	Girls	20	.89	.79	.84
	Combined	28	.88	.76	.72
8 year olds	Boys	10	.86	.81[b]	.78[b]
	Girls	17	.71	.83	.68[b]
	Combined	27	.74	.79	.72
Total	Boys	42	.70	.75	.64
	Girls	73	.89	.76	.76
	Combined	115	.84	.74	.72

Note 1 The alternative-form reliability coefficients are calculated on standard scores. Those on raw scores are reported in Chapter III/Section 3.1.3, Table 3.1.

Note 2 The correlation coefficients are significant at .001 except in the following cases
a = the correlation coefficient is significant at .05
b = the correlation coefficient is significant at .01
c = NS

Note 3 Underlined coefficients between the MS and the WS with SS represent the relation of two same-sex HFDs, differentiated in terms of age (man-boy and woman-girl).

Table R3 Correlation coefficients for the GH Scales by type of items broken down by age level and gender.

			Core items (n = 14)			Elaborate items (n = 23)			Common items (n = 44)		
Age		N	Man Woman	Man Self	Woman Self	Man Woman	Man Self	Woman Self	Man Woman	Man Self	Woman Self
4	Boys	14	.65[a]	.77	.74[b]	.50[c]	.34[c]	.05[c]	.70[b]	.78	.75[b]
	Girls	20	.91	.85	.75	.75	.65[b]	.57[b]	.92	.87	.80
	Combined	34	.80	.82	.74	.70	.61	.48[a]	.85	.84	.79
5½	Boys	10	.54[c]	.27[c]	.34[c]	.87	.83[b]	.75[a]	.69[a]	.71[a]	.33[c]
	Girls	16	.80	.78	.77	.77	.57[a]	.62[a]	.90	.80	.67[b]
	Combined	26	.70	.61	.62	.84	.65	.66	.85	.78	.59
7	Boys	8	.76[a]	.89[b]	.64[c]	.71[b]	.87[b]	.71[b]	.74[a]	.92	.73[a]
	Girls	20	.78	.71	.67	.76	.81	.78	.79	.76	.79
	Combined	28	.77	.75	.66	.75	.84	.76	.78	.81	.76
8	Boys	10	.87	.40[c]	.00[c]	.56[c]	.65[a]	.76[a]	.76[a]	.70[a]	.69[a]
	Girls	17	.73	.80	.57[a]	.71	.83	.82	.75	.91	.82
	Combined	27	.69	.55[b]	.36[c]	.67	.77	.79	.74	.84	.76
Total	Boys	42	.68	.74	.64	.71	.80	.75	.69	.79	.61
	Girls	73	.89	.85	.80	.69	.73	.75	.81	.83	.78
	Combined	115	.82	.82	.75	.70	.76	.74	.78	.82	.73

Note 1 The coefficients are calculated on raw scores.
Note 2 The correlation coefficients are significant at .001 except in the following cases.
a = the correlation coefficient is significant at .05.
b = the correlation coefficient is significant at .01.
c = NS.
Note 3 All correlations are bivariate, except those for the cells for totals; these are partial correlations, controlling for age.

Table M₁ Type of items scored from Man Scale and percentage of 4 year old children passing them (n = 34) presented against Groves & Fried's data (1991). Comparisons are based on Koppitz's classification of frequency categories for her 30 developmental items.

	Items		Location[a]	Dimension[b]	Percentage		Pattern of
					P data[c]	GF data[c]	Agreement[d]
1.	Head		1	C	76	95	D
2.	Neck		2	E	03	11	*
3.	Eyes		4	C	91	92	****
4.	Eyes I	(brow/lashes)	5	E	09	13	*
5.	Eyes II	(pupil)	6	–	38	25	**
6.	Nose		9	C	47	77	D
7.	Mouth		11	C	85	88	d
8.	Mouth	(lips 2D)	12	E	03	00	*
9.	Hair		18	C	38	68	D
10.	Ear		22	–	15	29	d
11.	Fingers		24	C	26	27	**
12.	Fingers	(correct number)	25	E	06	08	*
13.	Hands		28	E	09	26	D
14.	Arms		30	C	65	70	***
15.	Legs		35	C	76	84	***
16.	Feet I		39	C	29	50	**
17.	Feet II	(proportion)	40	E	03	00	*
18.	Attachment	(limbs I)	44	C	26		
19.	Attachment	(limbs II)	45	E	09		
20.	Trunk		46	C	53	50	d
21.	Trunk	(proportion)	47	C	35		
22.	Head/trunk	(proportion I)	48	E	03		
23.	Arms/trunk	(proportion I)	51	E	03		
24.	Clothing I		55	–	06	100	D

[a] Location of the item in GH man Scale

[b] Location of the item in either of the two developmental dimensions found by Strommen' s (1987) multidimensional scaling method (MDS) on GHDT test. Items assigned with a *C* load on the core dimension and those assigned with an *E* load on the elaborate dimension. The items with no indication (–) were excluded from MDS either by being unique to each scale or by having a very low frequency of appearance.

[c] Percentage of children in the present sample scoring each item.

[a, b, c] These explanatory notes apply to Tables W_1 to S_{w4}

[d] Percentage of children in the Groves and Fried's sample (1991) scoring each item. There data is used for comparison with the 4 year old children since Koppitz does not give normative data for children younger than 5 years of age. Cells without a value indicate that the feature was not included in Koppitz's developmental items (KDI) or GH item is not directly comparable with that of Koppitz's.

[e] D = discordance in the frequency category where the item falls comparing the present data with Groves & Fried 's data. d = disagreement is marginal.

Frequency categories
- * = Exceptional item (≤ 15%)
- ** = Not unusual items (16-50%)
- *** = Common items (51-85%)
- **** = Expected items (≥ 86%)

Notes

1. The percentages for each item have been rounded by .01 (i.e. .645 is recorded as .65)
2. 24 out of Koppitz's 30 developmental items (KDI) are comparable with GH-MS. The six items that are not directly equivalent are nostrils, arms 2D, legs 2D, arms at shoulders, good proportions and clothing ≥ 4 items.

In contrast with GH, Koppitz also scored clothing on the basis of three categories, which were mutually exclusive. To allow comparisons, Koppitz data on clothing categories has been treated independently. The table above provides information on 19 features from KDI. The remaining five items were never scored by the present sample and are presented next, against percentages from Groves & Fried's study. All of them were exceptional features, consistently in both studies. Arms at side: 0% vs. 14%; Elbows: 0% vs. 5%; Knees: 0% vs. 9%; Clothing II: 0% vs. 0%; Profile 0% vs. 0%.

Table W₁ Type of items scored from Woman Scale and percentage of 4 year old children passing them (n = 34), presented against Groves & Fried's data (1991). Comparisons are based on Koppitz's classification of frequency categories for her 30 developmental items.

	Items		Location[a]	Dimension[b]	Percentage		Pattern of
					P data	GF data	agreement
1.	Head		1	C	85	95	d
2.	Neck		2	E	06	11	*
3.	Eyes		4	C	94	92	****
4.	Eyes I	(brow/lashes)	5	E	06	13	*
5.	Eyes II	(pupil)	6	–	26	25	**
6.	Cheeks		8	–	03	–	–
7.	Nose		9	C	62	77	***
8.	Nose	(proportion)	10	E	03	–	–
9.	Nose	(nostrils)	12	–	06	03	*
10.	Mouth		13	C	77	88	d
11.	Mouth	(lips 2D)	14	E	06	00	*
12.	Jaw		18	–	03	–	–
13.	Hair I		19	C	50	68	d
14.	Arm		24	C	47	70	D
15.	Fingers		28	C	27	27	**
16.	Fingers	(correct number)	29	E	03	08	*
17.	Hands		32	E	09	26	D
18.	Legs		33	C	73	84	***
19.	Feet		35	C	38	50	**
20.	Attachment	(limbs I)	41	C	18	–	–
21.	Clothing		43	–	09	100	D
22.	Waist I		48	–	03	–	–
23.	Trunk		55	C	59	50	d
24.	Trunk	(proportion)	56	C	38	–	–
25.	Head/trunk	(proportion I)	57	E	03	–	–

Note 22 of Koppitz's 30 developmental items (KDI) are comparable with GH-WS. The eight items that are not directly equivalent are arms 2D, legs 2D, arms at shoulders, good proportions, clothing ≥ 4 items, ears, knees, profile. Data for the 18 KDI is presented in the table above. The remaining four items which were never scored by the present sample are reported next, against percentages from Groves & Fried's data. Arms at side: 0% vs. 14%; Elbows: 0% vs. 5%; Feet 2D: 0% vs. 0%; Clothing II: 0% vs. 0%.

Table M_2 Type of items scored from Man Scale and percentage of 5½ year old children passing them (n = 26), presented against Koppitz's normative data (1968). Comparisons are based on Koppitz's classification of frequency categories for her 30 developmental items.

		Items		Location	Dimension	Percentage P data	K data[a]	Pattern of Agreement
1.		Head		1	C	100	100	****
2.		Neck		2	E	12	29	D
3.	→	Neck	(proportion)	3	E	08		
4.		Eyes		4	C	100	99	****
5.		Eyes I	(brow/lashes)	5	E	19	33	**
6.		Eyes II	(pupil)	6	–	54	22	D
7.	→	Eyes III	(proportion)	7	E	04		
8.	→	Eyes IV	(gaze)	8	–	12		
9.		Nose		9	C	81	91	d
10.	→	Nose	(proportion)	10	E	12		
11.		Mouth		11	C	96	95	****
12.		Mouth	(lips 2D)	12	E	08	03	*
13.	→	Chin/forehead		14	–	39		
14.	→	Chin	(projection)	15	–	08		
15.		Hair I		18	C	85	76	***
16.	→	Hair II		19	C	42		
17.		Ear		22	–	23	27	**
18.		Fingers		24	C	77	62	***
19.		Fingers	(correct number)	25	E	15	20	d
20.	→	Fingers	(proportion)	26	E	15		
21.		Hands		28	E	54	37	d
22.		Arms		30	C	100	92	****
23.	→	Shoulders I		31	E	08		
24.	→	Shoulders II		32	–	04		
25.	→	Arms	(at sides)	33	E	31	31	**
26.		Legs		35	C	96	96	****
27.	→	Hip I		36	E	42		
28.		Feet I		39	C	89	79	d
29.		Feet II	(proportion)	40	E	46	14	D
30.	→	Feet V	(detail)	43	–	12		
31.		Attachment	(limbs I)	44	C	92		
32.		Attachment	(limbs II)	45	E	35		
33.		Trunk		46	C	96	93	****
34.		Trunk	(proportion)	47	C	96		
35.		Head/trunk	(proportion I)	48	E	27		
36.	→	Face	(proportion)	50	E	35		
37.		Arms/trunk	(proportion I)	51	E	08		
38.	→	Legs	(proportion)	53	–	27		
39.	→	Limbs	(proportion)	54	–	05		
40.		Clothing I		55	–	46	71	D
41.	→	Lines	(quality)	63	C	54		
42.	→	Junctures	(quality)	64	E	27		

Note Koppitz'z normative data is used for purpose of comparison with the present data. Arrows denote the appearance of novel items compared to the previous age group. These remarks apply to subsequent tables with the addition that sometimes an item which was present at a younger age group was absent at the next age group. For those cases the direction of the arrow is the reverse (←).

[a] K = Koppitz data. Since the mean age of the present sample falls between 5 and 6 years of age, the averaged proportional values from these two age groups were calculated from Koppitz's data. The following items were never scored by the present sample and are presented against Koppitz percentages. Elbow: 0% vs. 4%; knee: 0% vs. 2%; profile: 0% vs. 3%; clothing 2-3 items: 0% vs. 25%.

Table W_2 Type of items scored from Woman Scale and percentage of 5½ year old children passing them (n = 26), presented against Koppitz's normative data (1968). Comparisons are based on Koppitz's classification of frequency categories for her 30 developmental items.

		Items		Location	Dimension	Percentage		Pattern of
						P data	**K data**	**agreement**
1.		Head		1	C	100	100	****
2.		Neck		2	E	11	29	D
3.	→	Neck	(proportion)	3	E	04		
4.		Eyes		4	C	100	100	****
5.		Eyes I	(brow/lashes)	5	E	23	33	**
6.		Eyes II	(pupil)	6	–	42	22	**
7.	→	Eyes	(proportion)	7	E	04		
8.		Cheeks		8	–	08		
9.		Nose		9	C	85	91	d
10.		Nose	(proportion)	10	E	23		
11.		Nose	(nostrils)	12	–	08	07	*
12.		Mouth		13	C	92	95	****
13.		Mouth	(lips 2D)	14	E	04	03	*
14.	→	Mouth	(cosmetic lips)	15	–	08		
15.	→	Nose/mouth	(proportion)	16	E	04		
16.	→	Chin/forehead		17	–	35		
	←	Jaw		18				
17.		Hair I		19	C	92	76	D
18.	→	Hair II		20	C	62		
19.	→	Hair III		21	E	04		
20.	→	Jewellery	(necklace/earrings)	23	–	04		
21.		Arms		24	C	92	92	****
22.	→	Shoulders		25	E	04		
23.	→	Arms	(at side)	26	E	27	31	**
24.		Fingers		28	C	62	62	***
25.		Fingers	(correct number)	29	E	08	20	D
26.	→	Fingers	(proportion)	30	E	04		
27.		Hands		32	E	50	37	**
28.		Legs		33	C	73	96	D
29.	→	Hip		34	E	04		
30.		Feet I		35	C	92	79	D
31.	→	Feet II	(proportion)	36	E	31	14	d
32.	→	Feet III	(detail)	37	–	08		
33.	→	Shoe	(feminine)	38	–	04		
34.	→	Feet	(placement)	40	–	20		
35.		Attachment I	(limbs)	41	C	89		
36.	→	Attachment II	(limbs)	42	E	39		
37.		Clothing		43	–	65		
38.	→	Sleeve I		44	–	39		
39.	→	Sleeve II		45	–	04		
40.	→	Neckline I		46	–	04		
41.		Waist I		48	–	12		
42.	→	Garb I	(feminine)	52	–	23		
43.	→	Garb II	(complete)	53	–	04		
44.		Trunk		55	C	96	93	****
45.		Trunk	(proportion)	56	C	92		
46.		Head/trunk	(proportion I)	57	E	39		
47.	→	Face	(proportion)	58	E	19		
48.	→	Limbs	(proportion)	59	–	62		
49.	→	Arms/trunk	(proportion)	60	E	04		
50.	→	Waistline	(position)	61	–	04		
51.	→	Dress area	(proportion)	62	–	08		
52.	→	Junctures	(quality)	63	E	12		
53.	→	Lines	(quality)	64	C	31		

Table M₃ Type of items scored from Man Scale and percentage of 7 year old children (n = 28) passing them, presented against Koppitz's normative data (1968). Comparisons are based on Koppitz's classification of frequency categories for her 30 developmental items.

		Items		Location	Dimension	Percentage P data	Percentage K data	Pattern of agreement
1.		Head		1	C	100	100	****
2.		Neck		2	E	21	59	D
3.		Neck	(proportion)	3	E	04		
4.		Eyes		4	C	100	100	****
5.		Eyes I	(brow/lashes)	5	E	36	49	**
6.		Eyes II	(pupil)	6	–	57	45	D
7.		Eyes III	(proportion)	7	E	14		
8.		Eyes IV	(gaze)	8	–	07		
9.		Nose		9	C	93	91	****
10.		Nose	(proportion)	10	E	32		
11.		Mouth		11	C	100	99	****
12.		Mouth	(lips 2D)	12	E	07	13	*
13.		Chin/forehead		14	–	46		
	←	Chin	(projection)	15	–			
14.	→	Nose	(bridge)	17	E	12		
15.		Hair I		18	C	82	85	***
16.		Hair II		19	C	43		
17.	→	Hair III		20	E	04		
18.		Ear		22	–	21	30	**
19.	→	Ear	(proportion/position)	23	–	07		
20.		Fingers		24	C	79	74	***
21.		Fingers	(correct number)	25	E	36	38	**
22.		Fingers	(proportion)	26	E	36		
23.	→	Fingers	(detail: thumb)	27	E	04		
24.		Hands		28	E	50	54	d
25.		Arms		30	C	96	98	****
26.		Shoulders I		31	E	07		
27.		Shoulders II		32	–	04		
28.		Arms	(at sides)	33	E	50	58	d
29.		Legs		35	C	96	99	****
30		Hip I		36	E	57		
31.	→	Hip II		37	–	04		
32.		Feet I		39	C	89	91	****
33.		Feet II	(proportion)	40	E	54	39	D
34.	→	Feet III	(detail: heel)	41	–	04		
35.		Feet V	(detail)	43	–	43		
36.		Attachment	(limbs I)	44	C	96		
37.		Attachment	(limbs II)	45	E	61		
38.		Trunk		46	C	100	99	****
39.		Trunk	(proportion)	47	C	100		
40.		Head/trunk	(proportion I)	48	E	64		
41.	→	Head/trunk	(proportion II)	49	–	07		
42.		Face	(proportion)	50	E	36		
43.		Arms/trunk	(proportion I)	51	E	21		
44.		Legs	(proportion)	53	–	54		
45.		Limbs	(proportion)	54	–	93		
46.		Clothing I		55	–	86	41	D
47.	→	Clothing II		56	–	04	37	D
48.	→	Clothing III		57	–	04		
49.	→	Clothing IV		58	–	04		
50.	→	Clothing V		59	–	04		
51.		Lines	(quality)	63	C	39		
52.		Junctures	(quality)	64	E	46		

Note The features which were never scored were elbow: K-5%; knee: K-1%; profile: K-13%

Table W_3 Type of items scored from Woman Scale and percentage of 7 year old (n = 28) children passing them, presented against Koppitz's normative data (1968). Comparisons are based on Koppitz's classification of frequency categories for her 30 developmental items.

		Items		Location	Dimension	Percentage P data	Percentage K data	Pattern of agreement
1.		Head		1	C	100	100	****
2.		Neck		2	E	29	59	D
3.		Neck		3	E	07		
4.		Eyes		4	C	100	100	****
5.		Eyes I	(brow/lashes)	5	E	46	49	**
6.		Eyes II	(pupil)	6	–	54	45	d
7.		Eyes	(proportion)	7	E	18		
8.		Cheeks		8	–	18		
9.		Nose		9	C	86	91	****
10.		Nose	(proportion)	10	E	36		
11.	→	Nose	(shape: bridge)	11	E	18		
12.		Nose	(nostrils)	12	–	25	15	d
13.		Mouth		13	C	100	99	****
14.		Mouth	(lips 2D)	14	E	18	13	d
15.		Mouth	(cosmetic lips)	15	–	07		
16.		Nose/mouth	(proportion)	16	E	11		
17.		Chin/forehead		17	–	29		
18.	→	Jaw		18	–	04		
19.		Hair I		19	C	100	85	d
20.		Hair II		20	C	75		
21.		Hair III		21	E	11		
22.		Jewellery	(necklace/earrings)	23	–	21		
23.		Arms		24	C	96	98	****
24.		Shoulders		25	E	21		
25.		Arms	(at side)	26	E	54	58	***
26.		Fingers		28	C	79	74	***
27.		Fingers	(correct number)	29	E	21	38	**
28.		Fingers	(proportion)	30	E	25		
29.	→	Fingers	(detail: thumb)	31	E	04		
30.		Hands		32	E	68	54	***
31.		Legs		33	C	96	99	****
32.		Hip		34	E	11		
33.		Feet I		35	C	93	91	****
34.		Feet II	(proportion)	36	E	64	39	D
35.		Feet III	(detail)	37	–	14		
36.		Shoe	(feminine)	38	–	11		
37.		Feet	(placement)	40	–	50		
38.		Attachment I	(limbs)	41	C	96		
39.		Attachment II	(limbs)	42	E	64		
40.		Clothing		43	–	96	41	D
41.		Sleeve I		44	–	82		
42.		Sleeve II		45	–	11		
43.		Neckline I		46	–	21		
44.		Waist I		48	–	29		
45.	→	Waist II		49	–	07		
46.	→	Skirt	(draping)	50	–	14		
47.	→	Occlusion	(dress/body)	51	–	14		
48.		Garb I	(feminine)	52	–	68	37	D
49.		Garb II	(complete)	53	–	04		
50.	→	Garb III	(type)	54	–	11		

Table continues in the next page

Table W3 continues from the previous page

	Items		Location	Dimension	Percentage		Pattern of agreement
					P data	K data	
51.	Trunk		55	C	100	99	****
52.	Trunk	(proportion)	56	C	100		
53.	Head/trunk	(proportion I)	57	E	57		
54.	Face	(proportion)	58	E	29		
55.	Limbs	(proportion)	59	–	89		
56.	Arms/trunk	(proportion)	60	E	32		
57.	Waist	(location)	61	–	14		
58.	Dress area	(proportion)	62	–	25		
59.	Junctures	(quality)	63	E	25		
60.	Lines	(quality)	64	C	43		
61.	→ Breast outline		67	–	04		

Note Elbow was never scored by the present sample whereas was scored by 4% of Koppitz sample

Table M_4 Type of items scored from Man Scale and percentage of 8 year old children (n = 27) passing them, presented against Koppitz's normative data (1968). Comparisons are based on Koppitz's classification of frequency categories for her 30 developmental items.

	Items		Location	Dimension	Percentage		Pattern of
					P data	K data	agreement
1.	Head		1	C	100	100	****
2.	Neck		2	E	59	68	***
3.	Neck	(proportion)	3	E	44		
4.	Eyes		4	C	100	100	****
5.	Eyes I	(brow/lashes)	5	E	22	54	D
6.	Eyes II	(pupil)	6	–	74	49	D
7.	Eyes III	(proportion)	7	E	44		
8.	Eyes IV	(gaze)	8	–	19		
9.	Nose		9	C	96	93	****
10.	Nose	(proportion)	10	E	63		
11.	Mouth		11	C	100	97	****
12.	Mouth	(lips 2D)	12	E	11	13	*
13.	→ Nose/mouth	(proportion)	13	E	11		
14.	Chin/forehead		14	–	59		
15.	→ Chin	(projection)	15	–	11		
16.	→ Jaw		16	–	22		
17.	Nose	(bridge)	17	E	26		
18.	Hair I		18	C	93	88	****
19.	Hair II		19	C	70		
20.	Hair III		20	E	07		
21.	Ear		22	–	48	28	**
22.	Ear	(proportion/position)	23	–	19		
23.	Fingers		24	C	93	69	D
24.	Fingers	(correct number)	25	E	67	42	D
25.	Fingers	(proportion)	26	E	63		
26.	Fingers	(detail: thumb)	27	E	22		
27.	Hands		28	E	89	57	D
28.	Arms		30	C	100	98	****
29.	Shoulders I		31	E	63		
30.	Shoulders II		32	–	22		
31.	Arms	(at sides)	33	E	63	72	***
32.	Legs		35	C	100	96	****
33.	Hip I		36	E	71		
34.	Hip II		37	–	04		
35.	Feet I		39	C	93	88	****
36.	Feet II	(proportion)	40	E	63	42	D
37.	Feet III	(detail: heel)	41	–	07		
38.	Feet V	(detail)	43	–	48		
39.	Attachment	(limbs I)	44	C	100		
40.	Attachment	(limbs II)	45	E	78		
41.	Trunk		46	C	100	100	****
42.	Trunk	(proportion)	47	C	100		
43.	Head /trunk	(proportion I)	48	E	56		
44.	Head/trunk	(proportion II)	49	–	07		
45.	Face	(proportion)	50	E	37		
46.	Arms/trunk	(proportion I)	51	E	37		
47.	→ Arms	(proportion II)	52	–	04		
48.	Legs	(proportion)	53	–	52		

Table continues in the next page

Table M_4 continues from the previous page

		Items		Location	Dimension	Percentage		Pattern of agreement
						P data	K data	
49.		Limbs	(proportion)	54	–	100		
50.		Clothing I		55	–	100	100	****
51.		Clothing II		56	–	07		
52.		Clothing III		57	–	04		
53.		Clothing IV		58	–	07		
	←	Clothing V		59	–			
54.	→	Profile I		60	–	04	17	D
55.		Lines	(quality)	63	C	52		
56.		Junctures	(quality)	64	E	37		

Table W_4 Type of items scored from Woman Scale and percentage of 8 year old children (n = 27) passing them, presented against Koppitz's normative data (1968). Comparisons are based on Koppitz's classification of frequency categories for her 30 developmental items.

	Items		Location	Dimension	Percentage P data	Percentage K data	Pattern of agreement
1.	Head		1	C	100	100	****
2.	Neck		2	E	59	68	***
3.	Neck	(proportion)	3	E	37		
4.	Eyes		4	C	100	100	****
5.	Eyes I	(brow/lashes)	5	E	19	54	D
6.	Eyes II	(pupil)	6	–	78	49	D
7.	Eyes III	(proportion)	7	E	48		
8.	Cheeks		8	–	04		
9.	Nose		9	C	93	93	****
10.	Nose	(proportion)	10	E	67		
11.	Nose	(shape: bridge)	11	E	37		
12.	Nose	(nostrils)	12	–	30	18	**
13.	Mouth		13	C	100	97	****
14.	Mouth	(lips 2D)	14	E	19	13	d
15.	Mouth	(cosmetic lips)	15	–	04		
16.	Nose/mouth	(proportion)	16	E	11		
17.	Chin/forehead		17	–	33		
18	Jaw		18	–	22		
19.	Hair I		19	C	100	88	****
20.	Hair II		20	C	85		
21.	Hair III		21	E	19		
22.	Jewellery	(necklace/earrings)	23	–	04		
23.	Arms		24	C	96	98	****
24.	Shoulders		25	E	52		
25.	Arms	(at side)	26	E	52	72	***
26.	Fingers		28	C	93	69	D
27.	Fingers	(correct number)	29	E	63	42	D
28.	Fingers	(proportion)	30	E	56		
29.	Fingers	(detail: thumb)	31	E	22		
30.	Hands		32	E	82	57	D
31.	Legs		33	C	96	96	****
32.	Hip		34	E	15		
33.	Feet I		35	C	96	88	****
34.	Feet II	(proportion)	36	E	59	42	D
35.	Feet III	(detail)	37	–	22		
36.	Shoe	(feminine)	38	–	26		
37.	Feet	(placement)	40	–	56		
38.	Attachment I	(limbs)	41	C	96		
39.	Attachment II	(limbs)	42	E	74		
40.	Clothing		43	–	100	100	****
41.	Sleeve I		44	–	85		
42.	Sleeve II		45	–	12		
43.	Neckline I		46	–	18		
44.	→ Neckline II		47	–	04		
45.	Waist I		48	–	18		
46.	Waist II		49	–	15		
47.	Skirt	(draping)	50	–	04		
48.	Occlusion	(dress/body)	51	–	04		
49.	Garb I	(feminine)	52	–	70		
50.	Garb II		53	–	11		
51.	Garb III	(type)	54	–	04		

Table continues in the next page

Table W_4 continued from the previous page

	Items		Location	Dimension	Percentage P data	K data	Pattern of agreement
52.	Trunk		55	C	100	100	****
53.	Trunk	(proportion)	56	C	96		
54.	Head/trunk	(proportion I)	57	E	41		
55.	Face	(proportion)	58	E	33		
56.	Limbs	(proportion)	59	–	85		
57.	Arms/trunk	(proportion)	60	E	18		
58.	Waist	(location)	61	–	18		
59.	Dress area	(proportion)	62	–	15		
60.	Junctures	(quality)	63	E	37		
61.	Lines	(quality)	64	C	44		
	← Breast outline		67	–			
62.	→ Arms taper		69	–	11		

Table S_{M1} Type of items scored from Self Scale and percentage of 4 year old boys passing them (n = 14).

	Items		Location	Dimension	Percentage
1.	Head		1	C	86
2.	Neck		2	E	07
3.	Eyes		4	C	93
4.	Eyes I	(brow/lashes)	5	E	14
5.	Eyes II	(pupil)	6	–	21
6.	Nose		9	C	64
7.	Mouth		11	C	64
8.	Hair I		18	C	.50
9.	Hair II		19	C	07
10	Ear		22	–	14
11.	Fingers		24	C	21
12.	Arms		30	C	43
13.	Legs		35	C	93
14.	Feet I		39	C	29
15.	Attachment	(limbs I)	44	C	21
16.	Trunk		46	C	43
17.	Trunk	(proportion)	47	C	36
18.	Clothing I		55	–	07

Table S_{W1} Type of items scored from Self Scale and percentage of 4 year old girls passing them (n = 20).

	Items		Location	Dimension	Percentage
1.	Head		1	C	95
2.	Eyes		4	C	85
3.	Eyes I	(brow/lashes)	5	E	10
4.	Eyes II	(pupil)	6	E	35
5.	Cheeks		8	–	05
6.	Nose		9	C	55
7.	Nose	(shape: bridge)	11	E	05
8.	Mouth		13	C	70
9.	Hair I		19	C	50
10.	Hair II		20	E	05
11.	Arm		24	C	55
12.	Fingers		28	C	35
13.	Fingers	(correct number)	29	E	05
14.	Hands		32	E	15
15.	Legs		33	C	70
16.	Feet		35	C	40
17.	Attachment	(limbs I)	41	C	25
18.	Trunk		55	C	50
19.	Trunk	(proportion)	56	C	35
20.	Head/trunk	(proportion I)	57	E	15
21.	Face	(proportion)	58	E	05

Table S_{M2} Type of items scored from Self Scale and percentage of 5½ year old boys passing them (n = 10).

		Items		Location	Dimension	Percentage
1.		Head		1	C	100
	←	Neck		2	E	
2.		Eyes		4	C	100
	←	Eyes I	(brow/lashes)	5	E	
3.		Eyes II	(pupil)	6	–	40
4.		Nose		9	C	80
5.		Mouth		11	C	100
6.	→	Mouth	(lips 2D)	12	E	10
7.	→	Chin/forehead		14	–	40
8.		Hair I		18	C	100
9.		Hair II		19	C	20
10.		Ear		22	–	10
11.		Fingers		24	C	70
12.	→	Fingers	(correct number)	25	E	20
13.	→	Hands		28	E	50
14.		Arms		30	C	100
15.	→	Shoulders I		31	E	10
16.	→	Arms	(at sides)	33	E	10
17.		Legs		35	C	100
18.	→	Hip I		36	E	10
19.		Feet I		39	C	80
20.	→	Feet II	(proportion)	40	E	40
21.		Attachment	(limbs I)	44	C	90
22.	→	Attachment	(limbs II)	45	E	30
23.		Trunk		46	C	90
24.		Trunk	(proportion)	47	C	80
25.	→	Head/trunk	(proportion I)	48	E	30
26.	→	Face	(proportion)	50	E	10
27.	→	Arms/trunk	(proportion I)	51	E	10
28.	→	Legs	(proportion)	53	–	10
29.	→	Limbs	(proportion)	54	–	40
30.		Clothing I		55	–	10
31.	→	Junctures	(quality)	64	E	10

Table S_{W2} Types of items scored from Self Scale and percentage of 5½ year old girls passing them (n = 16).

		Items		Location	Dimension	Percentage
1.		Head		1	C	100
2.		Eyes		4	C	100
3.		Eyes I	(brow/lashes)	5	E	25
4.		Eyes II	(pupil)	6	–	56
5.	→	Eyes	(proportion)	7	E	13
6.		Cheeks		8	–	13
7.		Nose		9	C	81
8.	→	Nose	(proportion)	10	E	19
	←	Nose	(shape: bridge)	11	E	
9.	→	Nose	(nostrils)	12	–	06
10.		Mouth		13	C	100
11.	→	Mouth	(lips 2D)	14	E	13
12.	→	Mouth	(cosmetic lips)	15	–	06
13.	→	Chin/forehead		17	–	38
14.	→	Jaw		18	–	06
15.		Hair I		19	C	94
16.		Hair II		20	C	63
17.	→	Hair III		21	E	13
18.	→	Jewellery	(necklace/earrings)	23	–	07
19.		Arms		24	C	100
20.	→	Arms	(at side)	26	E	38
21.		Fingers		28	C	75
22.		Fingers	(correct number)	29	E	06
23.		Hands		32	E	50
24.		Legs		33	C	94
25.	→	Hip		34	E	06
26.		Feet I		35	C	94
27.	→	Feet II	(proportion)	36	E	44
28.	→	Feet III	(detail)	37	–	13
29.	→	Feet	(placement)	40	–	19
30.		Attachment I	(limbs)	41	C	100
31.	→	Attachment II	(limbs)	42	E	38
32.	→	Clothing		43	–	69
33.	→	Sleeve I		44	–	56
34.	→	Sleeve II		45	–	13
35.	→	Waist I		48	–	13
36.	→	Garb I	(feminine)	52	–	19
37.		Trunk		55	C	100
38.		Trunk	(proportion)	56	C	100
39.		Head/trunk	(proportion I)	57	E	13
40.		Face	(proportion)	58	E	25
41.	→	Limbs	(proportion)	59	–	75
42.	→	Waist	(location)	61	–	13
43.	→	Dress area	(proportion)	62	–	06
44.	→	Junctures	(quality)	63	E	06
45.	→	Lines	(quality)	64	C	31

Table S_{M3} Type of items scored from Self Scale and percentage of 7 year old boys passing them (n = 7).

		Items		Location	Dimension	Percentage
1.		Head		1	C	100
2.		Eyes		4	C	100
3.	→	Eyes I	(brow/lashes)	5	E	14
4.		Eyes II	(pupil)	6	–	43
5.	→	Eyes III	(proportion)	7	E	14
6.		Nose		9	C	86
7.	→	Nose	(proportion)	10	E	14
8.		Mouth		11	C	100
	←	Mouth	(lips 2D)	12	E	
9.		Chin/forehead		14	–	29
10.		Hair I		18	C	71
11.		Hair II		19	C	29
12.		Ear		22	–	29
13.	→	Ear	(proportion/position)	23	–	14
14.		Fingers		24	C	71
15.		Fingers	(correct number)	25	E	29
16.	→	Fingers	(proportion)	26	E	43
17.		Hands		28	E	29
18.	→	Wrist/ankle		29	–	14
19.		Arms		30	C	100
20.		Shoulders I		31	E	14
21.		Arms	(at sides)	33	E	29
22.		Legs		35	C	100
23.		Hip I		36	E	14
24.		Feet I		39	C	86
25.		Feet II	(proportion)	40	E	57
26.	→	Feet V	(detail)	43	–	14
27.		Attachment	(limbs I)	44	C	100
28.		Attachment	(limbs II)	45	E	14
29.		Trunk		46	C	100
30.		Trunk	(proportion)	47	C	100
31.		Head/trunk	(proportion I)	48	E	71
32.		Face	(proportion)	50	E	14
	←	Arms/trunk	(proportion I)	51	E	
33.		Legs	(proportion)	53	–	43
34.		Limbs	(proportion)	54	–	100
35.		Clothing I		55	–	86
36.	→	Clothing IV		58	–	14
37.	→	Lines	(quality)	63	C	43
38.		Junctures	(quality)	64	E	14

Table S_{W3} Type of items scored from Self Scale and percentage of 7 year old girls passing them (n = 20).

	Items		Location	Dimension	Percentage
1.	Head		1	C	100
2.	→ Neck		2	E	35
3.	→ Neck		3	E	15
4.	Eyes		4	C	100
5.	Eyes I	(brow/lashes)	5	E	35
6.	Eyes II	(pupil)	6	–	75
7.	Eyes III	(proportion)	7	E	20
8.	Cheeks		8	–	20
9.	Nose		9	C	85
10.	Nose	(proportion)	10	E	35
11.	Nose	(shape: bridge)	11	E	10
12.	Nose	(nostrils)	12	–	15
13.	Mouth		13	C	95
14.	Mouth	(lips 2D)	14	E	25
15.	Mouth	(cosmetic lips)	15	–	20
16.	→ Nose/mouth	(proportion)	16	E	10
17.	Chin/forehead		17	–	25
18.	Jaw		18	–	05
19.	Hair I		19	C	100
20.	Hair II		20	C	80
21.	Hair III		21	E	05
22.	Jewellery	(necklace/earrings)	23	–	15
23.	Arms		24	C	95
24.	→ Shoulders		25	E	10
25.	Arms	(at side)	26	E	55
26.	Fingers		28	C	80
27.	Fingers	(correct number)	29	E	20
28.	→ Fingers	(proportion)	30	E	30
29.	Hands		32	E	50
30.	Legs		33	C	95
31.	Hip		34	E	25
32.	Feet I		35	C	95
33.	Feet II	(proportion)	36	E	50
34.	Feet III	(detail)	37	–	35
35.	→ Shoe	(feminine)	38	–	05
36.	Feet	(placement)	40	–	35
37.	Attachment I	(limbs)	41	C	95
38.	Attachment II	(limbs)	42	E	70
39.	Clothing		43	–	95
40.	Sleeve I		44	–	70
41.	Sleeve II		45	–	15
42.	→ Neckline I		46	–	35
43.	→ Neckline II		47	–	05
44.	Waist I		48	–	45
45.	→ Waist II		49	–	10
46.	→ Skirt	(draping)	50	–	10
47.	→ Occlusion	(dress/body)	51	–	25
48.	Garb I	(feminine)	52	–	65
49.	→ Garb II		53	–	10
50.	→ Garb III	(type)	54	–	10

Table S_{W3} continues in the next page

Table S_{W3} continues from the previous page

		Items		Location	Dimension	Percentage
51.		Trunk		55	C	100
52.		Trunk	(proportion)	56	C	95
53.		Head/trunk	(proportion I)	57	E	40
54.		Face	(proportion)	58	E	40
55.		Limbs	(proportion)	59	–	85
56.	→	Arms/trunk	(proportion)	60	E	25
57.		Waist	(location)	61	–	10
58.		Dress area	(proportion)	62	–	25
59.		Junctures	(quality)	63	E	30
60.		Lines	(quality)	64	C	45

Table S_{M4} Type of items scored from Self Scale and percentage of 8 year old boys passing them(n = 10).

		Items		Location	Dimension	Percentage
1.		Head		1	C	100
2.	→	Neck		2	E	60
3.	→	Neck	(proportion)	3	E	50
4.		Eyes		4	C	100
5.		Eyes I	(brow/lashes)	5	E	30
6.		Eyes II	(pupil)	6	–	70
7.		Eyes III	(proportion)	7	E	30
8.	→	Eyes IV	(gaze)	8	–	20
9.		Nose		9	C	90
10.		Nose	(proportion)	10	E	60
11.		Mouth		11	C	100
12.	→	Mouth	(lips 2D)	12	E	20
13.		Nose/mouth	(proportions)	13	E	20
14.		Chin/forehead		14	–	40
15.	→	Chin		15	–	10
16.	→	Jaw		16	–	40
17.	→	Nose	(shape: bridge)	17	E	40
18.		Hair I		18	C	100
19.		Hair II		19	C	80
20.		Ear		22	–	60
21.		Ear	(proportion/position)	23	–	40
22.		Fingers		24	C	90
23.		Fingers	(correct number)	25	E	50
24.		Fingers	(proportion)	26	E	30
25.	→	Fingers	(detail: thumb)	27	E	10
26.		Hands		28	E	70
	←	Wrist/ankle		29	–	
27.		Arms		30	C	90
28.		Shoulders I		31	E	80
29.	→	Shoulders II		32	–	30
30.		Arms	(at sides)	33	E	60
31.		Legs		35	C	100
32.		Hip I		36	E	70
33.		Feet I		39	C	100
34.		Feet II	(proportion)	40	E	50
35.	→	Feet III	(detail: heel)	41	–	20
36.		Feet V	(detail)	43	–	40
37.		Attachment	(limbs I)	44	C	90
38.		Attachment	(limbs II)	45	E	80
39.		Trunk		46	C	100
40.		Trunk	(proportion)	47	C	100
41.		Head/trunk	(proportion I)	48	E	40
42.	→	Head	(proportion II)	49	–	10
43.		Face	(proportion)	50	E	60
44.	→	Arms/trunk	(proportion I)	51	E	20
45.	→	Arms	(proportion II)	52	–	10
46.		Legs	(proportion)	53	–	50
47.		Limbs	(proportion)	54	–	90
48.		Clothing I		55	–	100
49.	→	Clothing II		56	–	10
	←	Clothing IV		58	–	
50.		Lines	(quality)	63	C	60
51.		Junctures	(quality)	64	E	70

Table S_{W4} Type of items scored from Self Scale and percentage of 8 year old girls passing them (n = 17).

	Items		Location	Dimension	Percentage
1.	Head		1	C	100
2.	Neck		2	E	65
3.	Neck		3	E	29
4.	Eyes		4	C	94
5.	Eyes I	(brow/lashes)	5	E	06
6.	Eyes II	(pupil)	6	–	82
7.	Eyes III	(proportion)	7	E	53
8.	Cheeks		8	–	12
9.	Nose		9	C	100
10.	Nose	(proportion)	10	E	71
11.	Nose	(shape: bridge)	11	E	18
12.	Nose	(nostrils)	12	–	06
13.	Mouth		13	C	100
	← Mouth	(lips 2D)	14	E	
	← Mouth	(cosmetic lips)	15	–	
	← Nose/mouth	(proportion)	16	E	
14.	Chin/forehead		17	–	59
15.	Jaw		18	–	24
16.	Hair I		19	C	100
17.	Hair II		20	C	100
18.	Hair III		21	E	24
	← Jewellery	(necklace/earrings)	23	–	
19.	Arms		24	C	100
20.	Shoulders		25	E	41
21.	Arms	(at side)	26	E	47
22.	Fingers		28	C	94
23.	Fingers	(correct number)	29	E	53
24.	Fingers	(proportion)	30	E	59
25.	→ Fingers	(detail: thumb)	31	E	18
26.	Hands		32	E	77
27.	Legs		33	C	100
28.	Hip		34	E	29
29.	Feet I		35	C	100
30.	Feet II	(proportion)	36	E	65
31.	Feet III	(detail)	37	–	41
	← Shoe	(feminine)	38	–	
32.	Feet	(placement)	40	–	41
33.	Attachment I	(limbs)	41	C	100
34.	Attachment II	(limbs)	42	E	65
35.	Clothing		43	–	94
36.	Sleeve I		44	–	82
37.	Sleeve II		45	–	12
38.	Neckline I		46	–	35
	← Neckline II		47	–	
39.	Waist I		48	–	12
40.	Waist II		49	–	06
41.	Skirt	(draping)	50	–	12
	← Occlusion	(dress/body)	51	–	
42.	Garb I	(feminine)	52	–	47
43.	Garb II		53	–	06
44.	Garb III	(type)	54	–	12

Table S_{W4} continues in the next page

Table S_{W4} continues from previous page

	Items		Location	Dimension	Percentage
45.	Trunk		55	C	100
46.	Trunk	(proportion)	56	C	100
47.	Head/trunk	(proportion I)	57	E	35
48.	Face	(proportion)	58	E	24
49.	Limbs	(proportion)	59	–	100
50.	Arms/trunk	(proportion)	60	E	06
51.	Waist	(location)	61	–	06
52.	Dress area	(proportion)	62	–	12
53.	Junctures	(quality)	63	E	41
54.	Lines	(quality)	64	C	53

Table A_M Items from Man Scale which differentiate children at successive ages.

	Items		Status	4 vs. 5½ χ^2	4 vs. 5½ p	5½ vs. 7 χ^2	5½ vs. 7 p	7 vs. 8 χ^2	7 vs. 8 p
1.	Head		1 C	5.17	<.05				
2.	Neck		2 E					6.70	<.01
3.	Neck	(proportion)	3 E					10.56	<.001
4.	Eyes	(proportion)	7 E					4.69	<.05
5.	Nose		9 C	5.73	<.05				
6.	Nose	(proportion)	10 E					4.07	<.05
7.	Chin/forehead		14 –	13.05	<.001				
8.	Jaw		16 –					4.88	<.05
9.	Hair I		18 C	11.20	<.001				
10.	Hair II		19 C	14.90	<.001			3.19	=.07
11.	Ear		22 –					3.24	=.07
12.	Fingers		24 C	13.06	<.001				
13.	Fingers	(Detail: number)	25 E					4.10	<.05
14.	Fingers	(proportion)	26 E	3.40	=.06			3.07	=.08
15.	Hands		28 E	12.57	<.001			8.00	<.01
16.	Arms		30 C	9.37	<.01				
17.	Shoulders I		31 E					16.55	<.001
18.	Arms	(at sides)	33 E	9.55	<.01				
19.	Legs		35 C	3.07	=.08				
20.	Hip I		36 E	14.90	<.001				
21.	Feet I		39 C	18.44	<.001				
22.	Feet II	(proportion)	40 E	13.76	<.001				
23.	Clothing		43 –			5.12	<.05		
24.	Attachment	(limbs I)	44 C	23.21	<.001				
25.	Attachment	(limbs II)	45 E	4.62	<.05				
26.	Trunk		46 C	11.50	<.001				
27.	Trunk	(proportion)	47 C	20.58	<.001				
28.	Head/trunk	(proportion I)	48 E	5.40	<.05	6.14	<.05		
29.	Face	(proportion)	50 E	11.26	<.001				
30.	Legs	(proportion)	53 –	7.91	<.01	2.94	=.08		
31.	Limbs	(proportion)	54 –	18.86	<.001	10.30	<.001		
32.	Clothing I		55 –	11.20	<.001	7.80	<.01		
33.	Lines	(quality)	63 C	20.96	<.001				
34.	Junctures	(quality)	64 E	7.91	<.01				

Note Under "Status", the number denotes the location of the item in the Man scale and the letter where it loads on Strommen's dimensions (1987), C = core dimension, E = elaborate dimension

Table A_W Items from Woman Scale which differentiate children at successive ages.

	Items		Status	4 vs. 5½		5½ vs. 7		7 vs. 8	
				χ^{2b}	p	χ^2	p	χ^2	p
1.	Neck		2 E					4.09	<.05
2.	Neck	(proportion)	3 E					5.55	<.05
3.	Eyes	(detail I: brow/lashes)	5 E					3.68	=.05
4.	Eyes	(proportion)	7 E					4.43	<.05
5.	Nose	(proportion)	10 E	4.01	<.05			4.10	<.05
6.	Nose	(bridge)	11 E			3.21	=.07		
7.	Chin/forehead		17 –	11.26	<.001				
8.	Hair I		19 C	10.31	<.001				
9.	Hair II		20 C	25.47	<.001				
10.	Arms		24 C	11.61	<001				
11.	Shoulders		25 E					4.26	<.05
12.	Arms	(at sides)	26 E	7.91	<.01	2.94	=.09		
13.	Fingers		28 C	6.08	<.05				
14.	Fingers	(detail: number)	29 E					8.11	<.01
15.	Fingers	(proportion)	30 E			3.25	=.07	4.15	<.05
16.	Hands		32 E	10.75	<.001				
17.	Legs		33 C	3.92	<.05				
18.	Feet I		35 C	16.01	<.001				
19.	Feet II	(proportion)	36 E	9.55	<.01	4.80	<.05		
20.	Feet	(placement)	40 –	4.84	<.05	4.33	<.05		
21.	Attachment	(limbs I)	41 C	26.82	<.001				
22.	Attachment	(limbs II)	42 E	13.04	<.001				
23.	Clothing		43 –	18.74	<.001	6.68	<.01		
24.	Sleeve I		44 –	13.04	<.001	9.06	<.01		
25.	Garb I	(feminine)	52 –	6.34	<.05	9.15	<.01		
26.	Trunk		55 C	9.05	<.01				
27.	Trunk	(proportion)	56 C	16.01	<.001				
28.	Head/trunk	(proportion I)	57 E	10.16	<.001				
29.	Face	(proportion)	58 E	4.84	<.05				
30.	Limbs	(proportion)	59 –	25.47	<.001	4.26	<.05		
31.	Arms/trunk	(proportion)	60 E			5.40	<.05		
32.	Lines	(quality)	64 C	9.55	<.01				

Note Under "Status", the number denotes the location of the item in the Woman scale and the letter where it loads on Strommen's dimensions (1987), C = core dimension, E = elaborate dimension

Table I_1 Means and standard deviations of raw scores from WISC-III^{UK} and WPPSI subtests, and from EFT.

	Age Levels									
	4 (n = 34)		5½ (n = 26)		7 (n = 28)		8 (n = 27)		Total (n = 115)	
Instruments	X	SD	X	SD	X	SD	X	SD	X	SD
Wechler Scales										
Verbal Scale										
1. Vocabulary	14	5	22	5	19	4	21	6	19	6
2. Information	11	3	14	2	8	2	10	2	11	3
3. Comprehension/ Similarities	11	4	17	4	9	3	13	5	12	5
Performance Scale										
1. Picture Completion	11	3	16	2	14	4	17	3	14	4
2. Block Design	9	4	14	3	23	11	29	12	18	11
3. Picture Arrangment/Mazes	10	6	18	5	17	9	22	8.	16	8
CEFT/PEFT[a]	13	4	9	4	13	5	18	5	13	5

[a] CEFT = Children Embedded Figure Test, PEFT = Preschool Embedded Figure Test

Table I_2 Means and standard deviations of standard scores from WISC-III^{UK} and WPPSI subtests and scales.

	Age Levels									
	4 (n = 34)		5½ (n = 26)		7 (n = 28)		8 (n = 27)		Total (N = 115)	
Wechler Scales (WISC-R/WPPSI)	X	SD	X	SD	X	SD	X	SD	X	SD
Verbal Scale										
1. Vocabulary	11	3	12	3	11	3	11	3	11	3
2. Information	12	3	10	2	9	3	10	3	11	3
3. Comprehension/ Similarities	12	3	11	2	10	3	12	4	11	3
Performance Scale										
1. Picture Completion	13	2	12	2	10	3	10	3	11	3
2. Block Design	12	2	11	2	10	4	10	4	11	3
3. Picture Arrangement/Mazes	12	3	12	2	11	4	11	4	12	3
Verbal Score (VIQ)	110	13	108	11	98	14	104	17	105	14
Performance Score (PIQ)	117	13	113	9	103	21	105	18	110	17
Total Score (FSIQ)	115	11	111	9	100	16	104	17	108	15

Appendix A-Tables

Table V, Accuracy of classification using FSIQ scores and GH standard scores from Man, Woman and Self drawings.

	GH-Man classification				GH-Woman classification				GH-Self classification			
FSIQ classification	<85	85-100	101-115	>115	<85	85-100	101-115	>115	<85	85-100	101-115	>115
<85 (n = 7)	2	2	3	0	2	4	1	0	2	3	1	0
85-100 (n = 23)	1	10	10	2	3	10	9	1	2	11	8	2
101-115 (n = 48)	7	15	18	8	12	17	14	5	11	18	13	6
>115 (n = 37)	4	14	12	7	5	15	11	6	7	12	14	4

Note Diagonal: Underlined frequencies represent accurate classification; outside diagonal: Number of cases incorrectly classified
GH-Man Scale: Accurately classified cases 37, hit rate 32%; Best classified IQ category 85-100 (10/23; 44%)
GH-Woman Scale: Accurately classified cases 32, hit rate 28%; Best classified IQ category 85-100 (10/23; 44%)
GH-Self Scale: Accurately classified cases 30, hit rate 26%; Best classified IQ category 85-100 (11/23; 48%)

Table V_2 Correlation coefficients between the GHDT and criterion measures by age level and GH Scale.

Instruments	MAN SCALE					WOMAN SCALE					SELF SCALE				
	Age levels					Age levels					Age levels				
	4	5½	7	8	Total	4	5½	7	8	Total	4	5½	7	8	Total
	n = 34	n = 26	n = 28	n = 27	N = 115	n = 34	n = 26	n = 28	n = 27	N = 115	n = 34	n = 26	n = 28	n = 27	N = 115
Wechler Scales (WISC-III^{UK}/WPPSI)															
Verbal Scale															
1. Vocabulary	-.03	.20	.32	.36	.20[a]	.13	.05	.33	.26	.21[a]	-.05	.31	.07	.29	.15
2. Information	.02	.27	.48[b]	.20	.14	.03	.19	.42[a]	.12	.10	.08	.19	.38[a]	.03	.11
3. Comprehension/ Similarities	-.08	.25	.60[c]	.16	.14	-.10	.12	.52[b]	.07	.06	-.08	.12	.48[b]	.09	.10
Verbal score (VIQ)	-.04	.27	.56[b]	.29	.19[a]	.03	.12	.51[b]	.18	.15	-.03	.24	.36	.17	.14
Performance Scale															
1. Picture Completion	.31	.17	.61[c]	.31	.33[c]	.31	.18	.61[c]	.28	.31[c]	.36[a]	.17	.44[a]	.32	.30[c]
2. Block Design	.25	.27	.67[c]	.15	.29[b]	.21	.30	.61[c]	.17	.28[b]	.36[a]	.36	.60[c]	.21	.35[c]
3. Picture Arrangement/Mazes	.43[a]	.41[a]	.38[a]	.17	.33[c]	.39[a]	.26	.49[b]	.20	.33[c]	.29	.34	.32	.27	.28[b]
Performance score (PIQ)	.46[b]	.43[a]	.63[c]	.26	.40[c]	.41[a]	.36	.67[c]	.28	.39[c]	.46[b]	.43[a]	.53[b]	.35	.40[c]
Total score (FSIQ)	.26	.41[a]	.70[c]	.31	.36[c]	.28	.27	.70[c]	.26	.32[c]	.28	.39[a]	.52[b]	.29	.32[c]
CEFT/PEFT	.43[a]	.36	.43[a]	.29	.40[c]	.32	.28	.41[a]	.18	.37[c]	.24	.35	.21	.31	.37[c]

Note: Correlations between the GHDT and Wechsler subtests and Scales are computed on standard scores
Correlations between the GHDT and PEFT/CEFT are computed on raw scores.
The coefficients in the columns on totals are partial correlations, controlling for age.
[a] p <.05; [b] p <.01; [c] p <.001

Table V_3 Correlation coefficients between the GHDT, and criterion measures by age level and GH Scale using raw scores.

	MAN SCALE					WOMAN SCALE					SELF SCALE				
	Age levels					Age levels					Age levels				
Instruments	4	5½	7	8	Total[1]	4	5	7	8	Total	4	5	7	8	Total
Wechler Scales (WISC-III/WPPSI)															
Verbal Scale															
1. Vocabulary	-.17	.32	.32	.32	.31[c]	-.01	.21	.33	.20	.25[b]	-.06	.40	.04	.25	.23[a]
2. Information	.13	.39[a]	.42[a]	.18	.30[c]	.13	.32	.41[a]	.09	.14	.20	.31	.25	.08	.15
3. Comprehension/ Similarities	.02	.39[a]	.58[c]	.15	.34[c]	-.02	.27	.56[b]	.04	.18	-.06	.25	.42[a]	.09	.20[a]
Performance Scale															
1. Picture Completion	.43[a]	.36	.54[b]	.22	.44[c]	.38[a]	.33	.54[b]	.16	.32[c]	.48[b]	.38	.29	.20	.31[c]
2. Block Design	.57[c]	.36	.51[b]	.20	.30[c]	.51[b]	.38	.49[b]	.22	.32[c]	.56[c]	.41[a]	.36	.30	.31[c]
3. Picture Arrangement/Mazes	.58[c]	.49[b]	.36	.18	.41[c]	.56[c]	.33	.45[a]	.21	.37[c]	.40[a]	.42[a]	.26	.30	.34[c]
CEFT/PEFT	.43[a]	.36	.43[a]	.29	.40[c]	.32	.28	.41[a]	.18	.37[c]	.24	.35	.21	.31	.37[c]

Note The correlations on column for totals are partial controlling for age.

This section deals with certain problematic aspects in the scoring of particular items of GHDT. These features along with scoring suggestions are illustrated by example drawings from the sample. Its item is first specified by its place in the scales, followed by a short description. For example M_3/W_3: Neck, two dimensions, suggests that a neck with correct proportions is the third item in both the Man and Woman scales. The example drawings are marked according to the following key:

✓: The drawing passes according to GH manual.

×: The drawing fails according to GH manual.

? : The drawing is a marginal or unclear case.

M3/W3: Neck, two dimensions

a. Gemma, 7.1 b. Allan, 8.3 c. Gemma, 7.1

• According to GH "outline of neck should be continuous with that of the head, of the trunk or of both" while when the neck is interposed as a pillar between head and trunk it does not get any credit.

? The above examples illustrate a case where the two criteria are irreconcilable. In terms of construction, continuity is satisfied since the neck and the trunk have been drawn with one stroke. However, in terms of appearance the drawings should fail.

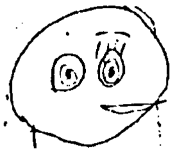
➪ **Suggestion**: Scorers should interpret the principle of continuity as one that refers to the appearance and not to the actual construction process.

a. Tamara, 6.0 b. Bobbie, 8.1 c-d. Nicola, 8.1

? In the manual, there are neither scoring instructions for the relative proportions between the neck and/or trunk (a, b) nor for the lack of co-ordination in the treatment of neck's lines (c, d) –incongruous stroke direction, asymmetrical attachment to the head.

➪ **Suggestion**: Scorers may be lenient on proportions and line junctures as long as both lines delineating neck contour converge at the head outline (a, b).

M_5/W_5: Brow or lashes



?

Luke, 5.11

• The item passes if brow, lashes or both are shown.
? It is unclear if both members of either pair should be present for the item to score.
➪ **Suggestion**: If both eyes are shown, both must have brows or lashes.

M_{12}/W_{14}: Lips two dimensions

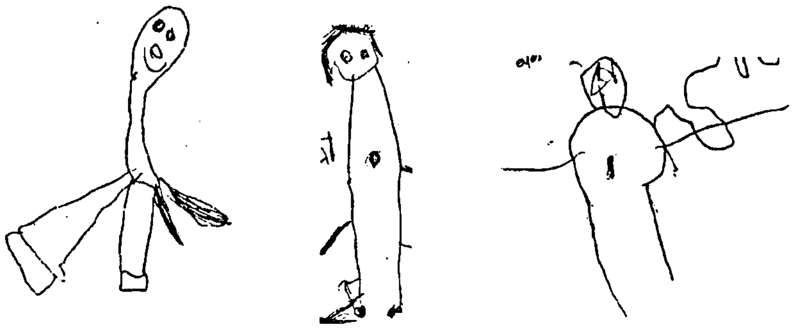
a. Sara, 7.5   b. Craig, 5.10   c. Lorna, 5.11   d. Jill, 7.3   e. Sara, 6.6

• The item's single definition "two lips should be clearly shown" is rather vague. According to the scoring examples an open elliptical region with or without "cupid's bow" never passes (a, b).
? However, when mouth is depicted as a single region it can be differentiated into upper and lower lips by the anchoring of denture at corresponding sides within the enclosed space (c, d).
➪ **Suggestion**: Credit those cases where teeth are attached to the mouth's outline. The "filling-in" of mouth with cross lines should not pass (e).

Katrina, 4.4 ✓

<u>Remark</u>: Keeping in mind that according to GH two parallel lines is a satisfactory representation of two lips in two dimensions, scorers should be cautioned against misinterpreting the lower lip for a boundary between a head and torso. Appearance can be misleading and children's description of their picture always deserves due attention.

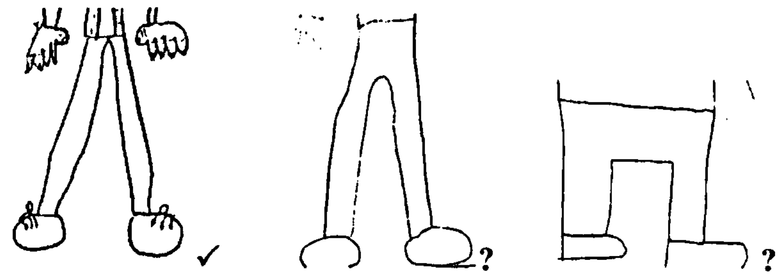
M46/W55: Trunk present



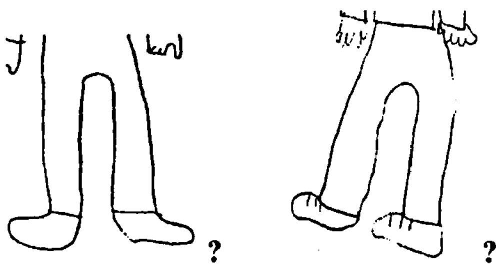
a. Sara, 4.4 b. Andrew, 4.3 c. Joanne, 4.0

Note: When there is no regional differentiation between the head and the trunk (a) or the trunk and legs (b), immature drawers can distinguish them by juxtaposing characteristic features at opposite ends (belly button, arms vs. feet). Both items (trunk, legs) should pass.

M36: Hip



a. Naomi, 7.11 b. James, 7.8 1 c. Allan, 8.3



d. Lucy, 7.4 e. Lorna, 8.3

• According to the instructions, the inner lines of the two legs should meet at a point of junction with the body, while legs placed as far apart as possible never score.
? Insufficient guidance is given for those cases where the legs are continuous with the body, but they are placed far apart without forming an acute angle (c). Uncertainty results as to whether legs should meet at a single point of junction or the later principle can be overruled as long as divergence of legs is satisfied (b).
➪ **Suggestions**: Lines of legs should either meet at a single point of junction with the body (a) or perceptibly diverge from crotch to feet (b). Parallel legs connected with a perpendicular line (c) or an arc (d, e) should never score. The instructions can be also complemented by the note that i) the trunk must be two dimensional, otherwise two legs meeting at a point of junction with the body is the only possible solution and ii) linear legs may be credited as long as they form an acute angle with body's axis.

W_{34}: Hip


?


?


?

a.Gemma, 7.1 b. Anna, 8.8 c. Kirstin, 8.3

• The principal axis of the legs should form an angle.
? Instructions and scoring examples are incomplete since they cover only the cases where the figure is depicted with feminine garment.
➪ **Suggestion**: Following the criteria from man's scale, a figure in slacks/trousers scores as long as the crotch is indicated.

$M_{24\text{-}28}$/$W_{28\text{-}32}$: Hands/Fingers (present, proportions, details).


?

Allan, 8.6

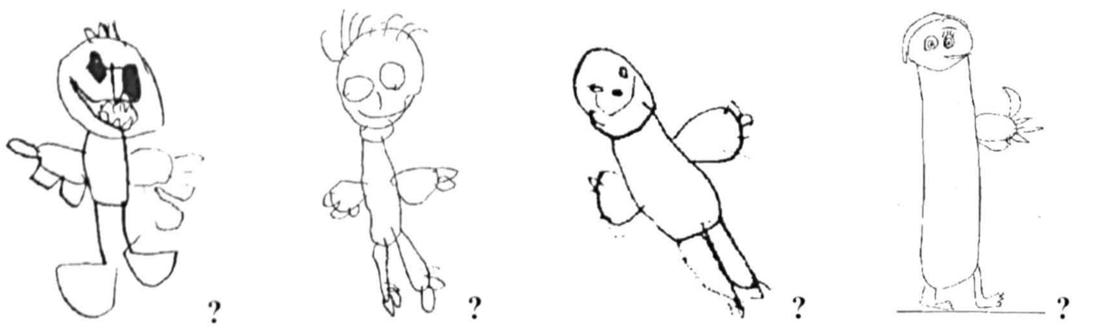
? Nowhere in the description of those items is there a note referring to those drawings where one of both hands is concealed behind the back or inside pockets.
➪ **Suggestion**: When there is a clear intention to depict the figure in a posture, score the 5 items on the basis of hands shown. Otherwise, if two hands are present, characteristics should appear on both.

M_{29}: Wrist/Ankle shown

a. Prue, 5.7 b. Rachel, 5.3 c. Kirsty, 6.0

• GH's criteria are the separation of the limbs from the wrist or ankle by some method other than of a line drawn across the limb.
? There is not an explicit account for cases where linear limbs protrude from the garment (a, b). Neither is there any guidance for drawings where only one member of either pair passes (c).
➪ **Suggestion**: Limbs must be two dimensional. If two hands or feet are shown, the characteristic should appear on both (c).

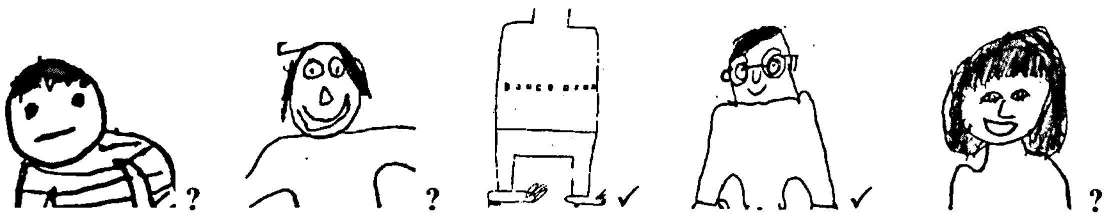
M_{30}/W_{24}: Arms present



a. Arran, 5.7 b. Michael, 5.6 c. Mark, 7.2 d. Luke, 5.11

Difficulty arises in interpreting the representational status of two dimensional forms, extending from the upper part of torso (a, b, c, d).
? The Manual made no mention of these cases, which are often produced by immature drawers.
➪ **Suggestion**: Score them as either arms or hands.

M31/W25: Shoulders I



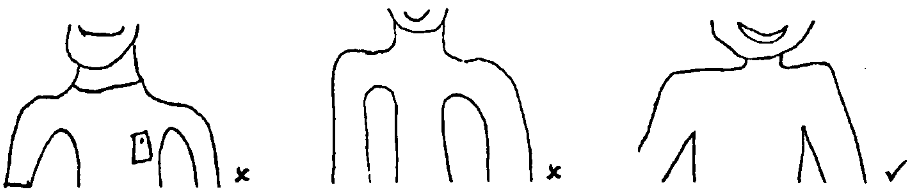
a. Ben, 5.9 b. Nicola, 8.8 c. Allan, 8.3 d. Raymond, 7.11 e. Lorna, 8.3

• The item passes if "there is a change in the direction of the outline of the upper part of the trunk which gives an effect of concavity". A square trunk does not score unless the corners have been rounded (c, d).

? The criterion might not be satisfied at both sides.

➪ **Suggestion**: Failure to reverse the orientation of trunk's outline at both sides below the head should be penalised (a, b, e).

M32: Shoulders II



a. David, 8.0 b. Stephanie, 8.3 c. Gary, 8.3

• According to the instructions, shoulders must be continuous with neck and arms and the armpit must be shown if the arm is held from the body.

Note: Only from one scoring example is there an indirect reference to the angular depiction of the armpit as a scoring principle (c).

M33/W26: Arms at side

Gemma, 8.5

• Arms perpendicular to the vertical axis of the body never passes. The angle should be no more than 10°.

? Sometimes arms are at resting posture, when the angle with trunk's axis is 0° degrees.

➪ **Suggestion**: Give credit to these depiction since they also deviate from the stereotypical stretching of the arms, seen in immature drawings.

M35/W33: Legs present

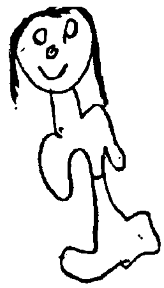
a. Kristin, 3.10

b. Steward, 4.4

Note: Peculiarities of construction may conceal the presence of legs.
Pencil slip produces a 2-D trunk when legs are intended (a).
The place of hair and legs is reversed in an upside down depiction. Appearance (tadpole with many legs and little hair) distorts intention (tadpole with two legs and lots of hair) (b).


?

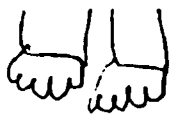
a. Mitchell, 6.11


?

b. Megan, 5.7

• Any figure clearly intended to represent legs receives credit. When long gown hides legs this item passes.
? GH do not specify the criteria that are sufficient for an adequate depiction of a long gown.
➪ **Suggestion**: Do not apply the second rule in immature drawings where trunk and garment are undifferentiated (a).

M43/W37: Feet detail


?

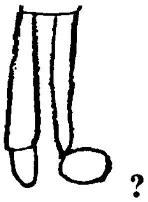
a. Jennifer, 6.7


?

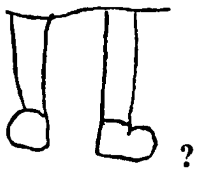
b. Luke, 5.11

• GH instructions give credit to any shoes that show some ornamentation.
? The scorer faces difficulty scoring barefoot depiction, since is not mentioned in the item's description.
➪ **Suggestion**: If the number of toes and proportions are accurate, the item may pass.

M_{40}/W_{36}: Feet proportion



a. Lean, 7.1



b. Luke, 5.11

• The horizontal measure of the feet must be greater than the vertical one in profile depiction. The reverse applies in front view.

? It is not clear whether one foot is sufficient for the item to pass (b). In addition feet drawn in mixed views are not discussed (a).

➪ **Suggestion**: If both feet are shown, the criteria should be satisfied for both. Feet in mixed views should not score.

M_{44}/W_{41}: Attachment I

Josh, 4.6

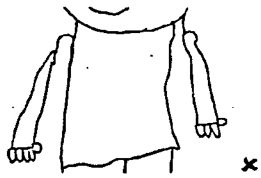
• Both arms and legs must be attached to the trunk at any point.

? The manual does not discuss the case where the arms and legs are absent but hands and feet are attached to the trunk.

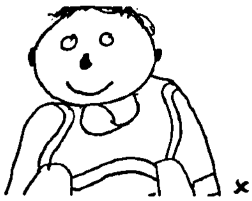
➪ **Suggestion**: Since the item differentiates between tadpole and conventional drawers, credit can be given as long as hands and feet are extensions of limbs.

M_{45}/W_{42}: Attachment II

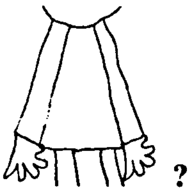
a. Lean, 7.1



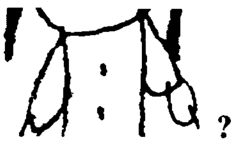
b. Nyree, 7.2



c. Heather, 8.5



d. Karen, 8.4



e. Laura, 7.4

? It is difficult to reconcile instructions in the first paragraph, which seem to be lenient, "When no neck is present the arms should be attached to the upper part of the trunk", with the second paragraph which appears much stricter, "the attachment must be exactly at the point which should be indicated as the shoulders". Examples (a), (b) and (c) fail based on the additional criterion of symmetrical attachment. However drawings like (d) and (e) are difficult to score.

➪ **Suggestion**: When there is a shared boundary between the arms' inner line and the torso's outline, the arms' outer line should be attached exactly at the junctures of head with the trunk.

W_{48}: Waist

a. Bobbie, 8.1

b. Nicola, 8.8

c. Loise,6.7

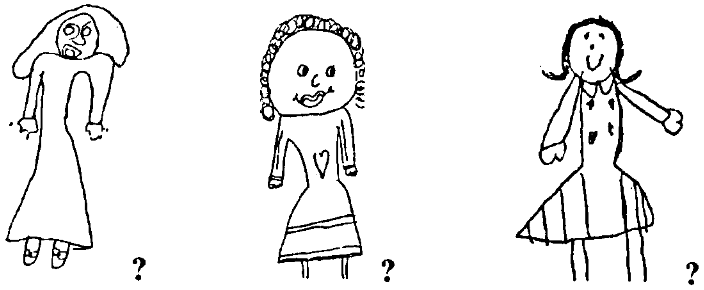
d. Laura, 7.4

• The direction of the body contour must change perceptibly at/or below waist (a).

? Profile depiction (b) and lack of angular co-ordination (c, d) are not discussed in the manual.

➪ **Suggestion**: Body contour should form obtuse angles at both sides (or at one in profile depiction). Do not penalise lack of angular symmetry in front view drawings.

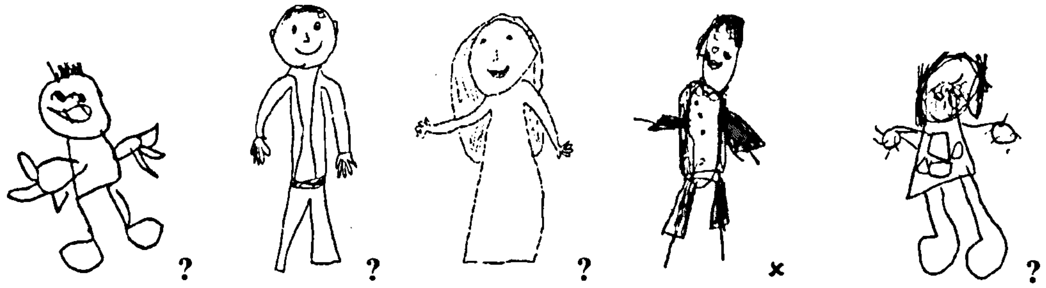
W_{52}: Garb feminine



a. Loise, 6.7 b. Georgia, 5.5 c. Laura, 7.4

? Instructions for younger children, "body must appear in two different segments" are surprisingly stricter than those for older ones, "Credit any dress or skirt".
➪ **Suggestion**: Credit single region bodies when the direction of the body's contour changes to differentiate upper from lower part (a, b and c).

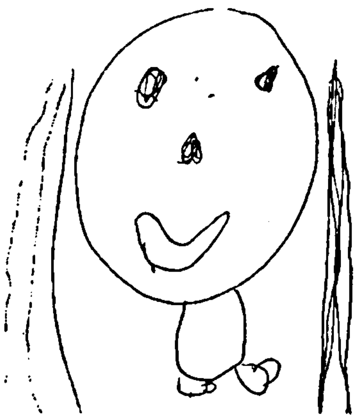
M_{54}/W_{59}: Limbs two dimensions



a. Arran, 5.7 b, c. Jenny, 8.4 d. Tamara, 6.0 e. Claire, 5.4

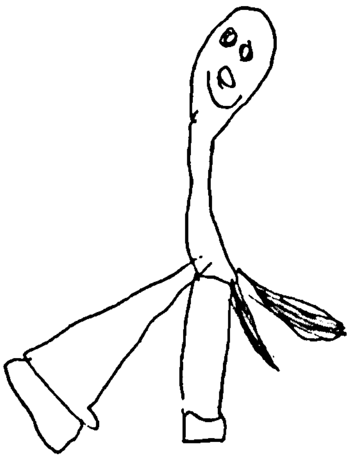
• The item receives credit when the length of the arms and the legs is greater than the width.
? There is not guidance as to how we score the item
i) in the absence of either of member of the limbs (b, c) and
ii) in the absence of arms but presence of hands and legs in two dimensions (a).
➪ **Suggestion**: Both pairs of limbs must satisfy the scoring criteria. Either pair can be substituted by its extending feature (hands/feet) if it is in 2-D.
<u>Note</u>: Linear limbs protruding from two dimensional garments do not pass (d). When arms are linear and hands are in two dimensions the item fails (e).

Appendix A-Sample drawings

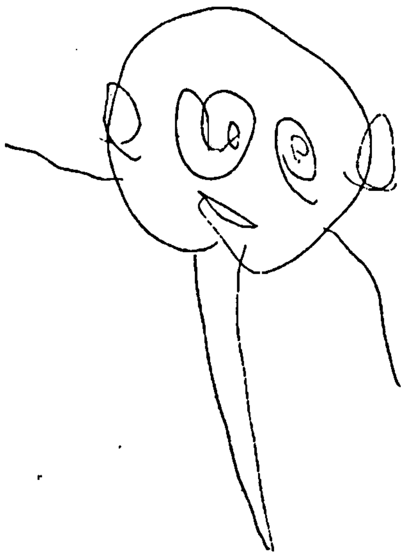


Louise, 4.5, Woman (89)[1]

Caroline, 4.2, Self (78)



Sara, 4.4, Man (100)



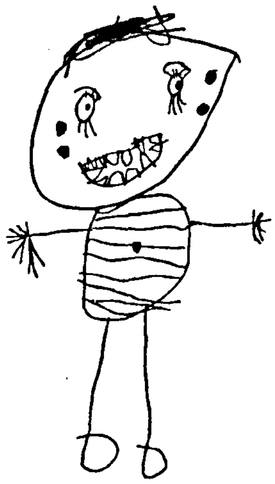
Andrew, 4.6, Man (87)

Jake, 4.2, Self (91)

Tim, 3.10, Self (100)

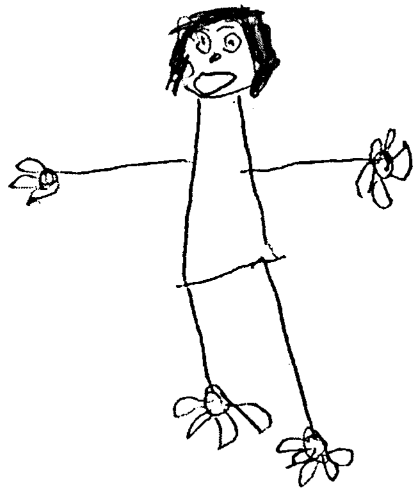
[1] The figure next to the type of drawing each child made is the GH standard score that the child earned.

Rebecca, 4.3, Man (115)



Scott, 4.8, Woman (87)

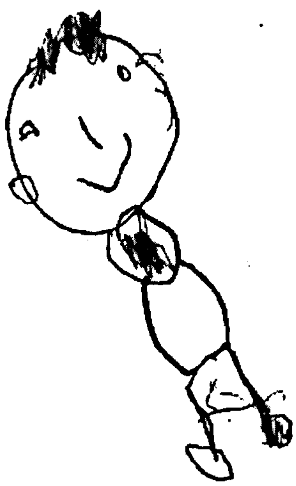
Heather, 4.5, Self (119)



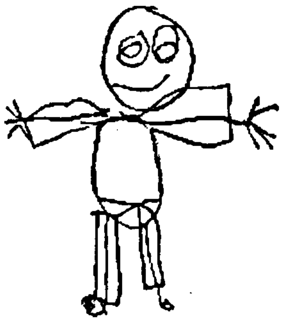
Any, 4.7, Woman (96)

Jake, 4.2, Woman (91)

Scott, 4.8,
Man (104)



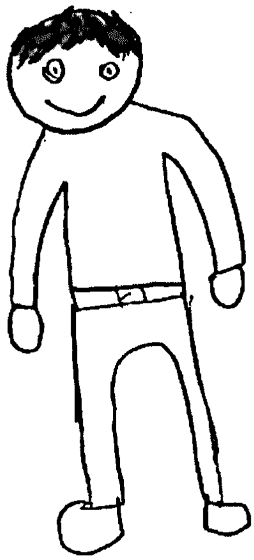
Prue, 5.7, Self (117)



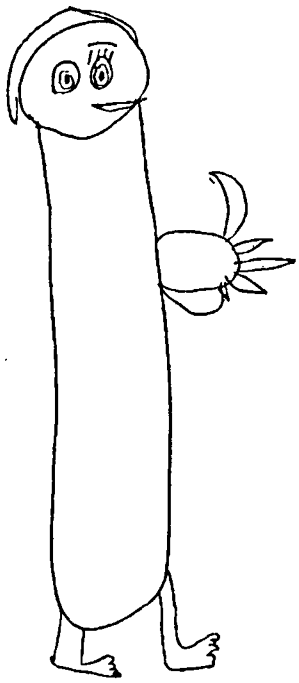
Rachel, 5.3, Man (99)

Adam, 5.9, Woman (111)

Ben, 5.9, Self (125)



Luke, 5.11, Man (107)



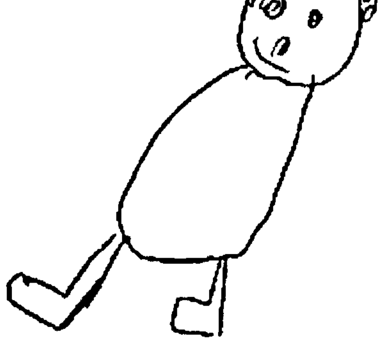
Stacey, 5.7, Woman (82)

Georgia, 5.5, Woman (139)

Daniel, 5.4, Man (119)

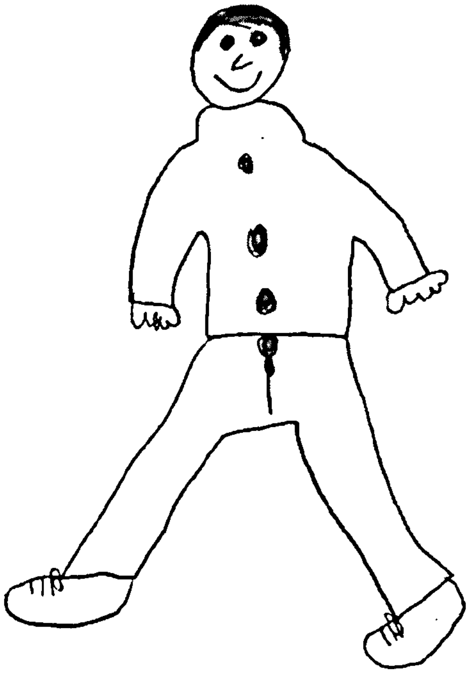
Lorna, 5.11, Self (107)

Linsday, 6.2, Man (115)



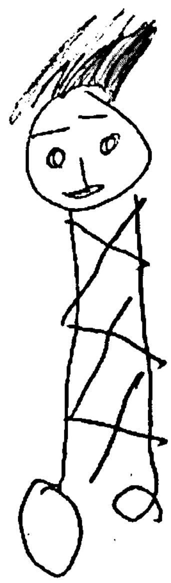
Graham, 5.2, Woman (86)

Amy, 5.8, Man (102)

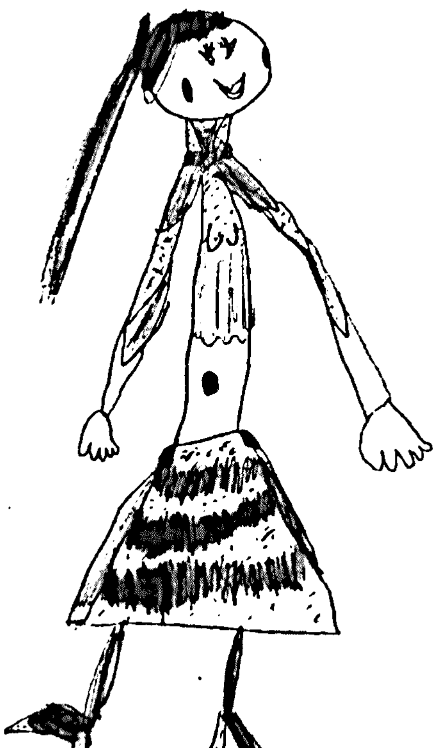


Alistair, 7:2, Man (112)

Jennifer, 6.7, Man (113)

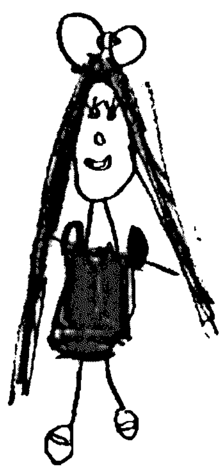


Michelle, 6.11, Man (75)



Jody, 6.11, Woman (139)

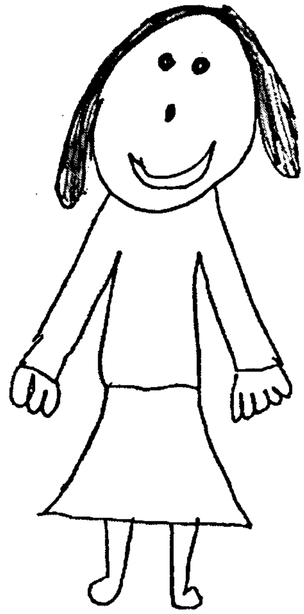
Luke, 6.0, Self (115)

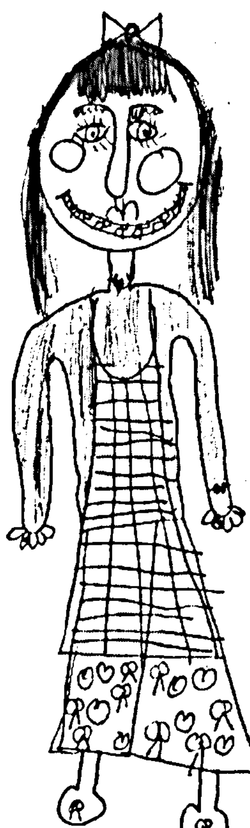


Tamara, 6.0, Self (103)

Alistair, 7.2, Woman (118)

Gillian, 7.1, Woman (106)

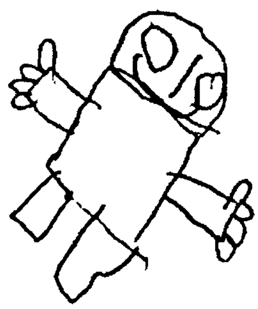




Jill, 7.3, Self (102)

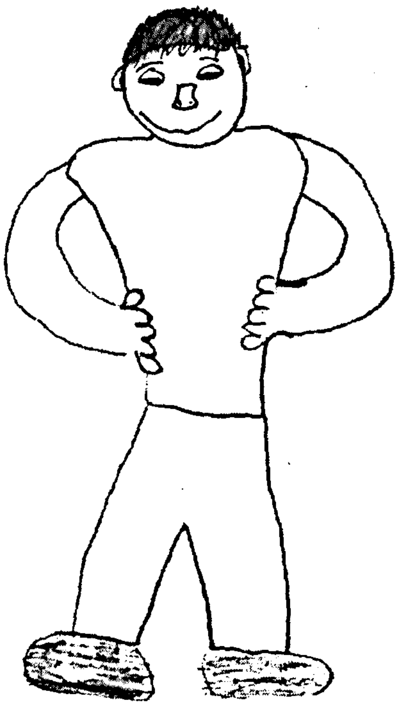
Jody, 6.11, Man (125)

Blair, 7.3, Man (97)

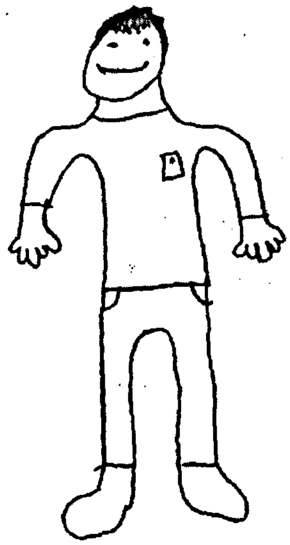


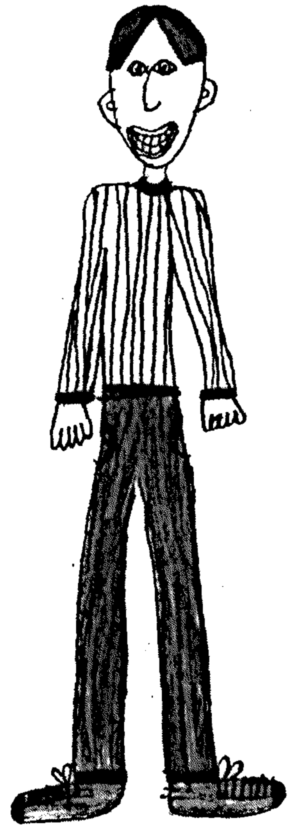
Joseph, 7.6, Man (77)

Andrew, 7.10, Self (114)



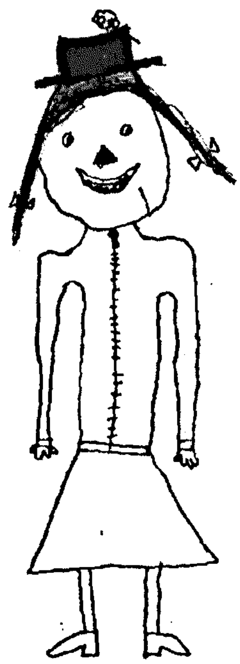
David, 8.8, Self (104)



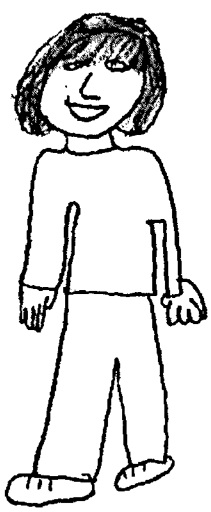


Anna, 8.8, Man (135)

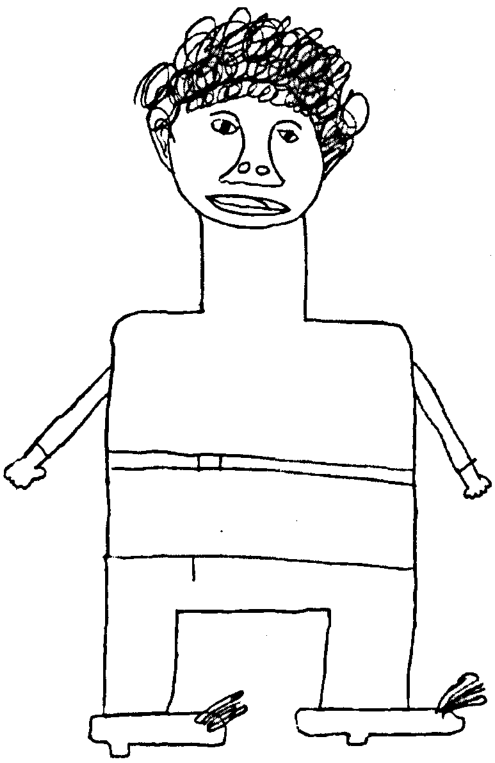
Kirsty, 8.3, Self (103)



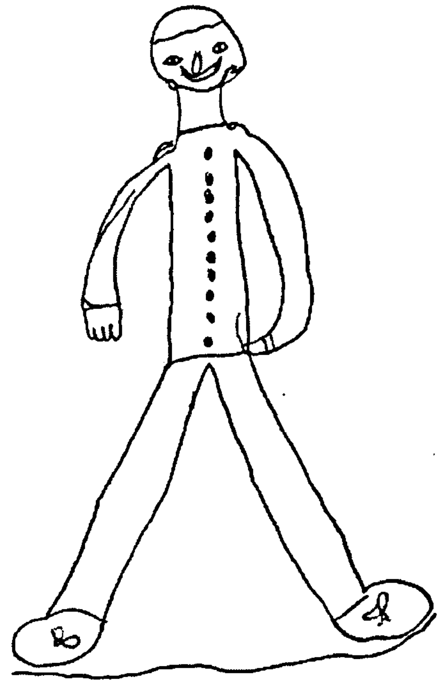
Christopher, 8.6,
Woman (114)



Lorna, 8.3, Self (106)

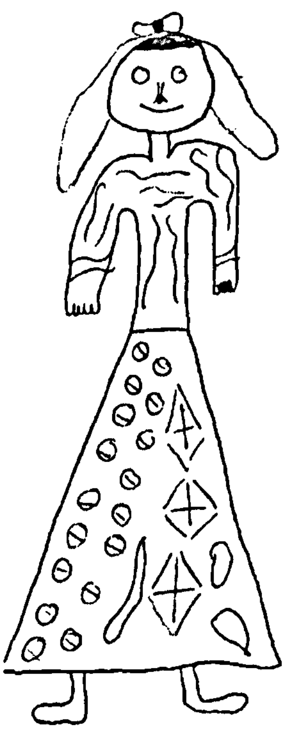


Allan, 8.3, Woman (106)

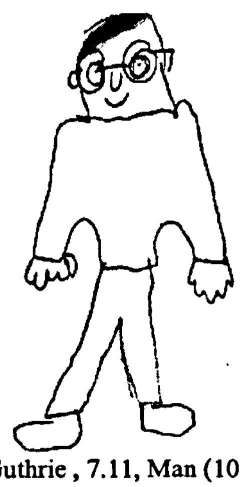


Allan, 8.6, Man (108)

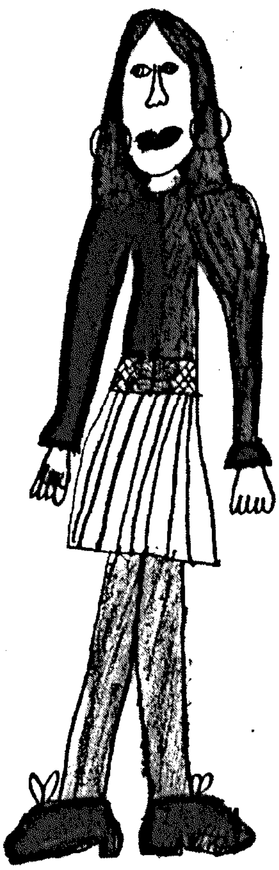
Naomi, 7.11, Man (125)



Heather, 8.5, Woman (108)

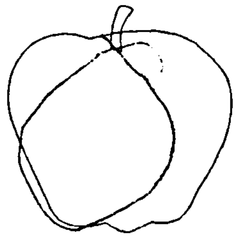


Guthrie , 7.11, Man (101)

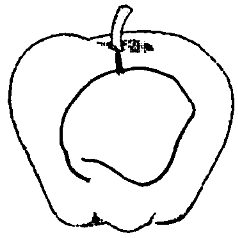


Anna, 8.8, Woman (139)

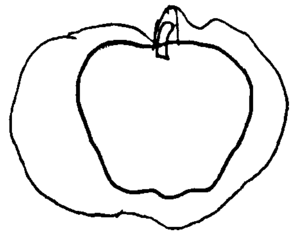
APPENDIX B



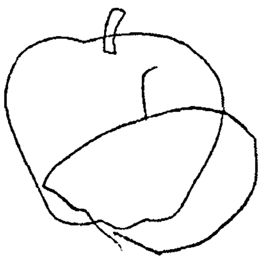
Enclosure: Sara, 3.7



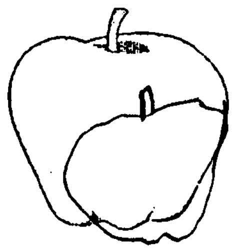
Sean, 5.6



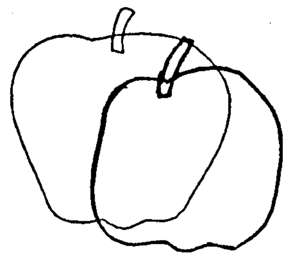
Sara, 7.5



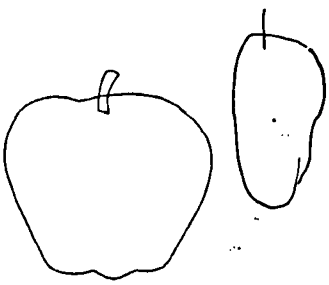
Overlap: Caroline, 4.2



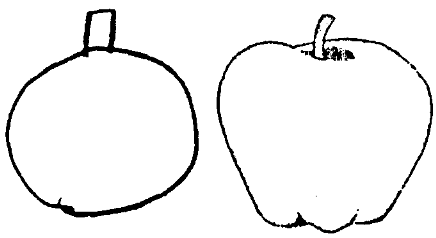
Ben, 5.8



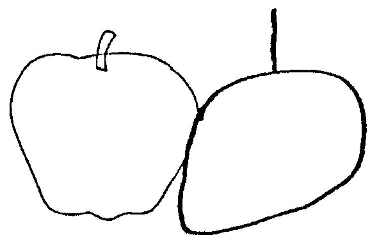
Gillian, 7.1



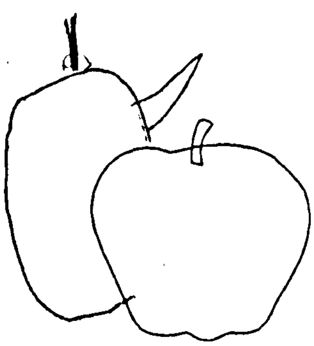
Segregation: Aileen, 3.8



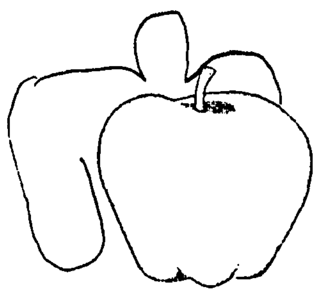
Daniel, 5.4



Attachm.: Heather, 4.4

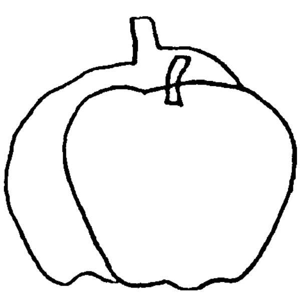


P. occlusion : Louise, 4.5

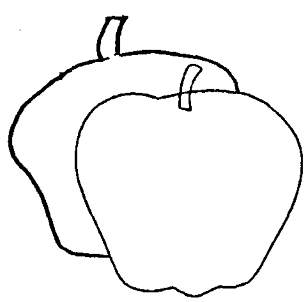


Prue, 5.6

Gemma, 7.0

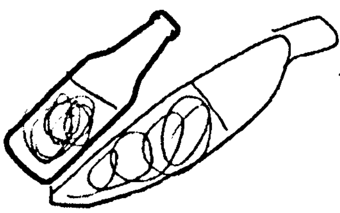
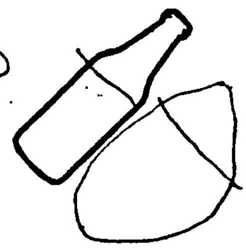
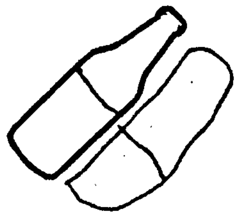


Fiona, 7.8



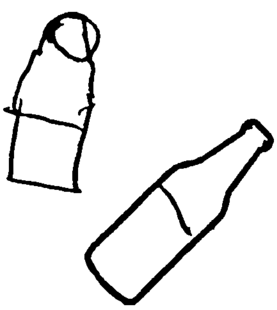
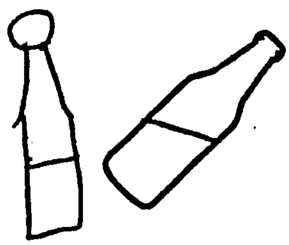
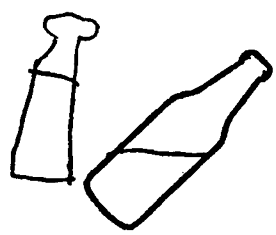
Caroline, 4.2 (Attachment & scribbles)

Separations

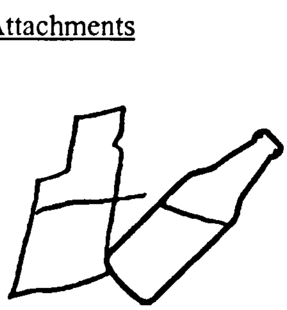
Context- driven solutions

Louise, 4.5 Amy, 4.7 Michelle, 7.0

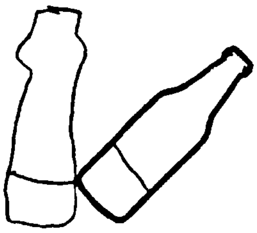
Sean, 5.6 (Excessive) Allan, 8.6 (Compromise) Linsday, 6.2 (Horizontal)

Attachments

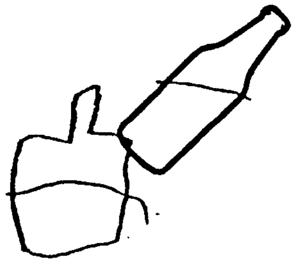


Compromise solutions

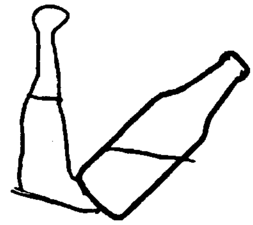
Louise, 5.3 Stacey, 5.6 Ewan, 8.3



Nyree, 7.3 (Excessive)

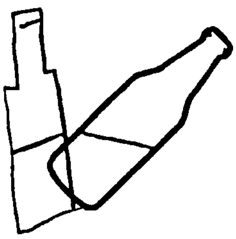


Michael, 5.6 (Horizontal)



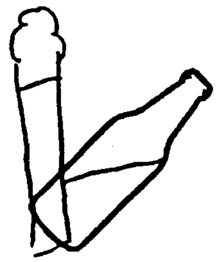
Emma, 7.0 (Horizontal)

Overlaps



Gemma, 5.6

Jennifer, 6.7

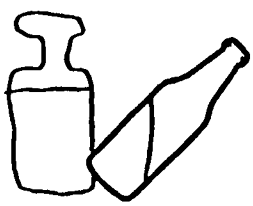


Gemma, 8.0

Compromise solutions

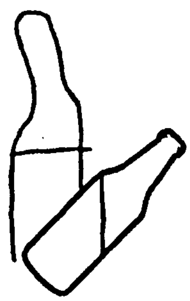
Partial Occlusions

A. With excessive solutions(>50)



Stephanie, 8.3

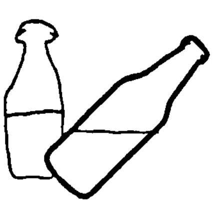
Naomi, 7.9



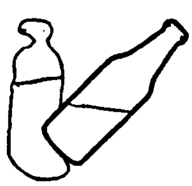
Kirsty, 8.3

B. With horizontal WLs (0°-11°)

Georgia, 5.5



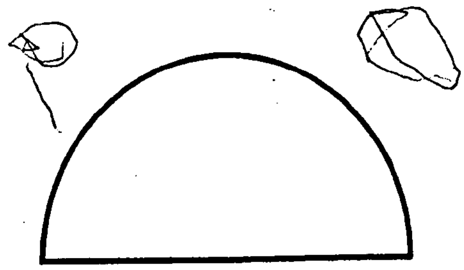
Alistair, 7.3



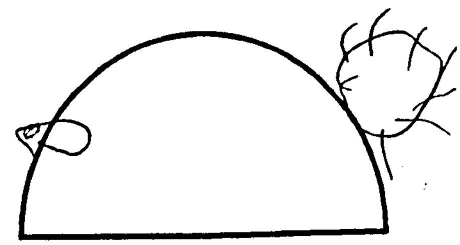
Jill, 7.3

David, 8.0

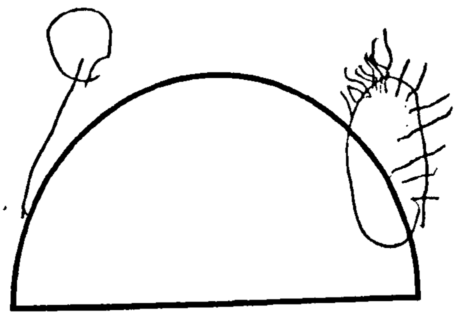
Anna, 8.8



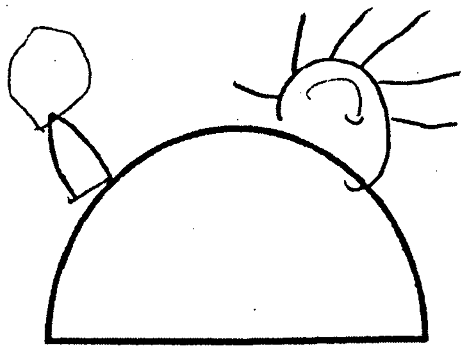
Rebecca, 3.9 (Tilted tree & segregation)



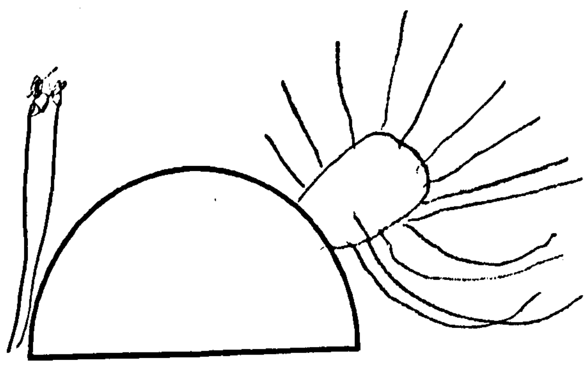
Caroline, 4.2, (Tilted tree & attachment)



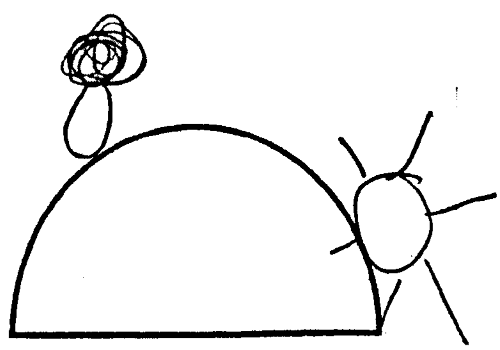
Elsa, 4.1 (Tilted tree & overlap)



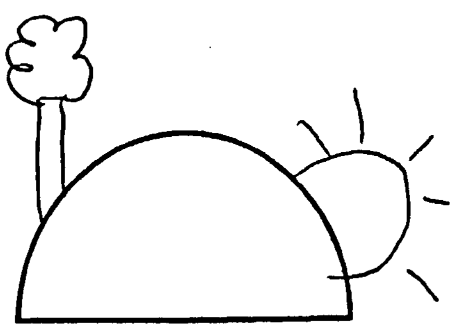
Natalie, 5.3 (Tilted tree & p. occlusion)



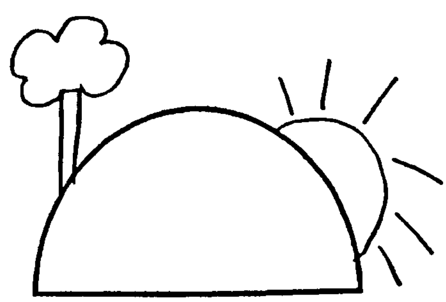
Louise, 4.5 (V. tree & p. occlusion)
Note: Trees aligned with mountain's base
were not included in analyses



Sara, 5.7 (V. tree & attachment)

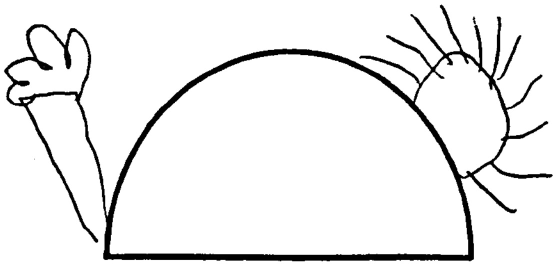


Andrew, 5.6 (V. tree & p. occlusion)

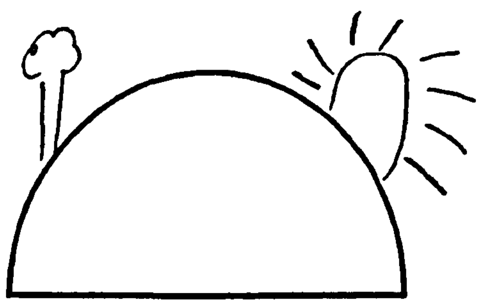


Georgia, 5.5 (V. tree & p. occlusion)

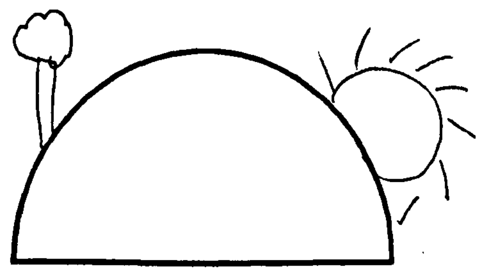
Joseph, 7.5 (Tilted tree & p. occlusion)



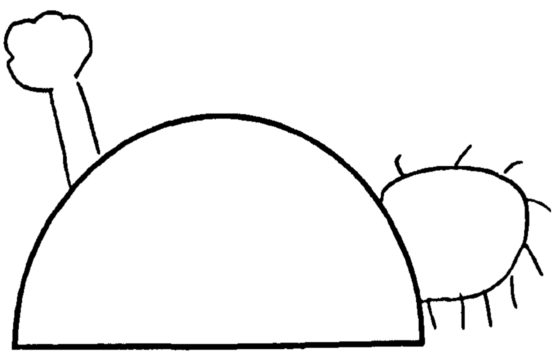
Mark, 7.3 (Tilted tree & p. occlusion)



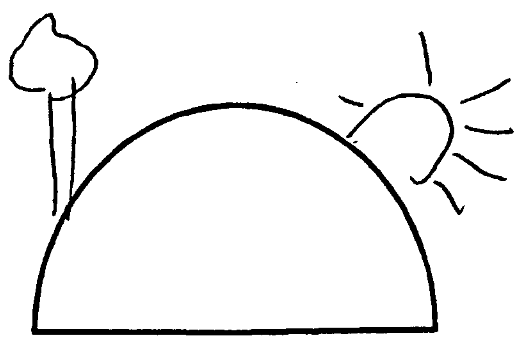
Lucy, 7.3 (V. tree & p. occlusion)



Gillian, 7.1 (V. Tree & p. occlusion)

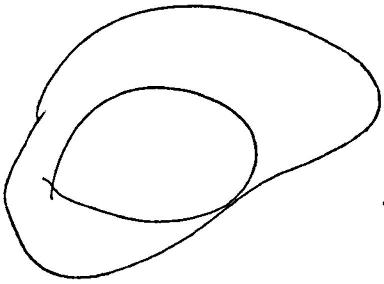


Gemma, 8.0 (Tilted tree & P. occlusion)



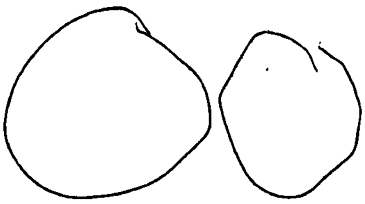
Allan, 8.6 (V. tree & p. occlusion)

Andrew, 7.8 (V. tree & p. occlusion)

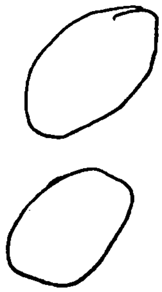


Josh, 4.6 (MO, Enclosure)

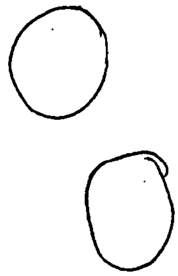
Segregations



Elsa, 4.1 (MO)

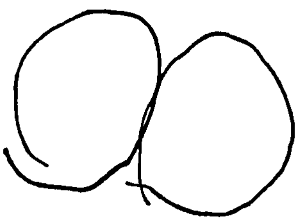


Caroline, 4.2 (FO)

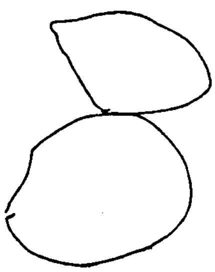


Susan, 6.8 (MO)

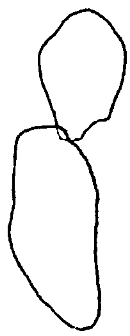
Attachments



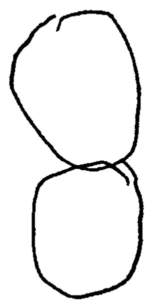
Arran, 5.6 (MO)



Sara, 4.4 (FO)

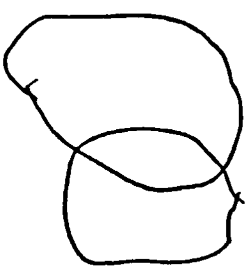


Kirsty, 4.2 (MO)

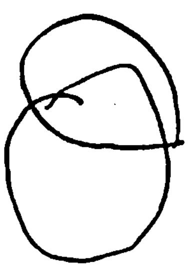


Joseph, 7.5 (MO)

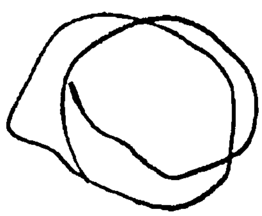
Overlaps



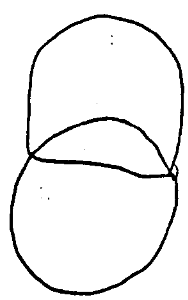
Katrina, 4.8 (MO)



Sarah, 4.3 (MO)

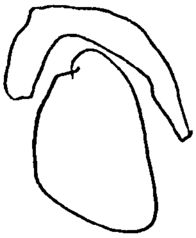


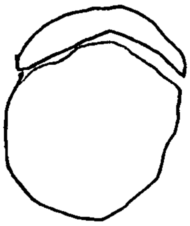
Sean, 5.6 (MO)

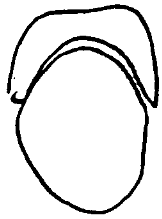


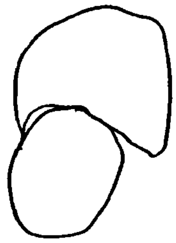
Nicola, 8.7 (MO)

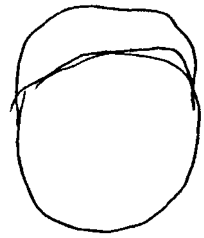
Incomplete forms with segregation and attachment (FO)


Joseph, 7.5

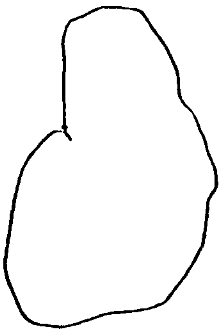

Jennifer, 6.1

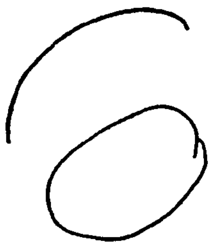

Ruth, 5.9

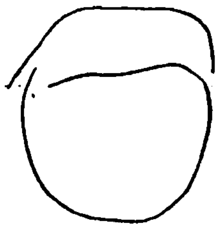

Louise, 4.5

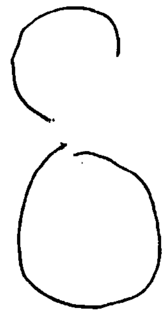

Graham, 7.0

Incomplete Partial Occlusions

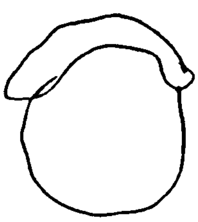

Greg, 5.8 (MO)

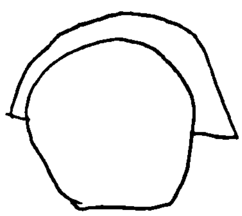

Andrew, 4.2 (FO)

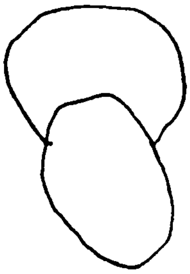

Megan, 5.6 (MO)

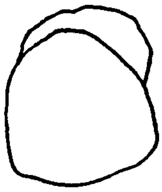

Graham, 5.3 (FO)

Partial Occlusions

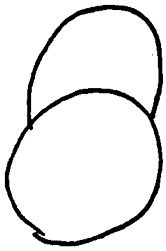

Sara, 7.1 (FO)

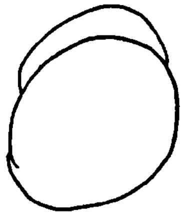

Sara, 7.5 (FO)


(MO)

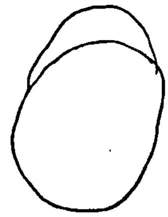

(FO)

Christopher, 8.5

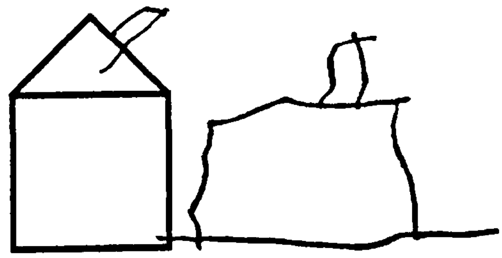

(MO)

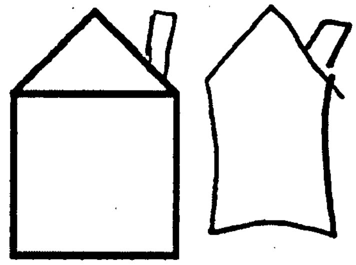

(FO)

Kirsty, 8.3

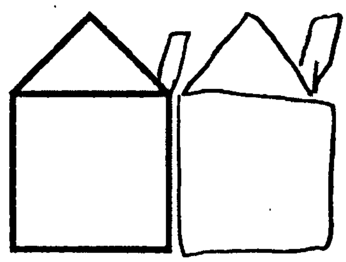

(MO)
Nicola, 8.3

Colin, 3.8 (Tilted chimney & segregation)

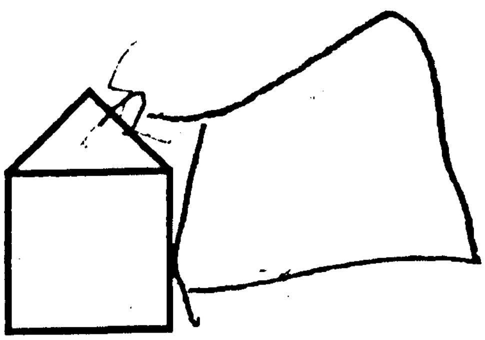




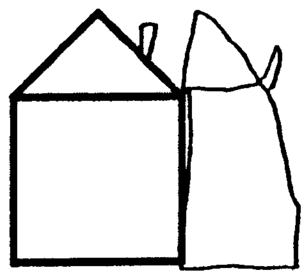
Prue, 5.6 (Tilted chimney & segregation)



Cory, 5.5 (Tilted chimney & segregation)

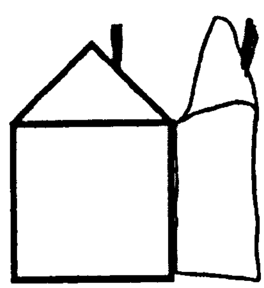


Joanne, 4.0 (Tilted chimney & attachment)

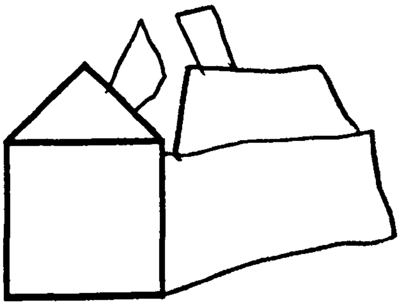


Tara, 6.0 (V. chimney & attachment)

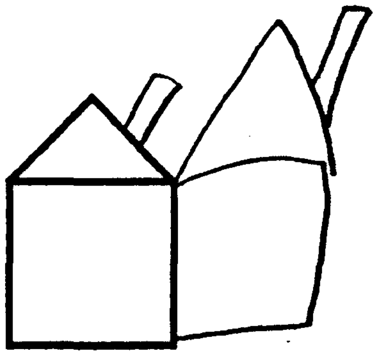
Raymond, 8.0 (V. chimney & attachment)



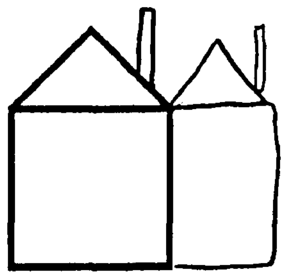
Kirsty, 6.7 (V. chimney & overlap)



Louise, 4.5 (Tilted chimney & p. occlusion)

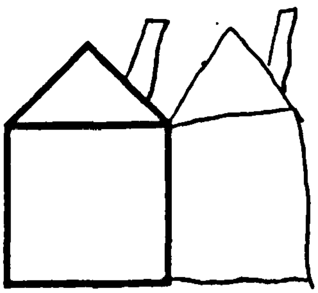


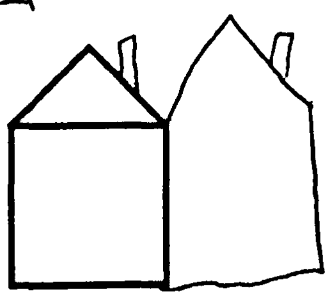
Emma, 6.2 (Tilted chimney & p. occlusion)



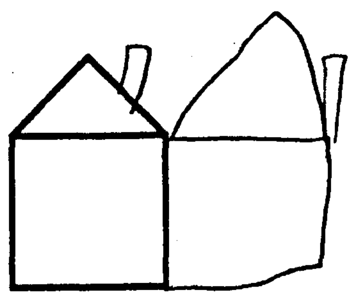
Jane, 5.6 (V. chimney & p. occlusion)

Gemma, 7.2 (Tilted chimney & p. occlusion)

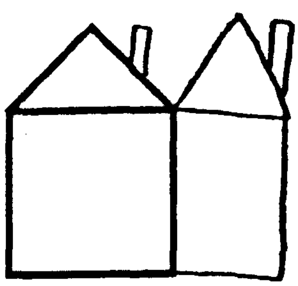




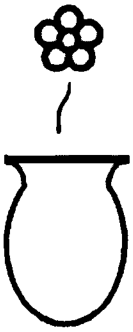
Emma, 7.0 (V. chimney & p. occlusion)



Liam, 7.9 (Tilted chimney & p. occlusion)



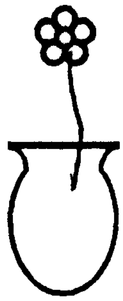
Donald, 8.3 (V. chimney & p. occlusion)



Andrew, 4.3 (Segregation)

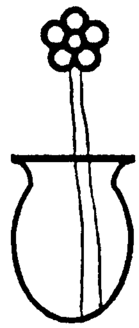
Overlap

Katrina, 4.8



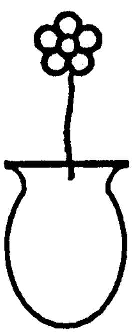
Amy, 4.7

Rachel, 5.3

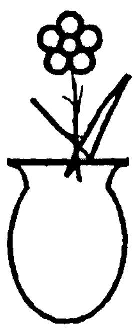


Ciaran, 6.7

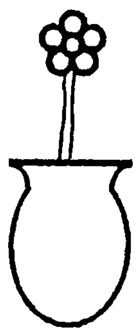
Partial Occlusion



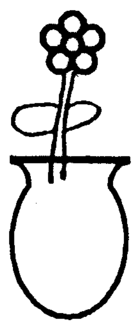
Claire, 4.2



Luke, 5.9

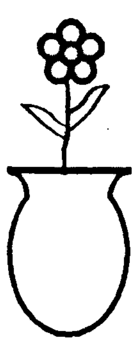


Georgia, 5.5

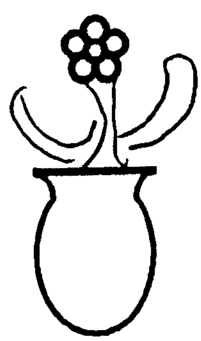


Jody, 7.0

Kirsty, 8.3

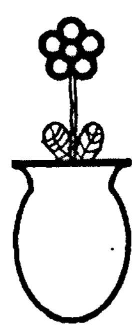


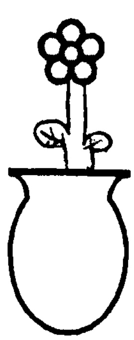
Jody, 7.0



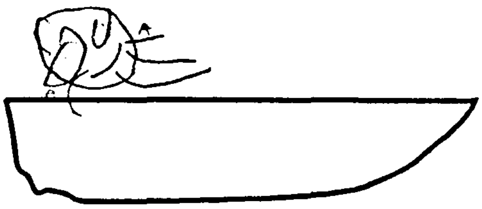
Karen, 8.3

Nicole, 8.3

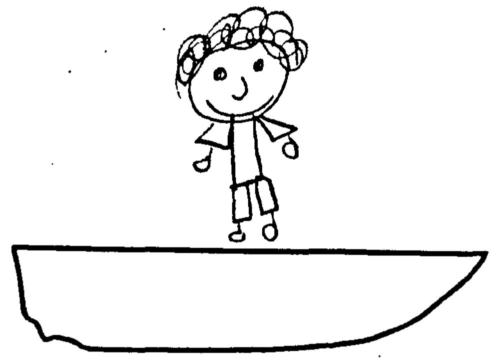




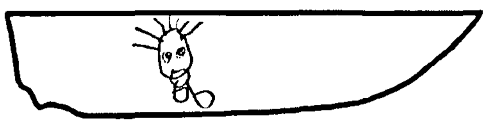
Gemma, 7.3



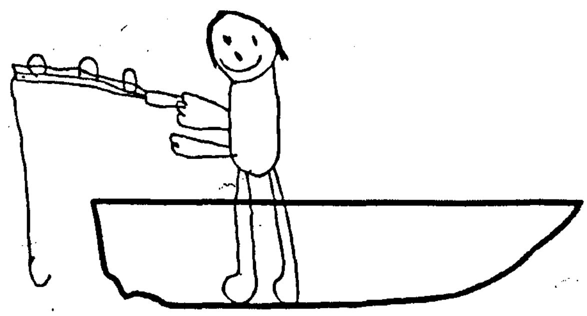
Louise, 4.5 (Segregation)



Prue, 5.6 (Segregation)

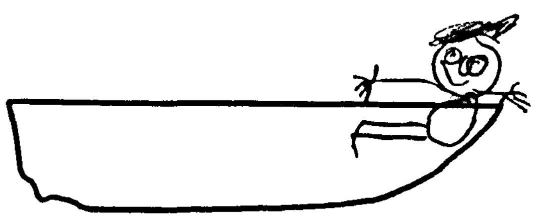


Rebecca, 3.9 (Enclosure)

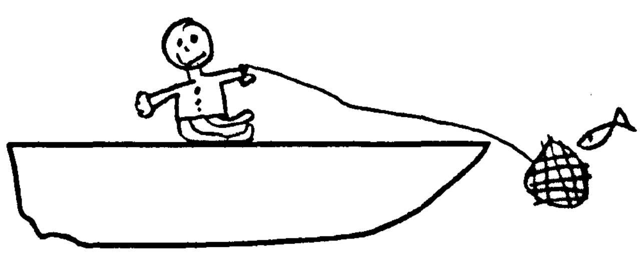


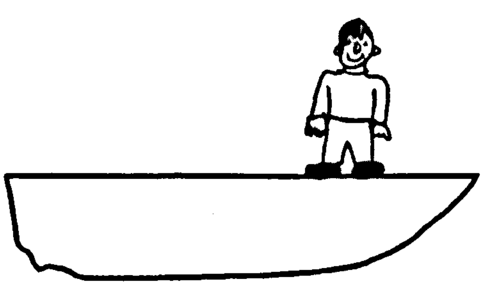
Gemma, 8.0 (Overlap)

Rachel, 5.3 (Overlap)

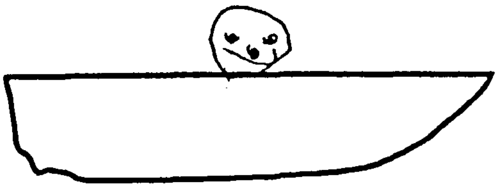


Louise, 7.4 (Attachment)

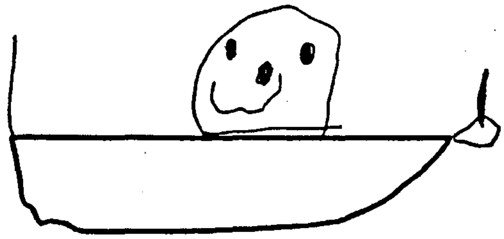




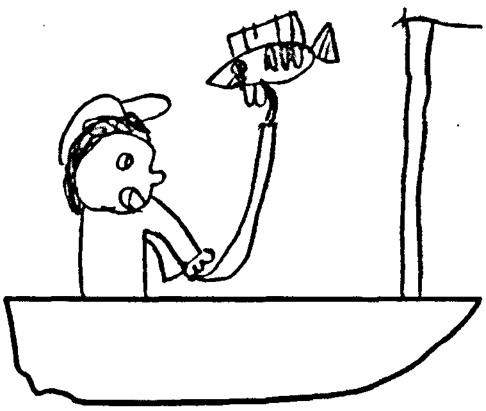
Anna, 8.8 (Attachment)



Andrew, 4.2 (P. occlusion)

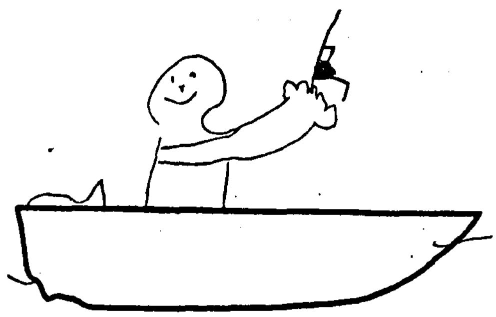


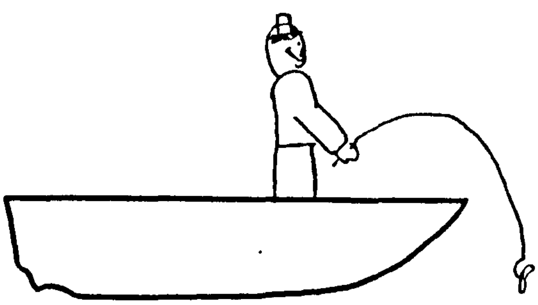
Sarah, 4.8 (P. occlusion)



Georgia, 5.5 (P. occlusion)

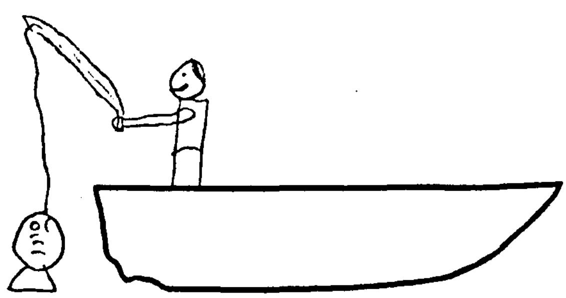
Alistair, 7.2 (P. occlusion)





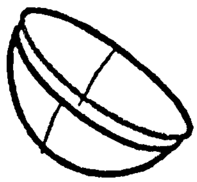
Gemma, 7.3 (P. occlusion)

Christopher, 8.5 (P. occlusion)

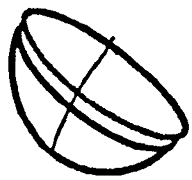


Josh, 3.8 (Scribbles)

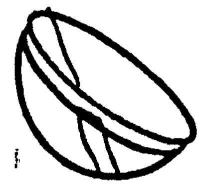
Excessive solutions (>50°)



Louise, 4.5

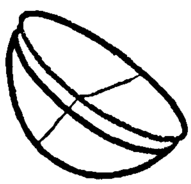


Nyree, 7.3

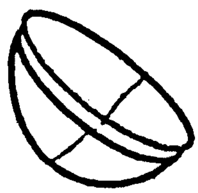


Stephanie, 8.3

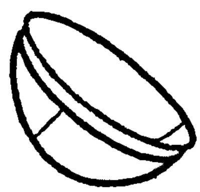
Context-driven solutions (34°-50°)



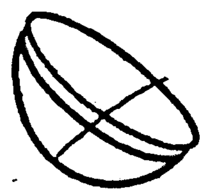
Sara, 4.5



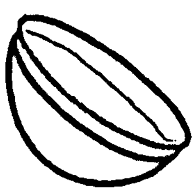
Ben, 5.8



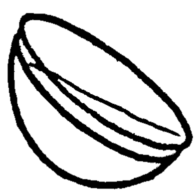
Amy, 5.8



Lindsey, 7.9



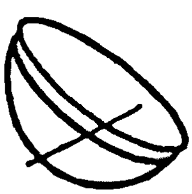
Claire, 4.2



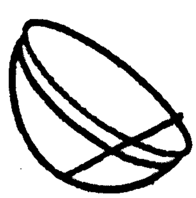
Lindsey, 6.2

Sara, 7.5

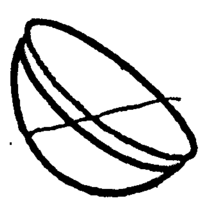
Compromise solutions (12°-33°)



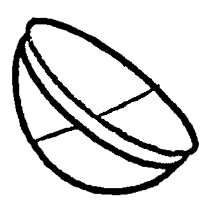
Amy, 4.7



Daniel, 5.4

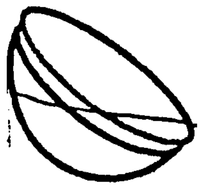


Rachel, 5.3

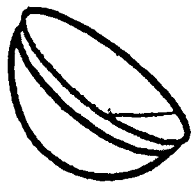


Jody, 7.0

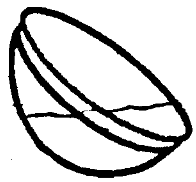
Horizontal water levels (0°-11°)



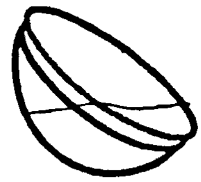
Louise, 5.3



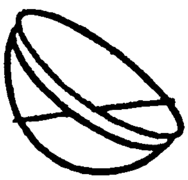
Sean, 5.6



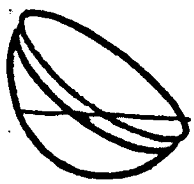
Louise, 6.6



Jennifer, 6.7

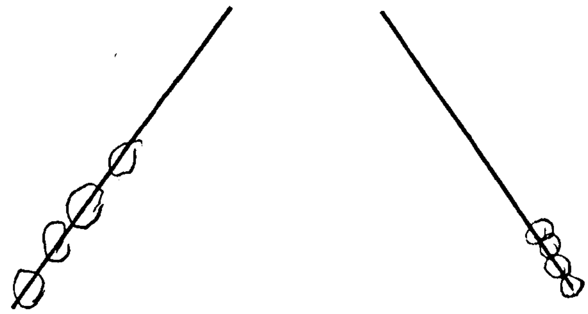


Anna, 8.8

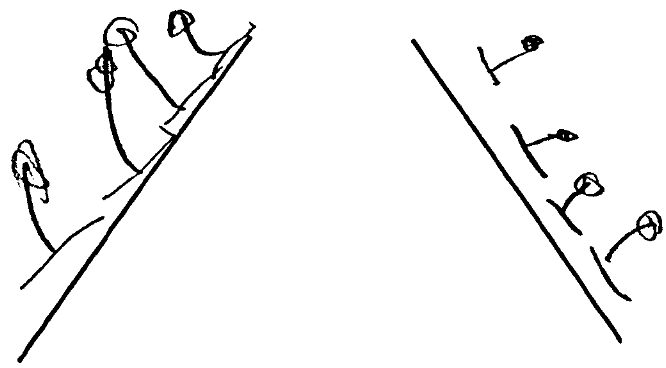


David, 8.0

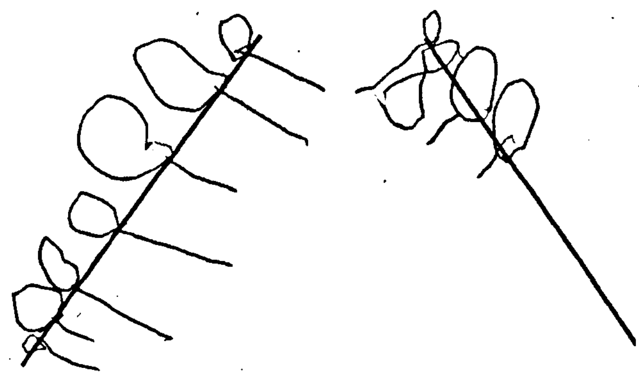
Jake, 3.8
(Unrepresentational drawing,
non-linear trees)



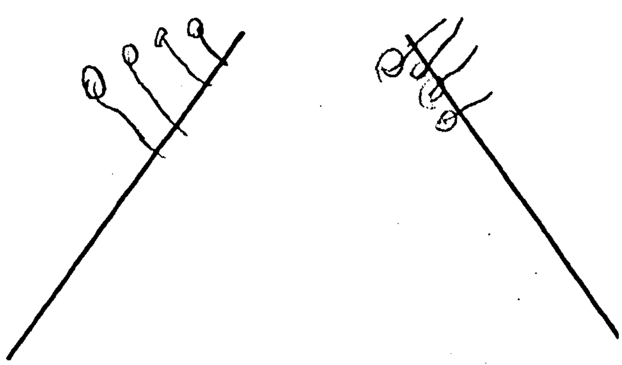
Heather, 4.4
(No road anchorage & tilted trees)

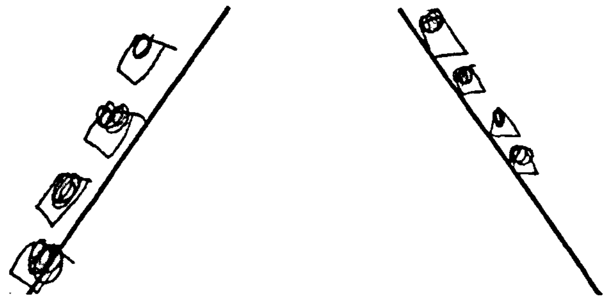


Sara, 3.7
(Incomplete anchorage & tilted trees)

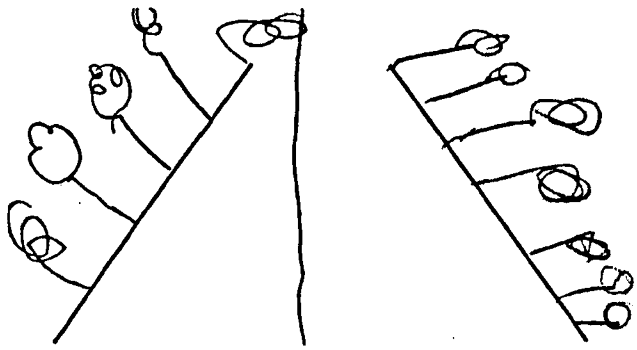


Elsa, 4.3
(Uncoordinated views & tilted trees)

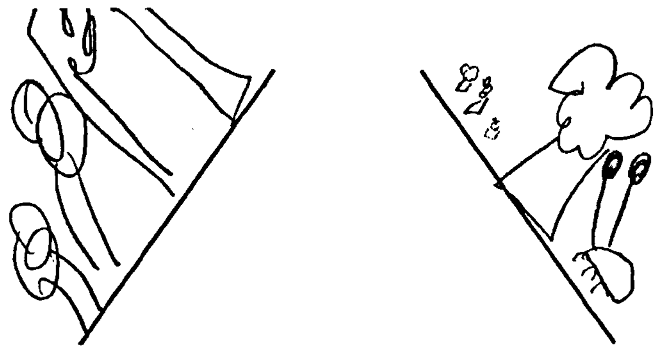




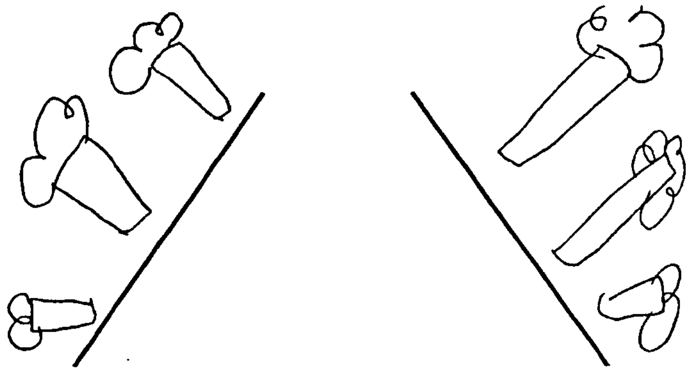
Amy, 5.8
(Alignment and anchorage parallel to road)



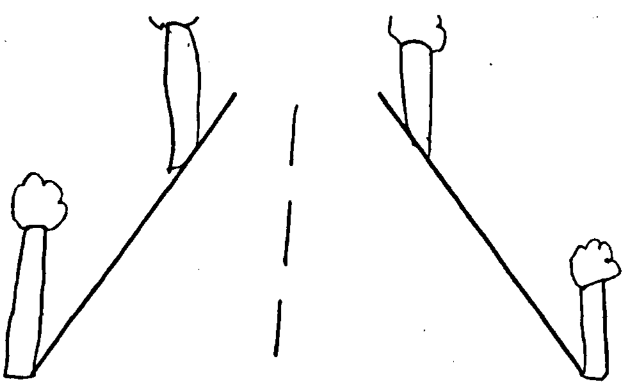
Daniel, 5.4
(Road anchorage & tilted trees)



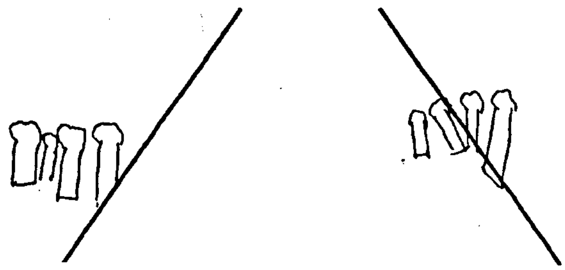
Luke, 5.9
(Road anchorage & tilted trees)



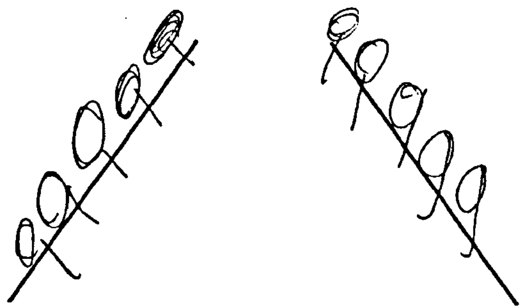
Emma, 6.2
No anchorage & tilted trees



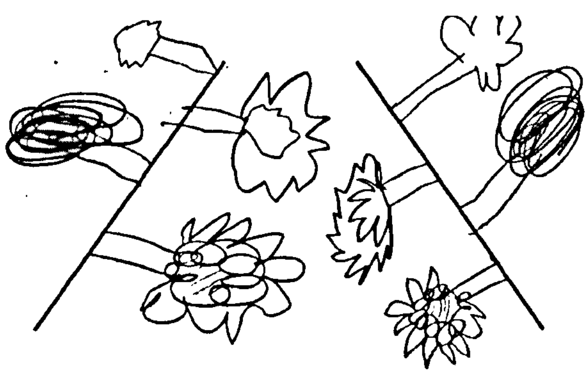
Lindsey 6.2
(Road anchorage & v. trees)



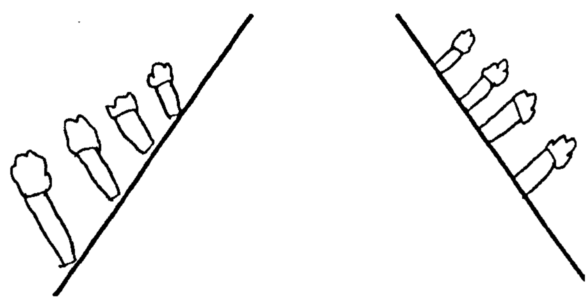
Katrina, 6.9
(Mixed anchorage & orienations)



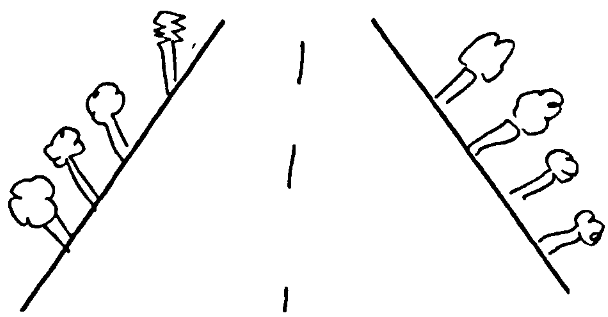
Michele, 7.0
(Incomplete anchorage
& mixed orientations)



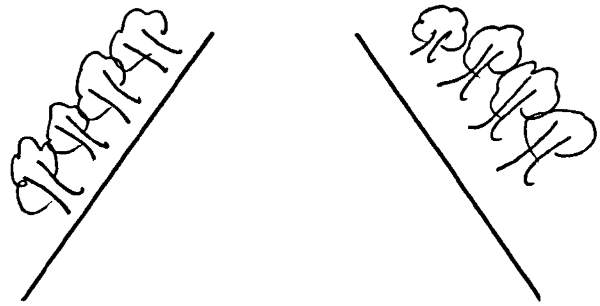
Jody, 7.0
(Road anchorage & tilted trees with
mixed views)



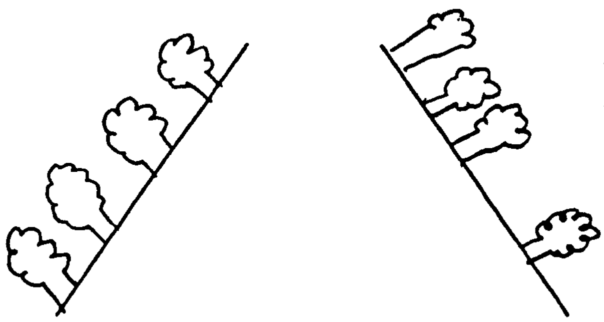
Tamara, 6.7
(Road anchorage & tilted trees)



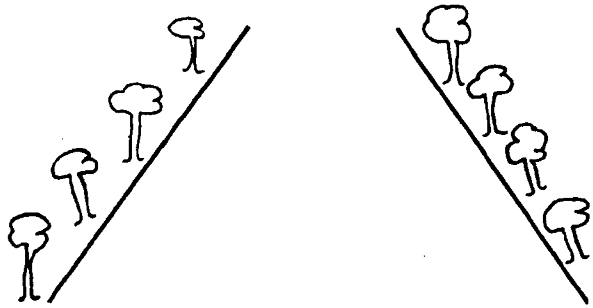
Jill, 7.3
(Road anchorage & mixed orientations)



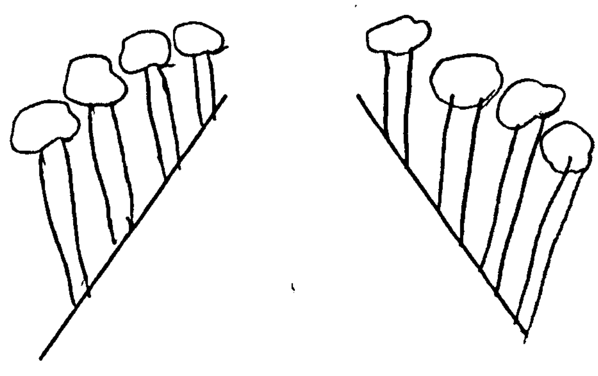
Heather, 8.5
(No anchorage & tilted trees)



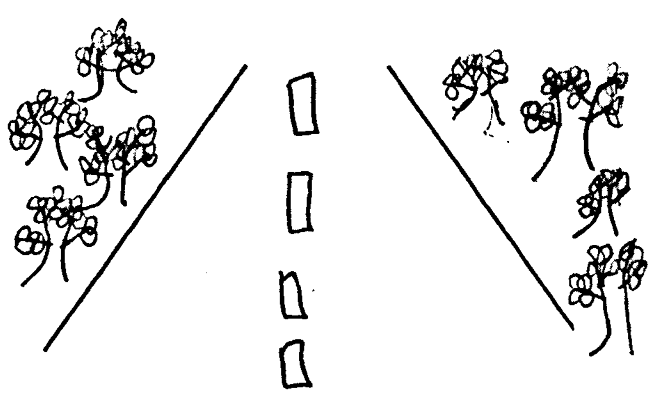
Nicola, 8.7
(Road anchorage & tilted trees)



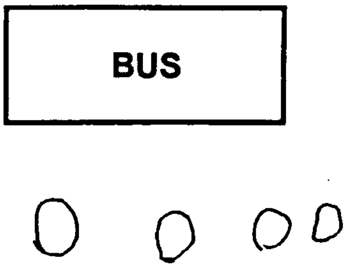
David, 8.0
(No anchorage & mixed orientations)



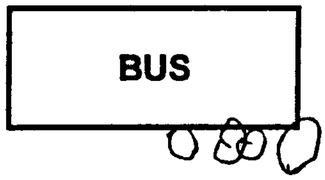
Anna, 8.8
(Road anchorage & mixed orientations)



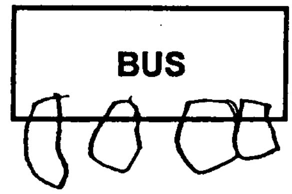
Ewan, 8.3
(No anchorage & v. trees)



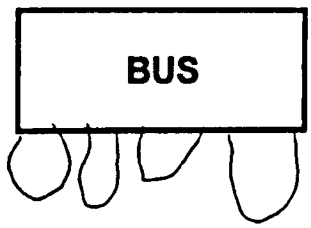
Jake, 3.8 (Segregation without anchorage)



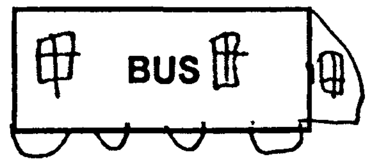
Elsa, 4.3 (Attachment)



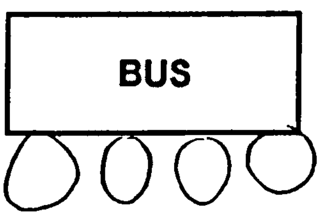
Claire, 4.2 (Overlap)



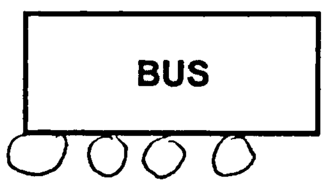
Louise, 4.5
(P. occlusion & segregation)



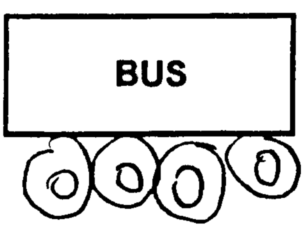
Ewan, 8.3
(P. occlusion & segregation)



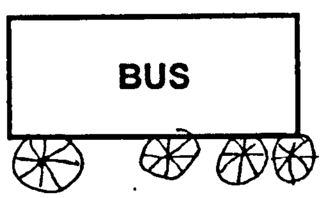
Amy, 5.8
(Attachment & segregation)



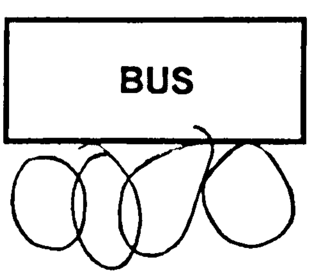
Lauren, 6.7
(Attachment & segregation)



Lorna, 5.9
(Attachment)

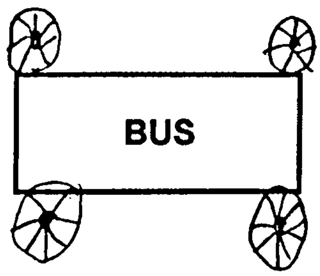


Jill, 7.3
(Attachment& segregation)

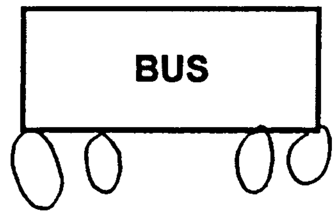


Natalie, 5.3
(Attachment & overlap)

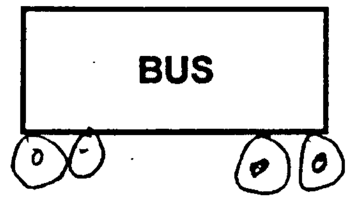
Paired Compositions



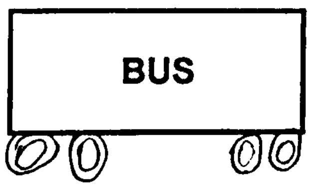
Lindsey, 6.2
(Folding over)



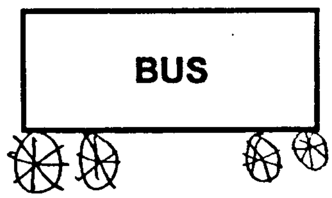
Allan, 8.6
(Attachment & segregation)



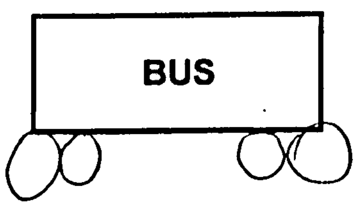
Claire, 7.4
(Attachment & segregation)



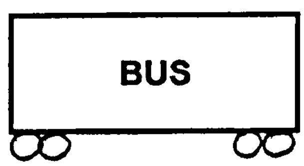
Louise, 5.3
(Attachment & segregation)



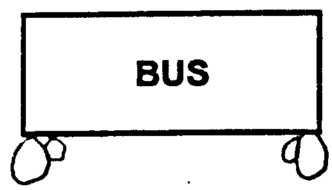
Gillian, 7.1
(Attachment & segregation)



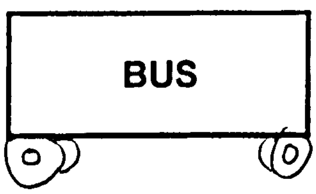
Megan, 5.6
(Attachments)



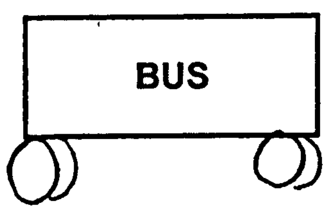
Anna, 8.8
(Attachments)



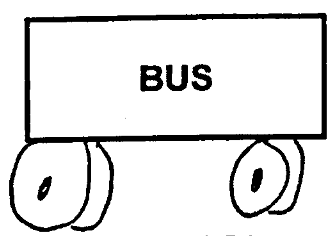
Lindsey, 7.9
(Attachments)



Kara, 5.6
(P. occlusion)



Nicola, 8.7
(P. occlusion)



Naomi, 7.9
(P. occlusion)

APPENDIX C

Table 1 Levels of significance for change of performance across pairs of tasks for the total sample

	Structurally separate scenes						Structurally integrated scenes		
	Bus	Bottles	Apples	Balls (MO)	Sunset	Balls (FO)	Boat	Houses	Vase
Bus									
Bottles	*NS*								
Apples	$<.001^{r}$	$<.01^{r}$							
Balls (MO)	$<.001^{r}$	$<.001^{r}$	$<.001^{r}$						
Sunset	$<.001^{r}$	$<.001^{r}$	$<.01^{r}$	$<.01^{r}$					
Balls (FO)	$<.001^{r}$	$<.001^{r}$	$<.01^{r}$	$<.01^{r}$	*NS*				
Boat	$<.001^{r}$	$<.05^{r}$	*NS*	$<.001^{c}$	$<.001^{c}$	$<.001^{c}$			
Houses	$<.001^{r}$	$<.001^{r}$	$<.05^{r}$	*NS*	$<.001^{c}$	$<.001^{c}$	$<.05^{r}$		
Vase	$<.001^{r}$	$<.001^{r}$	$<.001^{r}$	$<.01^{r}$	*NS*	*NS*	$<.001^{r}$	$<.001^{r}$	

Note *The tasks have been ordered on the basis of children's performance from the most to the least difficult within the two types of scene structures. In any cell, children's performance in the task nominated at the column is compared with performance in each task at the row. When among the inconsistent responders, the number of children who performed better in the task at the column was significantly greater than those who performed better in the task at the row, there is a 'c' next to the p value; contrary, there is an 'r'. The same denotation follows in subsequent tables.*

The sample size of each cell is 110, except with those cells involving comparisons with Bottles task where it is 109 due to a four year old who did not produce any data in this task.

Table 2 Levels of significance for change of performance of 4 year old children across pairs of tasks (N = 30)

	Structurally separate scenes						Structurally integrated scenes		
	Bus	Bottles	Apples	Balls (MO)	Sunset	Balls (FO)	Boat	Houses	Vase
Bus									
Bottles	*NS*								
Apples	*NS*	*NS*							
Balls (MO)	$<.05^{r}$	$<.05^{r}$	*NS*						
Sunset	$<.01^{r}$	$<.01^{r}$	$<.05^{r}$	*NS*					
Balls (FO)	$<.001^{r}$	$<.001^{r}$	$<.05^{r}$	*NS*	*NS*				
Boat	$<.01^{r}$	<.01	*NS*	*NS*	*NS*	*NS*			
Houses	*NS*	*NS*	*NS*	*NS*	*NS*	*NS*	*NS*		
Vase	$<.001^{r}$	$<.001^{r}$	$<.001^{r}$	$<.001^{r}$	*NS*	*NS*	$<.05^{r}$	$<.001^{r}$	

Note — *The sample size of cells involving comparisons with bottles task is 29 due to a child who did not produce any data in this task.*

Table 3 Levels of significance for change of performance of 5½ year old children across pairs of tasks (N = 25)

	Structurally separate scenes						Structurally integrated scenes		
	Bus	Bottles	Apples	Balls (MO)	Sunset	Balls (FO)	Boat	Houses	Vase
Bus									
Bottles	*NS*								
Apples	*NS*	$<.01^{r}$							
Balls (MO)	$<.001^{r}$	$<.001^{r}$	*NS*						
Sunset	$<.001^{r}$	$<.001^{r}$	$<.01^{r}$	$<.05^{r}$					
Balls (FO)	$<.001^{r}$	$<.001^{r}$	$<.01^{r}$	*NS*	*NS*				
Boat	*NS*	$<.05^{r}$	*NS*	$<.05^{c}$	$<.001^{c}$	$<.001^{c}$			
Houses	*NS*	$<.01^{r}$	*NS*	*NS*	$<.001^{c}$	$<.01^{c}$	*NS*		
Vase	$<.001^{r}$	$<.001^{r}$	$<.001^{r}$	*NS*	*NS*	*NS*	$<.001^{r}$	$<.01^{r}$	

Table 4 Levels of significance for change of performance of 7 year old children across pairs of tasks (N = 28)

	Structurally separate scenes						Structurally integrated scenes		
	Bus	Bottles	Apples	Balls (MO)	Sunset	Balls (FO)	Boat	Houses	Vase
Bus									
Bottles	*NS*								
Apples	$<.01^{r}$	*NS*							
Balls (MO)	$<.001^{r}$	$<.001^{r}$	$<.05^{r}$						
Sunset	$<.001^{r}$	$<.001^{r}$	$<.01^{r}$	*NS*					
Balls (FO)	$<.001^{r}$	$<.001^{r}$	$<.01^{r}$	*NS*	*NS*				
Boat	*NS*	*NS*	$<.05^{r}$	$<.001^{c}$	$<.001^{c}$	$<.001^{c}$			
Houses	$<.001^{r}$	$<.01^{r}$	*NS*	*NS*	*NS*	$<.05^{c}$	$<.01^{r}$		
Vase	$<.001^{r}$	$<.05^{r}$	*NS*	*NS*	*NS*	*NS*	$<.01^{r}$	*NS*	

Table 5 Levels of significance for change of performance of 8 year old children across pairs of tasks (N = 27)

	Structurally separate scenes						Structurally integrated scenes		
	Bus	Bottles	Apples	Balls (MO)	Sunset	Balls (FO)	Boat	Houses	Vase
Bus									
Bottles	*NS*								
Apples	*NS*	*NS*							
Balls (MO)	$<.001^{r}$	$<.05^{r}$	$<.05^{r}$						
Sunset	$<.001^{r}$	$<.01^{r}$	$<.01^{r}$	*NS*					
Balls (FO)	$<.001^{r}$	$<.01^{r}$	$<.01^{r}$	*NS*	*NS*				
Boat	*NS*	*NS*	*NS*	*NS*	$<.05^{c}$	$<.05^{c}$			
Houses	$<.001^{r}$	$<.01^{r}$	$<.05^{r}$	*NS*	*NS*	*NS*	*NS*		
Vase	$<.001^{r}$	$<.05^{r}$	$<.05^{r}$	*NS*	*NS*	*NS*	*NS*	*NS*	

APPENDIX D

General notes for Tables 1-5:

1. GHDT	Goodenough-Harris Drawing Test
2. Verbal scale	Vocab = Vocabulary Infor = Information Com/Sim = Comprehension subtest of WPPSI Similarities subtest of WISC-III^{UK} Verbal Score = Sum of raw scores from Verbal subtests VIQ = Verbal IQ
3. Performance scale	P.Com = Picture Completion BDesig = Block Design M/P.A = Mazes subtest of WPPSI Picture Arrangement subtest of WISC-III^{UK} Perform. Score = Sum of raw scores from Performance subtests PIQ = Performance IQ
4. *P* values	$^{a}p<.05$, $^{b}p<.01$, $^{c}p<.001$, $^{(a)}$ marginal failure, i.e. $p = .055$.
5. Tables N^{x} (a)- N^{x} (b)	Tables with the letter *a* following the relevant number present coefficients calculated on raw scores from criterion measures. Tables with the letter *b* following the relevant number report coefficients calculated using standard scores (IQ).
6. Tables 1a-b	The correlation coefficients for Table 1 are partial correlations, controlling for age.

Table 1a Correlation coefficients between scores from experimental tasks and standard measures of abilities for the total sample

			GHDT			WECHSLER SCALES								
						Verbal Scale				Performance Scale				
Experimental Tasks	Range	N	Man Scale	Wom. Scale	Self Scale	Vocab	Infor	CmSim	Verbal Score	P.Com	BDesig	M/P.A	Perform Score	Total Score
A. Point scores														
Occlusion (OPS)	0-9	110	.53[c]	.38[c]	.42[c]	.22[a]	.21[a]	.25[b]	.26[b]	.37[c]	.31[c]	.31[c]	.41[c]	.41[c]
Axial (APS)	0-5	110	.37[c]	.41[c]	.39[c]	.27[b]	.19[(a)]	.16	.25[b]	.24[a]	.46[c]	.26[b]	.44[c]	.44[c]
Composite (CPS)	0-14	110	.59[c]	.50[c]	.52[c]	.30[c]	.25[b]	.27[b]	.33[c]	.40[c]	.47[c]	.37[c]	.53[c]	.54[c]
B. Structural scores														
Occlusion (OSS)	0-27	110	.52[c]	.40[c]	.44[c]	.17	.12	.19[a]	.19[a]	.25[b]	.26[b]	.25[b]	.33[c]	.32[c]
Axial (ASS)	0-20	110	.45[c]	.40[c]	.40[c]	.28[b]	.24[b]	.22[a]	.29[b]	.29[b]	.28[b]	.28[b]	.36[c]	.40[c]

Table 1b Correlation coefficients between scores from experimental tasks and standard measures of abilities for the total sample

			GHDT			WECHSLER SCALES								
						Verbal Scale				Performance Scale				
Experimental Tasks	Range	N	Man Scale	Wom. Scale	Self Scale	Vocab	Infor	CmSim	VIQ	P.Com	BDesig	M/P.A	PIQ	FSIQ
A. Point scores														
Occlusion (OPS)	0-9	110	.48[c]	.42[c]	.44[c]	.16	.03	.14	.13	.26[b]	.33[c]	.26[b]	.35[c]	.29[b]
Axial (APS)	0-5	110	.33[c]	.43[c]	.41[c]	.25[b]	.15	.14	.21[a]	.20[a]	.38[c]	.22[a]	.34[c]	.34[c]
Composite (CPS)	0-14	110	.52[c]	.53[c]	.53[c]	.24[b]	.10	.17	.20[a]	.29[b]	.44[c]	.30[c]	.44[c]	.39[c]
B. Structural scores														
Occlusion (OSS)	0-27	110	.48[c]	.42[c]	.43[c]	.10	.02	.09	.08	.14	.33[c]	.19[a]	.28[b]	.22[b]
Axial (ASS)	0-20	110	.50[c]	.51[c]	.47[c]	.22[b]	.10	.07	.15	.17	.29[b]	.21[a]	.29[b]	.27[b]

Table 2a Correlation coefficients between scores from experimental tasks and standard measures of abilities for 4 year olds

						WECHSLER SCALES								
			GHDT			Verbal Scale				Performance Scale				
Experimental Tasks	Range	N	Man Scale	Wom. Scale	Self Scale	Vocab	Infor	CmSim	Verbal Score	P.Com	BDesig	M/P.A	Perform Score	Total Score
A. Point scores														
Occlusion (OPS)	0-9	30	.56[c]	.61[c]	.55[b]	-.21	.16	.11	-.02	.35	.57[c]	.51[b]	.61[c]	.45[a]
Axial (APS)	0-5	30	.08	.17	.05	.08	.14	-.07	.06	.13	.18	.01	.12	.13
Composite (CPS)	0-14	30	.56[c]	.65[c]	.54[b]	-.17	.21	.07	.01	.38[a]	.61[c]	.48[b]	.62[c]	.47[b]
B. Structural scores														
Occlusion (OSS)	0-27	30	.54[b]	.57[c]	.47[b]	-.20	.03	.07	-.07	.19	.49[b]	.41[a]	.48[b]	.31
Axial (ASS)	0-20	30	.38[a]	.47[b]	.35[(a)]	.04	.15	.01	.07	.28	.32	.41[a]	.45[a]	.39[a]

Table 2b Correlation coefficients between scores from experimental tasks and standard measures of abilities for 4 year olds

						WECHSLER SCALES								
			GHDT			Verbal Scale				Performance Scale				
Experimental Tasks	Range	N	Man Scale	Wom. Scale	Self Scale	Vocab	Infor	CmSim	VIQ	P.Com	BDesig	M/P.A	PIQ	FSIQ
A. Point scores														
Occlusion (OPS)	0-9	30	.37[a]	.37[a]	.45[a]	-.29	.02	-.03	-.15	.18	.44[a]	.40[a]	.47[b]	.22
Axial (APS)	0-5	30	.10	.16	.05	.03	.04	-.14	-.02	.10	.02	-.05	.02	.01
Composite (CPS)	0-14	30	.39[a]	.41[a]	.44[a]	-.26	.04	-.09	-.15	.21	.42	.35	.45[b]	.21
B. Structural scores														
Occlusion (OSS)	0-27	30	.35	.32	.33	-.27	-.07	-.03	-.17	.04	.39[a]	.30	.35[(a)]	.11
Axial (ASS)	0-20	30	.43[a]	.49[b]	.34	.00	.02	-.13	-.04	.17	.14	.32	.29	.17

Table 3a Correlation coefficients between scores from experimental tasks and standard measures of abilities for 5½ year olds

			GHDT			WECHSLER SCALES: Verbal Scale				WECHSLER SCALES: Performance Scale				
Experimental Tasks	Range	N	Man Scale	Wom Scale	Self Scale	Vocab	Infor	CmSim	Verbal Score	P.Com	BDesig	M/P.A	Perform Score	Total Score
A. Point scores														
Occlusion (OPS)	0-9	25	.64[c]	.48[a]	.45[a]	.27	.20	.25	.29	.04	.51[b]	.27	.38	.39[(a)]
Axial (APS)	0-5	25	.70[c]	.66[c]	.67[c]	.19	.30	.07	.19	.32	.26	.34	.41[a]	.34
Composite (CPS)	0-14	25	.77[c]	.63[c]	.61[c]	.28	.27	.22	.30	.16	.50[a]	.34	.45[a]	.44[a]
B. Structural scores														
Occlusion (OSS)	0-27	25	.66[c]	.52[b]	.49[a]	.19	.12	.15	.19	.08	.63[c]	.26	.43[a]	.35
Axial (ASS)	0-20	25	.62[c]	.54[b]	.62[b]	.18	.30	.05	.14	.23	.35	.15	.30	.24

Table 3b Correlation coefficients between scores from experimental tasks and standard measures of abilities for 5½ year olds

			GHDT			WECHSLER SCALES: Verbal Scale				WECHSLER SCALES: Performance Scale				
Experimental Tasks	Range		Man Scale	Wom Scale	Self Scale	Vocab	Infor	CmSim	VIQ	P.Com	BDesig	M/P.A	PIQ	FSIQ
A. Point scores														
Occlusion (OPS)	0-9	25	.66[c]	.50[a]	.47[a]	.22	.09	.17	.19	-.14	.27	.16	.16	.21
Axial (APS)	0-5	25	.67[c]	.63[c]	.63[c]	.09	.15	-.08	.05	.12	.05	.20	.20	.14
Composite (CPS)	0-14	25	.78[c]	.64[c]	.61[c]	.20	.13	.10	.16	-.07	.23	.21	.20	.22
B. Structural scores														
Occlusion (OSS)	0-27	25	.67[c]	.53[b]	.50[a]	.16	.04	.08	.10	-.08	.42[a]	.13	.23	.19
Axial (ASS)	0-20	25	.56[b]	.47[a]	.55[b]	.13	.23	.17	.05	.03	.22	.05	.14	.12

Table 4a Correlation coefficients between scores from experimental tasks and standard measures of abilities for 7 year olds

			GHDT			WECHSLER SCALES: Verbal Scale				Performance Scale				
Experimental Tasks	Range		Man Scale	Worn Scale	Self Scale	Vocab	Infor	CmSim	Verbal Score	P.Com	BDesig	M/P.A	Perform Score	Total Score
A. Point scores														
Occlusion (OPS)	0-9	28	.69^{c}	.62^{c}	.48^{b}	.20	.20	.44^{a}	.33	.64^{c}	.43^{a}	.39^{a}	.53^{b}	.50^{b}
Axial (APS)	0-5	28	.48^{a}	.46^{a}	.30	.31	.63^{c}	.45^{a}	.48^{b}	.34	.49^{b}	.50^{b}	.56^{b}	.63^{c}
Composite (CPS)	0-14	28	.73^{c}	.68^{c}	.50^{b}	.30	.46^{a}	.54^{b}	.48^{b}	.63^{c}	.55^{b}	.53^{b}	.67^{c}	.67^{c}
B. Structural scores														
Occlusion (OSS)	0-27	28	.70^{c}	.58^{c}	.53^{b}	.20	.22	.47^{b}	.35	.60^{c}	.40^{a}	.36	.49^{b}	.51^{b}
Axial (ASS)	0-20	28	.56^{b}	.43^{a}	.28	.33	.56^{b}	.46^{a}	.49^{b}	.32	.52^{b}	.43^{a}	.54^{b}	.60^{c}

Table 4b Correlation coefficients between scores from experimental tasks and standard measures of abilities for 7 year olds

			GHDT			WECHSLER SCALES: Verbal Scale				Performance Scale				
Experimental Tasks	Range		Man Scale	Worn Scale	Self Scale	Vocab	Infor	CmSim	VIQ	P.Com	BDesig	M/P.A	PIQ	FSIQ
A. Point scores														
Occlusion (OPS)	0-9	28	.62^{c}	.55^{b}	.40^{a}	.20	.14	.39^{a}	.30	.65^{c}	.37$^{(a)}$	.40^{a}	.52^{b}	.49^{b}
Axial (APS)	0-5	28	.56^{b}	.57^{b}	.42^{a}	.35	.62^{c}	.44^{a}	.56^{b}	.44^{a}	.55^{b}	.58^{c}	.64^{c}	.70^{c}
Composite (CPS)	0-14	28	.73^{c}	.68^{a}	.50^{b}	.32	.41^{a}	.50^{b}	.50^{b}	.69^{c}	.54^{b}	.58^{c}	.69^{c}	.70^{c}
B. Structural scores														
Occlusion (OSS)	0-27	28	.61^{c}	.49^{b}	.43^{a}	.20	.15	.45^{a}	.33	.60^{c}	.34	.38^{a}	.48^{b}	.48^{b}
Axial (ASS)	0-20	28	.65^{c}	.54^{b}	.40^{a}	-.37$^{(a)}$	.57^{c}	.49^{b}	.57^{c}	.44^{a}	.59^{c}	.48^{b}	.61^{c}	.69^{c}

Table 5a Correlation coefficients between scores from experimental tasks and standard measures of abilities for 8 year olds

			GHDT			WECHSLER SCALES: Verbal Scale				WECHSLER SCALES: Performance Scale				
Experimental Tasks	Range	N	Man Scale	Wom Scale	Self Scale	Vocab	Infor	CmSim	Verbal Score	P.Com	BDesig	M/P.A	Perform Score	Total Score
A. Point scores														
Occlusion (OPS)	0-9	27	.29	-.06	.25	.55[b]	.28	.23	.44[a]	.31	.37[(a)]	.17	.36	.44[a]
Axial (APS)	0-5	27	.26	.30	.42[a]	.42[a]	.14	.34	.39[a]	.28	.63[c]	.17	.52[b]	.53[b]
Composite (CPS)	0-14	27	.35[(a)]	.16	.44[a]	.62[c]	.27	.37[(a)]	.54[b]	.38[a]	.65[c]	.22	.57[b]	.63[c]
B. Structural scores														
Occlusion (OSS)	0-27	27	.31	.04	.29	.50[b]	.29	.21	.42[a]	.18	.41[a]	.08	.33	.40
Axial (ASS)	0-20	27	.25	.27	.39[a]	.30	.08	.38[a]	.33	.21	.40[a]	.04	.31	.36

Table 5b Correlation coefficients between scores from experimental tasks and standard measures of abilities for 8 year olds

			GHDT			WECHSLER SCALES: Verbal Scale				WECHSLER SCALES: Performance Scale				
Experimental Tasks	Range	N	Man Scale	Wom Scale	Self Scale	Vocab	Infor	CmSim	VIQ	P.Com	BDesig	M/P.A	PIQ	FSIQ
A. Point scores														
Occlusion (OPS)	0-9	27	.27	.03	.23	.53[b]	.23	.26	.40[a]	.19	.34	.11	.29	.40[a]
Axial (APS)	0-5	27	.22	.32	.47[b]	.39[a]	.06	.28	.30	.12	.56[b]	.07	.34	.39[a]
Composite (CPS)	0-14	27	.32	.19	.45[b]	.59[c]	.19	.35	.45[a]	.20	.58[b]	.12	.41[a]	.50[b]
B. Structural scores														
Occlusion (OSS)	0-27	27	.27	.04	.21	.48[a]	.26	.24	.38[a]	.07	.39[a]	.02	.23	.36
Axial (ASS)	0-20	27	-.22	.31	.46[b]	.28	.01	.34	.26	.06	.34	.05	.17	.26

Table 6 F values and level of significance from the analysis of experimental spatial scores by age (4) × IQ level (2)

	WISC-III			GH-Drawing Test		
	FSIQ	VIQ	PIQ	Man SS	Woman SS	Self SS
A. CPS						
Age	$F_{3,102} = 95.66^c$	$F_{3,102} = 89.55^c$	$F_{3,102} = 95.24^c$	$F_{3,102} = 116.22^c$	$F_{3,102} = 111.99^c$	$F_{3,102} = 100.74^c$
IQ level	$F_{1,102} = 14.08^c$	$F_{1,102} = 3.52^{(a)}$	$F_{1,102} = 12.95^c$	$F_{1,102} = 32.69^c$	$F_{1,102} = 32.46^c$	$F_{1,102} = 22.38^c$
Age × IQ level				$F_{3,102} = 3.00^a$		
B. OPS						
Age	$F_{3,102} = 66.89^c$	$F_{3,102} = 66.38^c$	$F_{3,102} = 68.80^c$	$F_{3,102} = 78.08^c$	$F_{3,102} = 78.85^c$	$F_{3,102} = 68.54^c$
IQ level	$F_{1,102} = 6.91^b$	$F_{1,102} = .55$	$F_{1,102} = 6.20^b$	$F_{1,102} = 20.48^c$	$F_{1,102} = 22.94^c$	$F_{1,102} = 11.15^c$
Age × IQ level					$F_{3,102} = 2.87^a$	
C. APS						
Age	$F_{3,102} = 44.62^c$	$F_{3,102} = 39.78^c$	$F_{3,102} = 42.14^c$	$F_{3,102} = 44.96^c$	$F_{3,102} = 44.11^c$	$F_{3,102} = 46.07^c$
IQ level	$F_{1,102} = 11.46^c$	$F_{1,102} = 6.94^b$	$F_{1,102} = 10.73^c$	$F_{1,102} = 14.38^c$	$F_{1,102} = 12.62^c$	$F_{1,102} = 16.79^c$
Age × IQ level	$F_{3,102} = 2.97^a$					
D. OSS						
Age	$F_{3,102} = 74.48^c$	$F_{3,102} = 79.30^c$	$F_{3,102} = 77.30^c$	$F_{3,102} = 87.80^c$	$F_{3,102} = 86.97^c$	$F_{3,102} = 77.62^c$
IQ level	$F_{1,102} = 3.61^{(a)}$	$F_{1,102} = .50$	$F_{1,102} = 4.40^a$	$F_{1,102} = 14.84^c$	$F_{1,102} = 16.21^c$	$F_{1,102} = 7.18^b$
Age × IQ level		.			$F_{3,102} = 2.90^a$	
E. ASS						
Age	$F_{3,102} = 66.79^c$	$F_{3,102} = 61.00^c$	$F_{3,102} = 65.80^c$	$F_{3,102} = 68.03^c$	$F_{3,102} = 67.84^c$	$F_{3,102} = 68.53^c$
IQ level	$F_{1,102} = 10.04^b$	$F_{1,102} = 5.38^a$	$F_{1,102} = 10.00^b$	$F_{1,102} = 8.44^b$	$F_{1,102} = 14.26^c$	$F_{1,102} = 19.20^c$

CPS = Composite Point Score, OPS = Occlusion Point Score, APS = Axial Point Score, OSS = Occlusion Structural Score, ASS = Axial Structural Score
FSIQ = Full scale IQ, VIQ = Verbal IQ, PIQ = Performance IQ, Man/Woman/Self – SS = Man/Woman/Self – Standard Scores
$^a p<.05$, $^b p<.01$, $^c p<.001$, $^{(a)}$ marginal failure e.g. $p = .06$